The effects of changes in the referential problem space of infants and toddlers (homo sapiens): implications for cross-species comparisons

Article (Accepted Version)

The Effects of Changes in the Referential Problem Space of Infants and Toddlers (*Homo sapiens*): Implications for cross-species comparisons

Hannah Clark[1], Zoe M. Flack[2] and David A. Leavens[1,*,]

[1]School of Psychology, University of Sussex, Falmer, East Sussex, BN1 9QH

[2]School of Applied Social Science, University of Brighton, Brighton, East Sussex, BN1 9PH

*Corresponding author: David A. Leavens, School of Psychology, University of Sussex, Falmer, East Sussex, BN1 9QH; davidl@sussex.ac.uk

Note: Accepted version of paper to appear in the Journal of Comparative Psychology

Abstract

Recent reviews have highlighted the tendency in the comparative literature to make claims about species' relative evolutionarily adaptive histories based on studies comparing different species tested with procedurally and methodologically different protocols. One particularly contentious area is the use of the Object Choice Task (OCT), used to measure an individual's ability to use referential cues, which is a core attribute of joint attention. We tested human children with versions of the OCT that have been previously used with dogs and nonhuman primates to see if manipulating the set-up would lead to behavioral changes. In Study 1, we compared the responses of 18-month-olds and 36-month-olds when tested with and without a barrier. The presence of a barrier between the child and the reward did not suppress performance but did elicit more communicative behavior. Moreover, the barrier had a greater facilitating effect on the younger children, who displayed more communicative behavior in comparison with older children, who more frequently reached through the barrier in acts of direct prehension. In Study 2, we compared the behavior of 36-month-olds when the reward was within reaching distance (proximal) and when it was out of reach (distal). The children used index-finger points significantly more in the distal condition and grabbed more in the proximal condition, showing that they were making spatial judgements about the accessibility of the reward rather than just grabbing *per se*. We discuss the implications of these within-species differences in behavioral responses for cross-species comparisons.

Key words: Object choice task; Use of experimenter-given cues; Referential problem space; Experimental methods; Social cognition.

Leavens, Bard and Hopkins (2019) discussed the tendency in comparative psychology literature to attribute differences in performance by apes and humans on socio-cognitive tasks to discontinuity in the hominid lineage that has endowed humans with species-unique cognitive skills in understanding others. Such theories, they argue, are not well justified by the empirical evidence that claims to support them due to a number of procedural and methodological confounds which arise from failing to match key selection, life history, and procedural variables across groups. In recent decades, there has been a proliferation of such rich interpretations of human responses to assays of social understanding, with many published claims for uniquely human, hypothetical psychological processes, typically with no consideration of plausible alternative explanations grounded in such confounded factors as level of pre-experimental preparation, life history stages, participant selection protocols, etc. (see also Leavens 2012a; Leavens, 2018). One contentious area in this regard relates to studies of joint attentional skills, in particular, the comprehension of declarative gestural cues (gestures whose motive is to inform or share information), the extent to which other species possess such skills, and their role in the emergence of verbal communication in humans (e.g. Corballis (1999). In attempts to measure such skills, the Object Choice Task (OCT) is frequently employed.

The OCT is used to assess an individual's ability to comprehend human gestural cues and involves an experimenter presenting a deictic cue to indicate to the subject in which of two or three containers a reward has previously been hidden (Anderson, Sallaberry & Barbier, 1995). Results of OCT studies have been used as the bases for theories pertaining to the evolutionary roots of social cognition in a number of species, especially humans (*Homo sapiens*; e.g, Povinelli, Bierschwale, & Čech, 1999; Tomasello, Call, & Gluckman, 1997), nonhuman primates (Primates; e.g., Anderson et al., 1995; Povinelli et al., 1999; Tomasello et al., 1997), and domestic dogs (*Canis familiaris*; e.g., Udell, Giglio, & Wynne, 2008). For example, Hare and Tomasello (2005), on the basis of domestic dogs' consistently adept performance, argued for a theory of convergent evolution between dogs and humans, in which the former have developed specialized socio-cognitive skills to comprehend human gestural cues as a result of centuries of artificial selection during domestication. The Cultural Intelligence hypothesis (Herrmann et al., 2007; van Schaik & Burkart, 2011), based, in part, on nonhuman primates' generally poor performance on the OCT, states that humans have developed species-specific socio-cognitive skills in order to facilitate cultural group living, and, as such, the comprehension of human gestural cues is a human-unique ability within the primates. Bräuer, Kaminski, Riedel, Call and Tomasello (2006) proposed that differences in the performance levels between dogs and nonhuman primates are due to species-specific specialisations where dogs have been selected for specialized social abilities, which thus enable them to follow human gestural cues, whereas apes' foraging behavior has led to increased physical abilities, which explains their ability to use physical but not social cues on the OCT. Leavens, Hopkins, and Bard (2005) posited a Referential Problem Space to account for the very sparse observations of pointing in wild chimpanzees, compared to captive apes, who very frequently point; they suggested that pointing emerges in circumstances in which pointers are forced to rely on the manipulation of others to obtain objects of interest, a situation that characterises both human infants, who are often restrained, and captive apes, whose prehension is often blocked by cage mesh.

Recent meta-analyses (e.g., Clark, Elsherif & Leavens, 2019; Krause, Udell, Leavens, & Skopos, 2018; Lyn, 2010; Mulcahy & Hedge, 2012), however, have identified procedural and methodological differences in the testing protocols used with different taxonomic groups on the OCT that may provide more comprehensive explanations of the performance differences found than theories that attribute them to phylogenetic causes. First, human infants' abilities to comprehend pointing gestures develop over the first year of life

(Butterworth, 2001; Butterworth & Morrisette, 1996) in an environment rich in human interactions. It is this developmental process that Bard and Leavens (2014) argued is essential to consider when making cross-species comparisons in socio-cognitive tasks, and they highlight the contemporary prevalence of basing phylogenetic theories on the performance of subjects unmatched for developmental experience. Indeed, Lyn (2010) argued that pre-experimental exposure to humans can differentially affect an individual's performance on the OCT, and evidence is accumulating to support this argument. Lyn, Russell and Hopkins (2010) found enculturated nonhuman primates to be successful at following human gestural cues and a growing body of work shows that domestic dogs with less exposure to humans perform significantly worse than their pet dog counterparts (e.g., D'Aniello et al., 2017; Duranton & Gaunet, 2016; Lazarowski & Dorman, 2015; Udell, Dorey & Wynne, 2008), whereas other canids raised in environments rich in human interactions perform well on the task (Barrera & Bentosela, 2016; Udell, Dorey, Spencer & Wynne, 2012; Udell et al., 2008).

Leavens et al. (2019) argued that a further inconsistency in the testing protocols adopted with different species is that of the presence of a barrier between the subject and the baited container in the testing environment. Due to safety issues surrounding the testing of nonhuman primates, species from this taxonomic group are tested from outside their cages, therefore imposing a barrier between the subject and the experimenter and test apparatus. Working with individuals from other taxonomic groups, domestic dogs or human infants, for example, does not necessitate the use of such safety precautions, and, as such, there is an absence of this barrier in the testing environment with these species; this constitutes a confound between experimental protocol and species classification in a significant number of contemporary studies. In a review of 71 published nonhuman primate and dog OCT studies, Clark et al. (2019) found that 99% of nonhuman primates were tested with a barrier present in the testing environment, compared with less than 1% of dogs. They therefore argued that this inconsistency in the test set-ups used across different species represents an experimental confound that may affect individuals' performances. This renders conclusions that group differences are attributable to differences in evolutionary histories implausible.

In the one study to date which measured domestic dogs' abilities to follow a pointing cue with a barrier present, Kirchhofer et al. (2012) found that those tested with a barrier had significantly lower success rates than those tested without. Similarly, Udell et al. (2008) found that domestic dogs tested with a partial visual barrier performed significantly worse than those tested without when required to follow a tapping cue on an OCT. Using a parental survey approach, Kishimoto (2013) found that parents who more frequently reported moving small or age-inappropriate objects, such as beads or personal computers, to locations that were out of reach of their children also reported higher frequencies of imperative pointing by their children, suggesting that the Referential Problem Space posited by Leavens et al. (2005) facilitated pointing in his sample. That is, Kishimoto (2013) argued that parents created a Referential Problem Space which facilitated the use of imperative points by placing such items out of reach of the children. Taken together, these findings demonstrate the reduced validity of interpreting group differences as phylogenetic traces without regard to the systematically confounded differences in experimental set-ups being used with different taxonomic groups (Leavens et al., 2019). We are not aware of any study with human children on the OCT to date in which a barrier has been present in the testing environment, although human children, at least in Western populations, are well-habituated to conditions of restraint in car seats, feeding chairs, playpens, cots, and so on (Leavens et al., 2005).

In order to investigate the possible confounding effects of this systematic difference in experimental protocols administered to representatives of different species, we tested children with and without a barrier on an OCT. Children from 14 months of age have been shown to reliably follow pointing cues on the OCT (Behne, Lizkowski, Carpenter & Tomasello, 2012) and so we tested children aged 18 months and 36 months in order to ensure that any differences in performance or behavioral responses between the two conditions were as a result of our experimental manipulation, rather than the lack of emergence of these skills.

Study 1: Barrier vs. No Barrier

In Study 1, we looked at the effects of the imposition of a barrier, testing human children aged 18 months and 36 months on a within-subjects design, in which participants completed an OCT with and without a barrier present. To recreate as closely as possible the conditions in which nonhuman primates are tested, that is, from within a test cage, in the barrier condition children were tested from within a child's playpen, thus imposing a physical barrier between the participant (inside the enclosure) and the experimenter and testing apparatus (outside the enclosure).

Method

Participants

The study was approved by the Science and Technology Cross-Schools Research Ethics Committee (C-REC) at the University of Sussex. Participants were nineteen 18-month-olds (*M* = 18 mos 18 days, *range* = 18 mos 3 days – 18 mos 27 days) and twenty 36-month-olds (*M* = 36 mos 8 days, *range* = 33 mos 10days– 39 mos 0 days), comprised of 22 males and 17 females (18-month-olds: 11 males, 8 females; 36-month-olds: 11 males, 9 females). Participants were recruited from a participant database where parents had registered their interest in participating in developmental studies with their children, and from advertisements on social media sites. Parents gave informed consent for their children to participate. Data were collected between April and November 2016.

Procedure

On arrival at the testing suite, participants and their parents were given time to become familiar with the surroundings, with participants playing freely in the playroom and interacting with the experimenter. When parents judged their child to be settled and comfortable, the experimenter, participant, and parent moved to the testing room, where the experimenter demonstrated a “ball run” toy to the child, and then encouraged the child to play with the toy. The test room was set up as shown in Figure 1. The experimenter then informed the child that they were going to play a fun hiding game with the balls, and that if the child found the balls they could put them in the ball run. The experimenter then asked the child to sit on the playmat with their parent and explained that she would hide the ball under one of two cups, and then give the participant a “clue” to see if they could find it. In the *barrier* condition, a child’s playpen was set up, such that the playmat was inside the pen, and the ball run was outside, but accessible to the child. In the *no barrier* condition, the playmat and ball run were in the same positions, but the playpen was not in place. The experimenter hid the ball under one of two cups, behind a cardboard occluder, then made eye contact with the participant, asking “are you ready for your clue?” The experimenter then presented an ipsilateral, dynamic, index-finger pointing cue, whilst alternating her gaze between the container and the experimenter. A dynamic point, according to Miklósi and Soproni’s (2006) definition of the different point types used on the OCT, is one in which the pointing gesture is carried out in front of the participant and remains in place until the participant makes a choice. The distance between the experimenter’s fingertip and the container was approximately 5cm. The experimenter maintained this position until the participant made a choice. If the participant was unresponsive, the experimenter encouraged the participant to make a choice by giving verbal encouragement such as “can you find that ball?” If the participant failed to respond after approximately 2 minutes, or was fussy (for example, trying to get out of the playpen), then the trial was terminated, and the experimenter attempted to increase motivation by again demonstrating the ball in the ball run. If the participant made a correct choice, they were given the ball and encouraged to put it in the ball run. If the participant made an incorrect choice, the experimenter lifted the incorrect cup and said, for example, "oh no! It's not in that one! Let's see if it was in the other one!" and then lifted the correct cup, showed the child the ball, and said "Never mind! Let's hide it again!" Participants received 8 trials in the *barrier* condition and 8 trials in the *no barrier* condition. Order of administration was counterbalanced across participants, and in between conditions, participants left the test room with their parent and were engaged in another task, such as looking at wall stickers of animals. The baited container was on the right or left an equal number of times in each condition, and the order was counterbalanced, such that the reward was never on the same side for more than two consecutive trials.

Materials

The playpen used in the barrier condition was a Dream Baby Royal Converta 3-in-1 Playpen Gate, measuring 380 x 4 x 74cm (Rosyth Business Centre, 16 Cromarty Campus, Rosyth, Fife, KY11 2WX). Children and their parents sat on a playmat made up of 16 interlocking JSG Accessories Outdoor/ Indoor Protective Flooring Mats (JSG Accessories, Unit 6 Hughes Business Centre, Wilverley Road, BH23 3RU). The containers used to hide the reward were two white opaque plastic cups measuring 7.8 x 10cm. A John Lewis Junior Ball Run was used as the stimulus, measuring 52 x 56 x 47.5cm (John Lewis Partnership, 71 Victoria Street, London, SW1E 5NN). The occluder was a piece of brown cardboard measuring 65 x 80cm. All testing sessions were recorded on two Sony Handycam HDR-PJ410 video-cameras (Sony, 1-7-1 Konan Minato-ku, Tokyo, 108-0075 Japan).

Data Scoring

Test sessions were video-recorded and coded at a later date. For each trial, data were coded for whether or not the choice made was correct, latency of response (from maximum extension of the index-finger by the experimenter to the participant choosing a cup), type of response, the direction of the participant's gaze whilst giving the response, and whether the response was accompanied by a vocalisation. Response types were categorized according to the following scheme:

Index-finger point: The arm and index-finger are extended towards the referent, with the other fingers curled under the hand (Masataka, 2003).

Whole-hand point: An indicative gesture categorized by outstretched arm and extended fingers, which is not a direct attempt to obtain the container (Leavens & Hopkins, 1999).

Indicative gesture other than index-finger/ whole-hand point: Where a participant indicated a choice using a gesture other than an index-finger point or whole-hand point. An example of this is one participant "pointed" to the container with their foot.

Direct Grab: A response was categorized as a grab when the participant reached for and contacted the container with their hand or fully grasped the container.

Reach: An attempt to obtain the container, categorized by hand outstretched and fingers grasping. Reaches are distinguished from whole-hand points by the presence of repetitive flexion, which is absent from the latter. (Leavens & Hopkins, 1999).

Other: Responses other than those described above. An example of an *other* response is a child who used the parent's arm to indicate the choice.

Analyses

Participants were excluded from the analyses if they failed to complete at least four trials in each condition. This led to the exclusion of four 18-month-olds and one 36-month-old. There was no significant difference in the number of trials completed between 18-month-olds (Mdn = 15) and 36-month-olds (Mdn = 16), Z = -2.70, p = .458. Due to non-normal distribution of the data, non-parametric tests were used throughout the analyses and

Wilcoxon's signed-ranks test was used for the within-subjects comparisons.

Reliability

An independent coder who was blind to the purpose of the study coded 20% of the videos (five participants, with sixteen trials each, for a total of 80 trials). For correct choices, There was complete agreement between the two coders for whether participants chose the correct cup on each trial, Cohen's kappa, $\kappa = 1.00$, $p < .001$. There was excellent agreement for response latency, $r_s = 0.8769$, $p < .001$, and substantial agreement for the type of response elicited from participants on each trial response type, $\kappa = 0.66$, $p < .001$ (Landis & Koch, 1977). In the event of disagreement between coders on a specific trial, the original coding was maintained.

Results

Correct Choices

Eighteen-month-olds

The 18-month-olds, as a group, performed above chance both with a barrier (binomial test, $p < .001$) and without a barrier (binomial test, $p < .001$). There was no significant difference in the proportion of correct choices made between the *barrier* (*Mdn* = 1.00) and *no barrier* (*Mdn* = 1.00) conditions, $Z = -0.60$, $p = .552$, $r = -.04$. This shows that the barrier did not have an effect on the younger groups' ability to use the pointing cue to find the hidden reward.

Thirty-six-month-olds

The 36-month-olds also performed above chance as a group in both the *barrier* (binomial test, $p < .001$) and the *no barrier* (binomial test, $p < .001$) conditions. There was no significant difference in the proportion of correct trials between the *barrier* (*Mdn* = 1.00) and the *no barrier* (*Mdn* = 1.00) conditions, $Z = -1.36$, $p = .175$, $r = -.08$. This shows the older children were also able to use the pointing cue despite the presence of a barrier.

Age comparisons

There was no significant difference in proportion of correct trials between the 18-month-olds (*Mdn* = 1.00) and the 36-month olds (*Mdn* = 1.00) in the *barrier* condition, Mann-Whitney $U = 112.5$, $p = .302$, $r = -.08$, nor in the *no barrier* condition (18-month-old *Mdn* = 1.00, 36-month-old *Mdn* = 1.00), Mann-Whitney $U = 135.0$, $p = .811$, $r = -.02$. This shows that the barrier did not have an effect on performance for either age group and that the children of both age groups were equally adept at using the cue to find the hidden reward.

Response Latency

Eighteen-month-olds

There was a significant effect of barrier on mean latency to respond within 18-month-olds, with increased latencies in the *barrier* condition (*Mdn* = 13.00*s*) compared with the *no barrier* condition (*Mdn* = 4.75*s*), $Z = -2.44$, $p = .015$, $r = -.16$. This shows that the younger children were slower in responding when a barrier was present.

Thirty-six-month-olds

There was no significant difference in response latency between the two conditions for the 36-month-olds (*barrier Mdn* = 2.88*s*, *no barrier Mdn* = 3.13*s*), $Z = -0.22$, $p = .825$, $r = -.01$. This shows that the older children's response times were unaffected by the barrier.

Age comparisons

In the *barrier* condition, the 18-month-olds (*Mdn* = 13.00*s*) were significantly slower to respond than the 36-month-olds (*Mdn* = 2.88*s*), Mann-Whitney $U = 18.0$, $p < .001$, $r = -.27$. The 18-month-olds (*Mdn* = 4.75*s*) were also significantly slower to respond than the 36-month-olds (*Mdn* = 3.13*s*) in the *no barrier* condition, Mann-Whitney $U = 61.5$, $p = .005$, $r = -.18$. This shows that the 18-month-olds were generally slower to respond than the older children, and these response times were further increased by the presence of a barrier in the testing environment.

Response Type

Eighteen-month-olds

In the barrier condition, there was a significant difference in the proportion of the types of responses elicited, Friedman's χ^2 (3) = 25.41, p <.001. Participants used significantly more *index-finger point*s than grabs, Z = -3.19, p = .001, and reaches, Z = -2.90, p = .005, and significantly more *whole-hand points* than grabs, Z = -2.94, p = .003, and reaches, Z = -2.58, p = .010. There were no other significant differences. In the no barrier condition, there was a significant difference in the proportion of the different response types elicited, Friedman's χ^2 (3) = 38.06, p < .001. Participants used significantly more index-finger points than whole-hand points, Z = -2.55, p = .011 and reaches, Z = -3.30, p = .001, and significantly more grabs than index-finger points, Z = -3.30, p = .001, and reaches, Z = -3.45, p = .001. There were no other significant differences.

There were a number of differences in the response types elicited from the younger group as a function of the presence of a barrier. Eighteen-month-olds used significantly more *index-finger points* in the *barrier* (*Mdn* = .43) than in the *no barrier* (*Mdn* = .13) condition, Z = -3.18, p = .001, r = -.39, as well as significantly more *whole-hand points* in the *barrier* (*Mdn* = .43) than in the *no barrier* condition (*Mdn* = .00), Z = -2.94, p = .003, r = -.43. They grabbed the container significantly less in the *barrier* condition (*Mdn* = .00) than in the *no barrier* (*Mdn* = .88) condition, Z = -3.45, p = .001, r = -.35. There was no significant difference in 18-month-olds' tendency to reach for the container between the *barrier* (*Mdn* = .00) and *no barrier* (*Mdn* = .00) conditions, Z = -1.34, p = .180, r = -.54. This shows that the younger group were more likely to respond using a communicative cue such as an index-finger point or a whole-hand point when there was a barrier present, and more likely to grab the container when there was no barrier present. Analyses were not performed where responses were categorized as *other indicative gesture* or *other*, as these only constituted 0.36% and 1.47% of the total responses (for both age groups combined), respectively. Figure 2a shows the distribution of the response types for the younger children.

Thirty-six-month-olds

There was no significant difference in the proportion of the different response types elicited in the barrier condition, Friedman's χ^2 (3) = 3.08, p = .380. In the no barrier condition, there was a significant difference in the proportion of the different response types elicited, Friedman's χ^2 (3) = 47.24, p < .001. Participants used significantly more index-finger points than whole-hand points, Z = -2.11, p = .035, and reaches, Z = -2.69, p = .007, and significantly more grabs than index-finger points, Z = -3.60, p = <.001, and reaches, Z = -3.89, p = <.001.

There was no significant difference in the proportion of responses that were index-finger points for the 36-month-olds between the *barrier* (*Mdn* = .00) and *no barrier* (*Mdn* = .00) conditions, Z = -1.83, p = .066, r = -.24, but they did use significantly more whole-hand points when the barrier was present (*Mdn* = .00) than when it was not (*Mdn* = .00), Z = -2.20, p = .028, r = -.44. The 36-month-olds grabbed significantly more in the *no barrier* condition (*Mdn* = 1.00) than in the *barrier* condition (*Mdn* = .38), Z = -3.24, p = .001, r = -.24. Thirty-six-month-olds were significantly more likely to reach in the *barrier* (*Mdn* = .00) than in the *no barrier* condition (*Mdn* = .00), Z = -2.69, p = .007, r = -.51. This shows that the older children were also more likely to use some communicative gestures when the barrier was present and again, more likely to grab, or try to grab, the container when the barrier was absent. Figure 2b shows the distribution of response types for the older group.

Age comparisons

The proportion of 18-month-olds' responses that were *index-finger points* was significantly higher than that of 36-month-olds in the *barrier* condition, Mann-Whitney U =

80.5, p = .030, r = -.23, but there were no significant effects of age on use of *index-finger points* in the *no barrier* condition, Mann-Whitney U = 125.5, p = .560, r = -.13. This shows that there were effects of both age and barrier on the use of this type of response, with 18-month-olds using *index-finger points* to indicate the container in which they thought the reward was hidden more than the 36-month-olds in the barrier condition.

The proportion of 18-month-olds' responses that were *whole-hand points* was significantly greater than that of 36-month-olds' in the *barrier* condition, Mann-Whitney U = 81.0, p = .022, r = -.27, but not in the *no barrier* condition, Mann-Whitney U = 140.5, p = .945, r = -.12. This shows that children from both age groups were more likely to indicate their choice using a *whole hand point* when there was a barrier in the testing environment than when there was not, and that this effect of the barrier was particularly pronounced for the 18-month-olds.

Eighteen-month-olds were significantly less likely to grab in the *barrier* condition than the 36-month-olds, Mann-Whitney U = 64.5, p = .006, r = -.41 but not in the *no barrier* condition, Mann-Whitney U = 133.0, p = .758, r = -.02. This shows that both age groups tended to grab the container in which they thought the reward was hidden more when there was no barrier present in the testing environment than when there was, but that this effect was less pronounced for the 36-month-olds. There was no significant effect of age on reaching behaviors in either the *barrier* condition, Mann-Whitney U = 95.5, p = .050, r = -.34. Participants from neither age group reached in the *no barrier* condition. This shows that the 36-month-olds were more likely to reach for the container in which they thought the reward was hidden in the barrier condition, however the 18-month-olds were not.

Discussion

There were no differences in performance (correct responding) as a function of either age or the imposition of a barrier. That both groups of children demonstrated ceiling-level performance when there was no barrier present was expected; however, it is interesting to find that the imposition of a barrier did not have a decreasing effect on success levels for the human children in the way that Kirchhofer et al. (2012) found for domestic dogs. This shows that human children from 18 months are reliably and flexibly able to follow index-finger pointing cues, even with a partial visual barrier, although it must be noted that the two studies differed procedurally in terms of the distances between the containers (1.5m in Kirchhofer et al.) and the locomotor demands on the participants in retrieving the rewards (dogs in Kirchhofer et al.'s study were required to retrieve the object, turn around, and locomote to give it to the experimenter).

The 18-month-olds were slower than the 36-month-olds to choose a container in both the barrier and no barrier conditions, and also showed a marked difference in latency to respond between the conditions, being significantly slower when there was a barrier present than when there was not. Interestingly, however, differences in latency were not associated with performance differences, likely due to the ceiling level performances by both age groups. It may be that these differences in response latencies were due to the unfamiliarity of the situation affecting the younger children more than the older children, or alternatively, due to superior skill in responding to deictic gestures in older children as a function of increased experience with such cues. Leung and Rheingold (1981) found an increase in the ability to comprehend pointing cues associated with age and pointing production in children from 10.5 to 16.5 months, suggesting that comprehension abilities increase with children's own use of these cues. It is worth noting that response latency is not discussed in any of the OCT studies with humans that we reviewed, but according to this explanation, it seems evident that as children become more proficient in both producing and comprehending and gestural cues, they also become quicker to interpret them.

There were differences in the types of response produced by the children as a function of both the imposition of a barrier and age. In the no barrier condition, children of both ages showed a preference for grabbing the container, that is, they overturned the container themselves in order to look inside for the reward. When there was a barrier present, however, both age groups showed an increase in gesturing behavior, that is, they were more likely in this condition to indicate their choice to the experimenter by gesturing, in the form of an index-finger or whole hand point, rather than reaching through the bars to overturn the container themselves. This bias towards gesturing in the barrier condition was particularly prominent in the 18-month-olds, with 36-month-olds often choosing to grab the container themselves, despite the presence of the barrier, something which the younger children did significantly less frequently. Interestingly, in previous studies of human children's performance on the OCT (e.g. Behne, Carpenter & Tomasello, 2005; Behne et al., 2012; Herrmann et al., 2007; Pflandler, Lakatos & Miklósi, 2013), descriptions of the children's behavior when responding to the cue tend to refer to them "searching" or "looking" in the containers. Thus, it can be inferred that typically on the OCT, when no barrier is present, children choose to look inside the container for themselves because previous studies make no mention of children using gestural responses when making a choice. Studies with nonhuman primates differ in the ways in which subjects make their choices, varying from the subject being able to reach through a plexiglass hole to overturn the container themselves (e.g. Barth, Reaux & Povinelli, 2005) to them being required to "indicate" the correct container by reaching their finger through wire mesh (e.g. Herrmann et al., 2007, supplemental material p. 7). That human children are varying their behavioral responses according to whether or not a barrier is present demonstrates that there is an effect of this experimental manipulation and demonstrates the need for consistency in testing environments when comparing across species. The increased use of gesturing, particularly by the younger children, may be explained in terms of the Referential Problem Space hypothesis proposed by Leavens et al.

(2005), that the children see the bars of the playpen as a barrier between themselves and a desirable, but out-of-reach object, and thus use a communicative gesture in order to influence another to retrieve said desirable object. That the object itself was not actually out of reach, but was instead simply partially obstructed, has interesting implications for the way the children perceived the barrier, perhaps as a form of psychological restraint.

Regarding the use of whole-hand points, 18-month-olds used significantly more of these than did 36-month-olds, consistent with Cochet and Vauclair's (2010) finding, in a sample of French children, that the incidence of whole-hand points tends to decrease with age, with a preference for index-finger points emerging. They found that, when points were analysed separately according to function, this correlation between age and hand shape remained for declarative points, but not for imperative points, and this, they suggest, can be taken as evidence for distinctive origins of these two pointing types. Specifically, similarly to Franco and Butterworth (1996), they hypothesized that declarative gestures have a communicative root, whereas imperative gestures originate in failed grasps. They thus assert that the absence of a correlation of age with the use of whole-hand points in an imperative context can be explained by the children preferentially utilising a hand shape that would permit them to grasp the desired object, rather than an index-finger point, which would not allow them to do so. It may also be the case that the function of, and intention behind, each gesture develops between the two age groups tested- the design used in the current study does not allow us to disentangle the children's intention in their use of each response type. It may be that the children, particularly the 18-month-olds, are pointing imperatively to retrieve the ball from the adult, whereas older children may intend to show the adult the location of the ball. Alternatively, our finding that the incidence of whole-hand pointing decreased with age could be explained through its being a product of increasing experience with conventionalized gestures. That is, the 36-month-olds, as a result of their superior level of experience of, and exposure to, index-finger points, are responding in a more conventionalized manner than the 18-month-olds, in terms of the shape of the hand when gesturing (Leavens & Hopkins, 1999).

Study 2 Manipulating the distance of the reward

Given our finding that the children's behavioral responses differed as a function of the imposition of a barrier, which they may have perceived as a physical restraint to obtaining the reward themselves, we thought it would be of interest to investigate the effects of the distance of the reward. In Study 2, therefore, we focused on manipulating the distance between the participant and the containers, specifically whether the participant was able to reach the reward or not, in order to examine the effects of placing the reward out of reach on children's communicative behavior. All participants were tested from within the playpen, and took part in a *proximal* condition, comparable to the *barrier* condition in Study 1, in which the containers were outside of the playpen but within reach of the participant, and a *distal* condition, in which the containers were placed outside of the barrier and out of reach of the participant.

Method

Participants

Participants were seventeen 36-month-olds (*M* = 36 mos 4days, *range* = 31 mos 30 days – 39 mos 26 days), comprised of 6 males (*M* = 37 mos 4 days, *range* = 31 mos 30 days – 39 mos 19 days) and 11 females (*M* = 36 mos 11 days, *range* = 32 mos 27 days – 39 mos 26 days). Participants were recruited from a participant database, where parents had registered their interest in participating with their children in cognitive studies, and from advertisements on social media sites. Data were collected between December 2017 and January 2018.

Procedure

After the same "settling in" period as in Study 1, the participant and caregiver entered the test room with the experimenter and were seated inside the playpen, as in the *barrier* condition in Study 1. The children were given a bowl of stickers and asked to choose one to keep in order for the child to become familiar with the available rewards. The experimenter then hid a sticker in one of two opaque plastic containers, and this baiting took place behind an occluder, in the form of a large sheet of brown cardboard. The experimenter then placed the containers in either the *proximal* or *distal* position (see Figure 3) and informed the participant that they were going to give them a clue to see if they could find the hidden sticker. The experimenter then used an ipsilateral dynamic pointing cue to indicate to the participant in which container the reward was hidden, and this cue was held until the participant made a choice. If the participant chose the correct container, the experimenter gave verbal praise, opened the container and gave the sticker to the participant to keep. If the participant chose the incorrect container, the experimenter opened that container, said "Oh dear, it's not in here, let's see if it's in the other one", opened the other container and showed the contents to the child, then returned the sticker to the bowl. If the child did not respond within two minutes or was fussy and trying to get out from the playpen, the trial was terminated, the bowl of stickers shown to the child again in order to increase motivation, and a new trial commenced. Participants completed two proximal and two distal trials and order of administration was counterbalanced across participants. The baited container was on the right or left an equal number of times in each condition, and the order was counterbalanced.

Data Scoring

Data were coded according to the same coding scheme as in Study 1.

Analysis

Five participants were excluded from the final analyses due to experimenter error during testing. Three of these cases were due to the experimenter using a momentary, rather than a dynamic, point, and two were because the cameras were placed such that the experimenter was not in shot in the videos, and therefore the moment of pointing could not be ascertained.

An independent coder who was blind to the purpose of the study coded 20% of the videos (2 participants, with four trials each, for a total of eight trials). For correct choices, there was complete agreement between the two coders, $\kappa = 1.00$, $p = .005$. There was also complete agreement for response type, $\kappa = 1.00$, $p < .001$ and excellent agreement for latency, $r_s = .88$, $p < .001$.

Results

Correct Choices

The data were not normally distributed and so non-parametric tests were used throughout the analyses. There was no significant difference in the proportion of correct choices made in the *proximal* (*Mdn* = 1.00) and *distal* (*Mdn* = 1.00) conditions, $Z = -1.00$, $p = .317$, $r = -.15$. This shows that the children performed at ceiling level in both conditions, as expected.

Response Latency

There was no significant difference in the mean latency to respond between the *proximal* (*Mdn* = 3.00 secs) and *distal* (*Mdn* = 3.75 secs) conditions, $Z = 0.00$, $p = 1.00$, $r = 0$. This shows that the children were equally quick to respond in both conditions.

Type of Response

Only two of the possible response types were used by the children, these were index-finger points and grabs. Figure 4 shows the mean proportion of each type of response used in the two conditions. The proportion of trials in which the children used an index-finger point to indicate their choice of container was significantly lower in the *proximal* condition (*Mdn* = 1.00) than in the *distal* condition (*Mdn* = 1.00), $Z = -2.24$, $p = .025$, $r = -.35$. This shows that when the containers were out of reach, the children were more likely to respond by using an index-finger point than when the containers were within reach.

The proportion of trials in which the children grabbed the container was significantly higher in the *proximal* (*Mdn* = 1.00) than in the *distal* condition (*Mdn* = .00), $Z = -2.24$, $p = .025$, $r = -1.00$. This shows that the children were more likely to indicate their choice by grabbing the container when the containers were within reach.

Order of administration and grabbing behavior

In order to investigate whether the order of presentation had an effect on the response types used, participants were categorized as 'grabbers' (grabbed the container on at least one trial in the *proximal* condition or 'non-grabbers' (did not grab on either trial in the proximal condition). There was no significant effect of order of administration on the likelihood of grabbing (*proximal* first *Mdn* = 1.00; *distal* first *Mdn* = .00), Mann-Whitney $U = 9.00$, $p = .093$, $r = -.25$. There was a trend, however, such that those tested with the *proximal* condition first were more likely to grab in the *proximal* condition, whereas those tested with the *distal* condition first were slightly less likely to grab at all in the *proximal* condition (see Figure 5).

Discussion

As expected, participants performed at ceiling level in both conditions, further demonstrating that 3-year-olds were able to reliably follow a pointing gesture to find a hidden reward. There were no differences in the children's response latencies between the two conditions, showing that this ability is flexible even across increased distances between the child and the object being signalled.

The children in this study used only two response types to indicate the container in which they thought the reward was hidden, index-finger points and direct grabs of the container. These were the responses most often utilized by the 3-year-olds in Study 1, however, it is notable that there was an absence of the use of whole-hand points and reaches in the current study. The absence of whole-hand points is congruent with Cochet and Vauclair's (2010) findings that the incidence of whole-hand pointing in a declarative context decreases with age and that, here, the children were responding to the experimenter by demonstrating where they believed the reward to be hidden, rather than demanding the cup in an imperative manner. As in Study 1, the children chose to grab the container to look inside themselves in the *proximal* condition on a number of trials. That they did not try to reach for the container on any of the *distal* trials, nor did they exhibit any whole-hand points - which Cochet and Vauclair (2010) argued could be the result of failed grasping attempts in an imperative context- demonstrates that 3-year-olds are not categorically 'grabbers', but rather that their grabbing responses are a result of a spatial evaluation. When the container is not in reach, they do not try to grab it.

Although there was no significant difference, there was a statistically non-significant trend towards increased grabbing in the *proximal* condition when this was the first condition administered than when it followed the *distal* condition. Specifically, only one of the six participants tested with the *distal* condition first grabbed the container in either of the two trials in the *proximal* condition, compared with four out of six children tested with the *proximal* condition first. This, like Study 1, has interesting implications for the way the children perceive the barrier, with one possible explanation being that those tested with the *distal* condition formed a perception of the barrier as a restraint that prevented them from being able to retrieve the reward themselves, and maintained this perception once the containers were actually moved within reach, such that they continued to use communicative cues to indicate their choice rather than grab for it themselves. An alternative explanation may be that the index-finger pointing became a perseverative response- once this had been effective as tool to retrieve the desired out-of-reach object in the *distal* condition, they habitually continued to use this response in the *proximal* condition.

General Discussion

The children, all of an age at which their comprehension of the pointing cue is at mastery level, displayed differential behavioral responses according to the configurational set up of the experiment, showing that these manipulations do have an effect on communicative behavior. This has implications for the wealth of OCT literature which compares across species with little to no regard for matching experimental conditions (Leavens et al., 2019). Human children, with 18- or 36-month-long histories rich in human interactions and exposure to human pointing cues and who are experts in using these cues to influence the behavior of others, react differently when tested with a barrier in the testing environment to when tested without.

Here we present the results of two studies with children in which elements of the configuration of the OCT were manipulated, in order to investigate whether such manipulations affected the children's behavioral responses, especially their decisions to either elicit aid from the experimenter or to act directly on the apparatus. In Study 1, we tested 18-month-olds and 36-month-olds on a standard version of the OCT, in which participants were tested with and without a barrier in an attempt to mimic the testing conditions used with nonhuman primates. In Study 2, we tested 36-month-olds with a barrier, manipulating the distance of the reward, such that it was placed either within or out of reach of the participant. We found that we could manipulate apparent changes in the children's motivations through the imposition of a permeable barrier between them and the target containers. Absent a barrier, the children tended to adopt a praxic mode of interaction with the apparatus, grabbing the containers directly, whereas imposition of the barrier tended to foster a communicative mode of engagement with the experimenter. This switch from manipulating objects to manipulating an agent is consistent with a longstanding interpretation of intentional communication as a kind of social tool use—the use of an agent in goal-directed activities (e.g., Bard, 1990; Bates, Camaioni, & Volterra, 1975; Leavens, Hopkins, & Bard, 1996). Notably, in Study 1, the 18-month-olds displayed a virtually complete shift from a praxic mode to a communicative mode in the Barrier condition, whereas the 3-year-olds displayed a mixture of tactics in the Barrier condition, with some still grabbing the containers through the bars of the playpen. This demonstrates that, in representatives of our own species, life history stage influences the size of the palette of response forms, with older children displaying a proportionately lower propensity to use communication in the face of a permeable barrier, relative to younger children. In our studies, both praxic and communicative responses were deemed “correct,” but it is commonplace to present adult nonhumans with versions of an OCT in which only communicative responses or only praxic responses were considered “correct” (e.g., Call & Tomasello, 1994; Hopkins, Russell, McIntyre, & Leavens, 2013; Povinelli et al., 1999). Thus, when presenting similar tasks across representatives of different species, it is important to consider how response requirements or expectations interact with life history stage and pre-experimental life experience, to ensure that organisms are given the best opportunity to display their cognitive competencies; for example, it is well-demonstrated that enculturated apes significantly outperform institutionalised apes in similar experimental contexts (Lyn et al., 2010; Russell et al., 2011; and see, in a different context, arguments by Horowitz, 2003, and Thomas, Murphy, Pitt, Rivers, & Leavens, 2008, to the effect that younger humans are not representative of adult humans in some cognitive assays). Thus, because previous cross-species comparisons have generally not controlled for life history stage or task-relevant pre-experimental experience (Clark et al., 2019; Krause et al., 2018; Leavens et al., 2019), and because the present study shows a developmental shift in human children toward a reduced reliance on the use of communication in the presence of a barrier, therefore, we recommend a systematic revision to the OCT that permits both communicative and praxic responses. This adjustment will foster best performance to be captured and reduce the existing bias towards false negatives in some testing circumstances.

In conclusion, here we add to and extend the arguments put forward in recent reviews that detail the procedural and methodological flaws in the OCT literature (Lyn, 2010; Mulcahy & Hedge, 2012), and analyses of ape-human comparisons more generally (Leavens, 2014, 2018; Leavens et al., 2019) and emphasize the necessity of ensuring matched conditions, as well as selection of correct response patterns that are appropriate to life history stage, in experimental testing. Furthermore, we demonstrate that these subtle manipulations of the testing environment can lead to large differences in the behavioral responses of members of the taxonomic group most experienced in the use of human gestural cues. This significantly challenges theories that rely on generalizations of the ability of representatives of a single sample to their entire species in studies which fail to adequately control for testing environment.

References

Anderson, J. R., Sallaberry, P. & Barbier, H. (1995) Use of experimenter-given cues during object-choice tasks by capuchin monkeys. *Animal Behaviour, 49*, 201-208. https://doi.org/10.1016/j.applanim.2012.07.005.

Bard, K. A. (1990). "Social tool use" by free-ranging orangutans: A Piagetian and developmental perspective on the manipulation of an animate object. In S. T. Parker & K. R. Gibson (Eds.), *"Language" and intelligence in monkeys and apes: Comparative developmental perspectives* (pp. 356-378). Cambridge, England: Cambridge University Press. https://doi.org/10.1017/cbo9780511665486.015

Bard, K. A. & Leavens, D. A. (2014) The importance of development for comparative psychology. *Annual Review of Anthropology*, *43*, 183-200. https://doi.org/10.1146/annurev-anthro-102313-030223

Barrera, G. & Bentosela, M. (2016) Interspecific social skills in a highly socialized Aguará Guazú (maned wolf, *Chrysocyon brachyurus*). A pilot subject study. *Current Ethology*, *15*(1), 55-63.

Barth, J., Reaux, J. E. & Povinelli, D. J. (2005) Chimpanzees' (*Pan troglodytes*) use of gaze cues in object-choice tasks: Different methods yield different results. *Animal Cognition, 8*, 84-92. https://doi.org/10.1007/s10071-004-0235-x

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly of Behavior and Development*, *21*, 205–226.

Behne, T., Carpenter, M. & Tomasello, M. (2005) One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science, 8*(6), 492-499. https://doi.org/10.1111/j.1467-7687.2005.00440.x

Behne, T., Liszkowski, U., Carpenter, M. & Tomasello, M. (2012) Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental*

Psychology, 30, 359-375. https://doi.org/10.1111/j.2044-835X.2011.02043.x

Bräuer, J., Kaminski, J., Riedel, J., Call, J. & Tomasello, M. (2006) Making inferences about the location of hidden food: Social dog, causal ape. *Journal of Comparative Psychology, 120*(1), 38-47. http://dx.doi.org/10.1037/0735-7036.120.1.38

Butterworth, G. (2001) Joint visual attention in infancy. In J. G. Bremner & A. Fogel (Eds.), *Blackwell handbook of infant development* (pp. 213-240). Hove: Blackwell.

Butterworth, G. & Morissette, P. (1996) Onset of pointing and the acquisition of language in infancy. *Journal of Reproductive and Infant Psychology*, *14*(3), 219-231. https://doi.org/10.1080/02646839608404519

Call, J., & Tomasello, M. (1994). Production and comprehension of referential pointing by orangutans (*Pongo pygmaeus*)." *Journal of Comparative Psychology*, *108*, 307-317.

Clark, H., Elsherif, M., & Leavens, D. A. (2019). Ontogeny vs. phylogeny in Primate/Canid comparisons: An Object Choice Task review. *Neuroscience and Biobehavioural Reviews, Advance online publication.* https://dx.doi.org/10.1016/j.neubiorev.2019.06.001.

Cochet, H. & Vauclair, J. (2010) Pointing gestures produced by toddlers from 15 to 30 months: Different functions, hand shapes and laterality patterns. *Infant Behaviour and Development*, *33*, 431-441. https://doi.org/10.1016/j.infbeh.2010.04.009

D'Aniello, B., Alterisio, A., Scandurra, A., Petremolo, E., Iommelli, M. R. & Aria, M. (2017) What's the point? Golden and Labrador Retrievers living in kennels do not understand pointing gestures. *Animal Cognition*, *20*, 777-787. https://doi.org/10.1007/s10071-017-1098-2

Duranton, C. & Gaunet, F. (2016) Effects of shelter housing on dogs' sensitivity to human social cues. *Journal of Veterinary Behaviour, 14*, 20-27. https://doi.org/10.1016/j.jveb.2016.06.011.

Franco, F. & Butterworth, G. (1996) Pointing and social awareness: Declaring and requesting in the second year. *Journal of Child Language*, *23*, 307-336. https://doi.org/10.1017/S0305000900008813

Hare, B. & Tomasello, M. (2005) Human-like social skills in dogs? *TRENDS in Cognitive Sciences*, *9*(9), 439-444. https://doi.org/10.1016/j.tics.2005.07.003

Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B. & Tomasello, M. (2007) Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, *317*, 1360-1367. https://doi.org/10.1126/science.1146282

Hopkins, W. D., Russell, J. L., McIntyre, J., & Leavens, D. A. (2013). Are chimpanzees really so poor at understanding imperative pointing? Some new data and an alternative view of canine and ape social cognition. *PLoS ONE*, *8*, e79338. https://doi.org/10.1371/journal.pone.0079338

Horowitz, A. C. (2003). Do humans ape? Or do apes human? Imitation and intention in humans (*Homo sapiens*) and other animals. *Journal of Comparative Psychology, 117,* 325–336. https://doi.org/10.1037/0735-7036.117.3.325

Kirchhofer, K. C., Zimmermann, F., Kaminski, J. & Tomasello, M. (2012) Dogs (*Canis familiaris*), but not chimpanzees (*Pan troglodytes*), understand imperative pointing. *PLoS ONE*, *8*(2), e30913. https://doi.org/10.1371/journal.pone.0030913

Kishimoto, T. (2013). Do referential problem spaces affect the frequency of imperative pointing by infants? *Psychology of Language and Communication*, *17*, 295-307. https://doi.org/10.2478/plc-2013-0019

Krause, M. A., Udell, M. A. R., Leavens, D. A., & Skopos, L. (2018). Animal pointing: Changing trends and findings from 30 years of research. *Journal of Comparative Psychology*, *132*, 326-345. http://dx.doi.org/10.1037/com0000125

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159-174. https://doi.org/10.2307/2529310

Lazarowski, L. & Dorman, D. C. (2015) A comparison of pet and purpose-bred research dog (*Canis familiaris*) performance on human-guided object-choice tasks. *Behavioural Processes*, *110*, 60-67. https://doi.org/10.1016/j.beproc.2014.09.021

Leavens, D. A. (2014). The plight of the sense-making ape. In M. Cappuccio & T. Froese (Eds.), *Enactive cognition at the edge of sense-making* (pp. 81-104). Basingstoke, U.K.: Palgrave Macmillan. https://doi.org/10.1057/9781137363367_4

Leavens, D. A. (2018). The cognitive implications of intentional communication: A multi-faceted mirror. In L. Di Paolo, F. Di Vincenzo, & F. De Petrillo (Eds.), *Evolution of primate social cognition* (pp. 59-77). New York: Springer. https://doi.org/10.1007/978-3-319-93776-2

Leavens, D. A. & Hopkins, W. D. (1999) The whole-hand point: The structure and function of pointing from a comparative perspective. *Journal of Comparative Psychology*, *113*(4), 417- 425. http://dx.doi.org/10.1037/0735-7036.113.4.417

Leavens, D. A., Bard, K. A., & Hopkins, W. D. (2019). The mismeasure of ape social cognition. *Animal cognition*, *22*, 487-504. https://doi.org/10.1007/s10071-017-1119-1

Leavens, D. A., Hopkins, W. D., & Bard, K. A. (1996). Indexical and referential pointing in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *110*, 346-353. https://doi.org/10.1037/0735-7036.110.4.346

Leavens, D. A., Hopkins, W. D. & Bard, K. A. (2005) Understanding the point of chimpanzee pointing: Epigenesis and ecological validity. *Current Directions in Psychological Science*, *14*(4), 185-189. https://doi.org/10.1111/j.0963-7214.2005.00361.x

Leung, E. H. L. & Rheingold, H. L. (1981) Development of pointing as a social gesture. *Developmental Psychology, 17*(2), 215-220. http://dx.doi.org/10.1037/0012-

1649.17.2.215

Lyn, H. (2010) Environment, methodology, and the object choice task in apes: Evidence for declarative comprehension and implications for the evolution of language. *Journal of Evolutionary Psychology*, *8*(4), 333-349. https://doi.org/10.1556/JEP.8.2010.4.3

Lyn, H., Russell, J. L. & Hopkins, W. D. (2010) The impact of environment on the comprehension of declarative gestures in apes. *Psychological Science*, *21*(3), 360-365. https://doi.org/10.1177/0956797610362218

Masataka, N. (2003) From index-finger extension to index-finger pointing: The ontogenesis of pointing in preverbal infants. In S. Kita (Ed.), *Pointing: Where language, culture and cognition meet* (pp. 9-33). Mahwah, NJ: Lawrence Erlbaum Associates.

Miklósi, A., & Soproni, K. (2006). A comparative analysis of animals' understanding of the human pointing gesture. *Animal Cognition*, *9*(2), 81-93. https://doi.org/10.1007/s10071-005-0008-1

Mulcahy, N. J. & Hedge, V. (2012) Are great apes tested with an abject object-choice task? *Animal Behaviour*, *83*, 313-321. https://doi.org/10.1016/j.anbehav.2011.11.019

Pflandler, E., Lakatos, G. & Miklósi, A. (2013) Eighteen-month-old human infants show intensive development in comprehension of different types of pointing gestures. *Animal Cognition, 16*, 711-719. https://doi.org/10.1007/s10071-013-0606-2

Povinelli, D. J., Bierschwale, D. T., & Čech, C. G. (1999). Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. *British Journal of Developmental Psychology, 17,* 37- 60. https://doi.org/10.1348/026151099165140

Russell, J.L., Lyn, H., Schaeffer, J.A., Hopkins, W.D. (2011). The role of socio-communicative rearing environments in the development of social and physical cognition in apes. *Developmental Science*, *14*, 1459‑1470. https://doi.org/10.1111/j.1467-7687.2011.01090.x.

van Schaik, C. P., & Burkart, J. M. (2011). Social learning and evolution: The cultural intelligence hypothesis. *Philosophical Transactions of the Royal Society, Series B*, *366*, 1008-1016. https://doi.org/10.1098/rstb.2010.0304

Thomas, E., Murphy, M., Pitt, R., Rivers, A., & Leavens, D. A. (2008). Understanding of visual attention by adult humans (*Homo sapiens*): A partial replication of Povinelli, Bierschwale, and Čech (1999). *Journal of Comparative Psychology*, *122*, 428-436. https://doi.org/10.1037/0735-7036.122.4.428

Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of novel communicative signs by apes and human children. *Child Development, 68*(6)*,* 1067- 1080. https://doi.org/10.1111/j.1467-8624.1997.tb01985.x

Udell, M. A. R., Dorey, N. R. & Wynne, C. D. L (2008) Wolves outperform dogs in following human social cues. *Animal Behaviour*, *76*, 1767-1773. https://doi.org/10.1016/j.anbehav.2008.07.028

Udell, M. A. R., Giglio, R. F., & Wynne, C. D. L. (2008). Domestic dogs (*Canis familiaris*) use human gestures but not nonhuman tokens to find hidden food. *Journal of Comparative Psychology, 122,* 84–93. http://dx.doi.org/10.1037/0735-7036.122.1.84

Udell, M. A. R., Spencer, J. M., Dorey, N. R. & Wynne, C. D. L. (2012) Human-socialized wolves follow diverse human gestures... And they may not be alone. *International Journal of Comparative Psychology*, *25*, 97-117. https://doi.org/10.1016/j.anbehav.2008.07.028

Acknowledgements

We thank the School of Psychology and the Doctoral School, University of Sussex for funding this research. Dan Hyndman and Martha Casey provided technical support. We thank Dr. Jessica Horst and Professor Nicola Yuill for advice and sharing of laboratory and office space. We are also grateful to the editor and four anonymous reviewers for their insightful and authoritative remarks. Finally, we are grateful to the parents and children who volunteered to take part in these studies.

Figure Captions

Figure 1. The experimental set-up. P = participant; CG = caregiver; E = experimenter. Barrier represented by dashed line. Drawing not to scale; positions and distances approximate.

Figure 2. The mean proportion of response types in the *barrier* and *no barrier* conditions by a) 18-month-olds and b) 36-month-olds, with standard errors. IFP= Index-finger point; WHP = whole-hand point. Means and standard errors are depicted, here, to more clearly display the effects, although nonparametric statistical tests were applied. $^{*}p$ <.05.

Figure 3. The experimental set-up for a) the proximal condition and b) the distal condition in Study 2. CG = caregiver, P = participant, E = experimenter. Barrier represented by dashed line. Drawing not to scale, positions and distances approximate.

Figure 4. The mean proportion of response types, with standard errors, in the *proximal* and *distal* conditions. Means and standard errors are depicted, here, to more clearly display the effects, although nonparametric statistical tests were applied. IFP = index finger point. * p <.05.

Figure 5. The number of participants tested with either the *proximal* or *distal* condition first who grabbed in at least one trial.

Figure 1

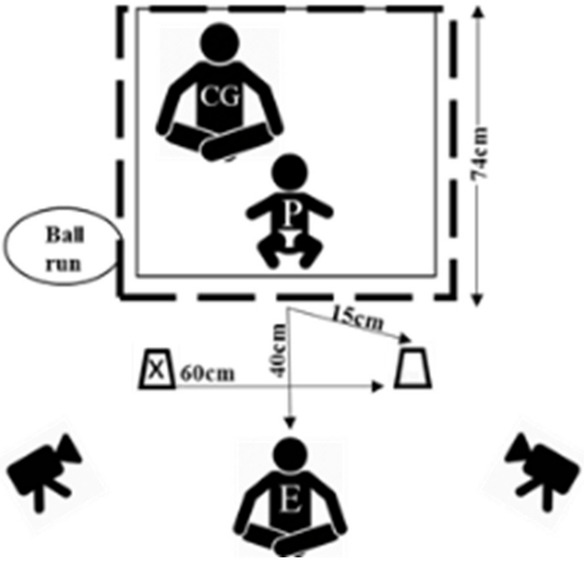


Figure 2

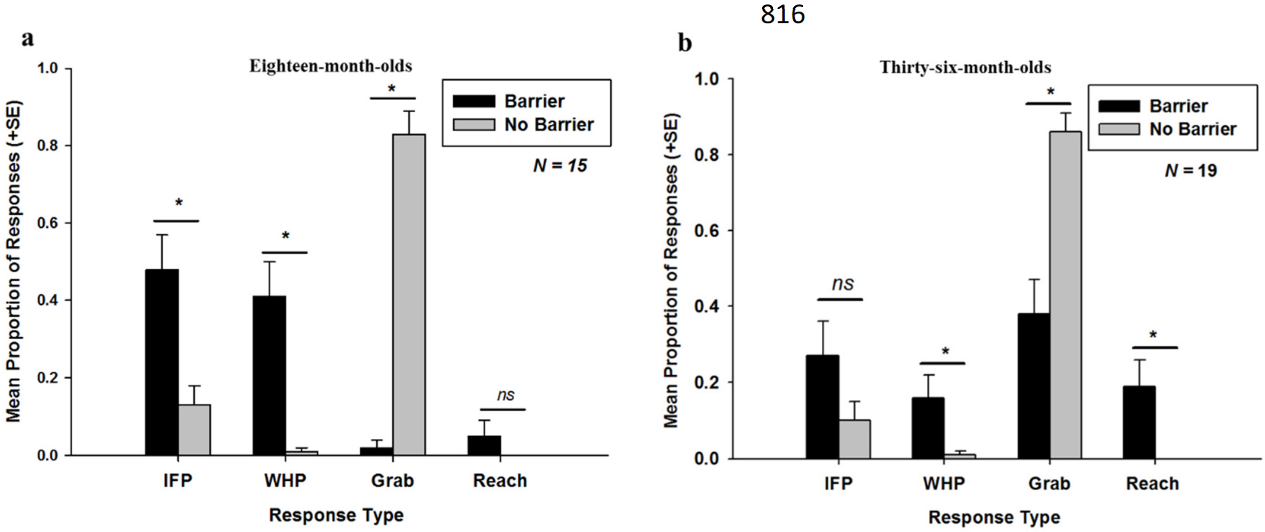
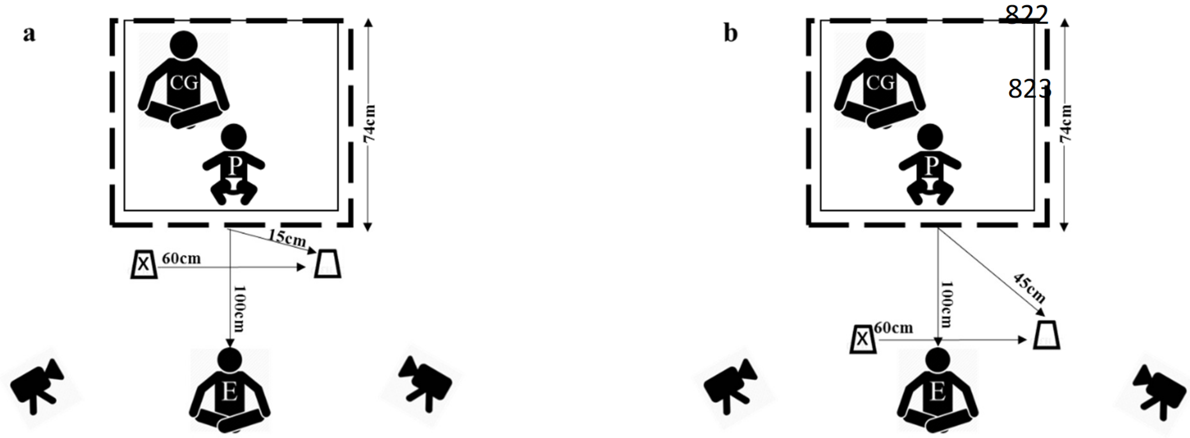

a
Eighteen-month-olds
Barrier
No Barrier
N = 15
Mean Proportion of Responses (+SE)
1.0
0.8
0.6
0.4
0.2
0.0
*
ns
IFP
WHP
Grab
Reach
Response Type
b
Thirty-six-month-olds
Barrier
No Barrier
N = 19
Mean Proportion of Responses (+SE)
1.0
0.8
0.6
0.4
0.2
0.0
*
ns
IFP
WHP
Grab
Reach
Response Type

Figure 3



824 Figure 4

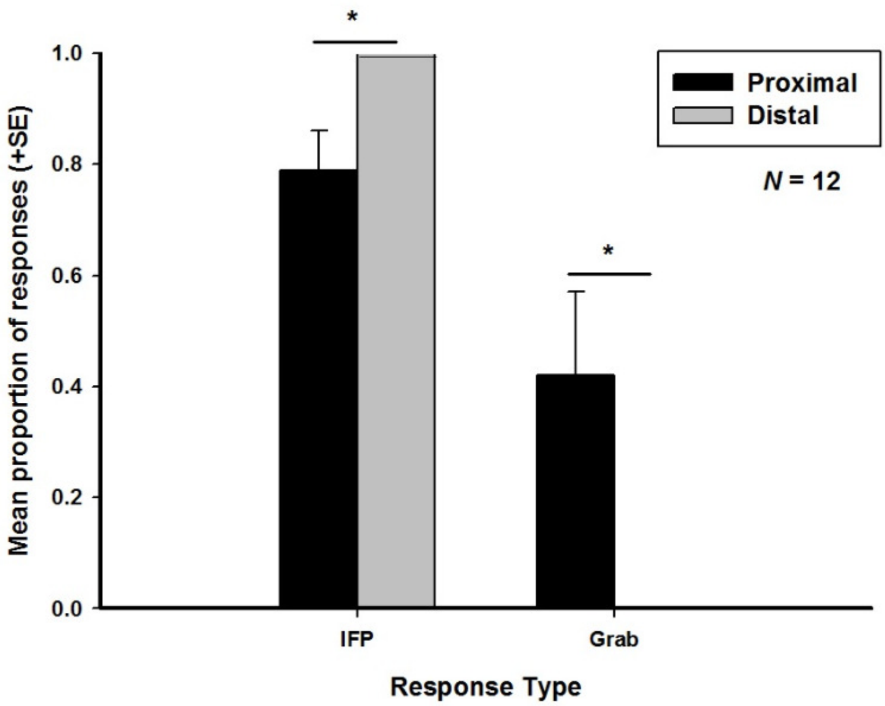
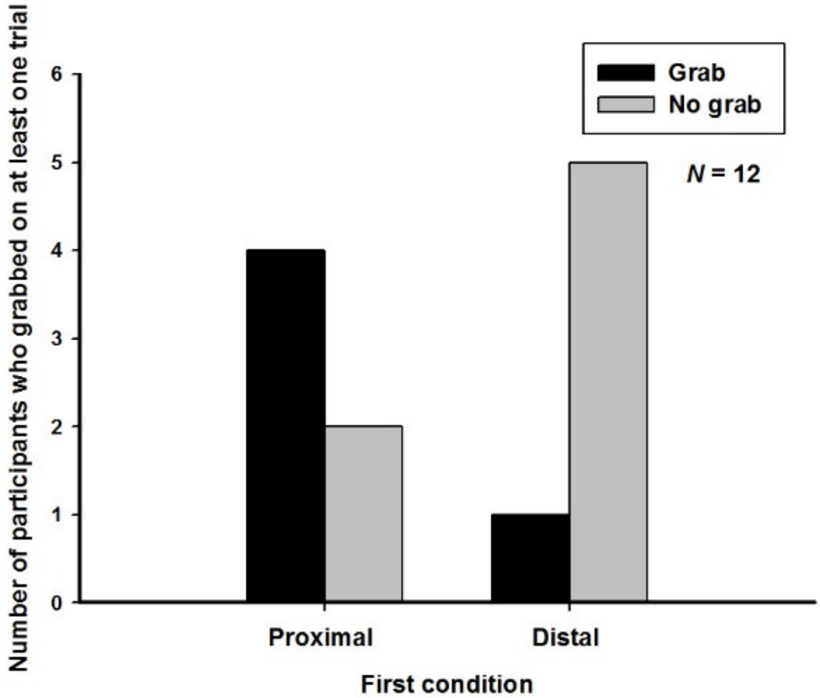

*
*
1.0
0.8
0.6
0.4
0.2
0.0
Mean proportion of responses (+SE)
Proximal
Distal
N = 12
IFP
Grab
Response Type

Figure 5


Grab
No grab
N = 12
Number of participants who grabbed on at least one trial
0
1
2
3
4
5
6
Proximal
Distal
First condition