

1 Running head: HUMAN REFERENTIAL PROBLEM SPACE

2

3

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

6

7

8

9 Hannah Clark<sup>1</sup>, Zoe M. Flack<sup>2</sup> and David A. Leavens<sup>1,\*</sup>,

10

11 <sup>1</sup>School of Psychology, University of Sussex, Falmer, East Sussex, BN1 9QH

12 <sup>2</sup>School of Applied Social Science, University of Brighton, Brighton, East Sussex, BN1 9PH

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

15

16

17

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

19

20

21

22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

**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.

## HUMAN REFERENTIAL PROBLEM SPACE

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

64           The OCT is used to assess an individual's ability to comprehend human gestural cues  
65 and involves an experimenter presenting a deictic cue to indicate to the subject in which of  
66 two or three containers a reward has previously been hidden (Anderson, Sallaberry &  
67 Barbier, 1995). Results of OCT studies have been used as the bases for theories pertaining to  
68 the evolutionary roots of social cognition in a number of species, especially humans (*Homo*  
69 *sapiens*; e.g. Povinelli, Bierschwale, & Čech, 1999; Tomasello, Call, & Gluckman, 1997),  
70 nonhuman primates (Primates; e.g., Anderson et al., 1995; Povinelli et al., 1999; Tomasello et  
71 al., 1997), and domestic dogs (*Canis familiaris*; e.g., Udell, Giglio, & Wynne, 2008). For

## HUMAN REFERENTIAL PROBLEM SPACE

72 example, Hare and Tomasello (2005), on the basis of domestic dogs' consistently adept  
73 performance, argued for a theory of convergent evolution between dogs and humans, in  
74 which the former have developed specialized socio-cognitive skills to comprehend human  
75 gestural cues as a result of centuries of artificial selection during domestication. The Cultural  
76 Intelligence hypothesis (Herrmann et al., 2007; van Schaik & Burkart, 2011), based, in part,  
77 on nonhuman primates' generally poor performance on the OCT, states that humans have  
78 developed species-specific socio-cognitive skills in order to facilitate cultural group living,  
79 and, as such, the comprehension of human gestural cues is a human-unique ability within the  
80 primates. Bräuer, Kaminski, Riedel, Call and Tomasello (2006) proposed that differences in  
81 the performance levels between dogs and nonhuman primates are due to species-specific  
82 specialisations where dogs have been selected for specialized social abilities, which thus  
83 enable them to follow human gestural cues, whereas apes' foraging behavior has led to  
84 increased physical abilities, which explains their ability to use physical but not social cues on  
85 the OCT. Leavens, Hopkins, and Bard (2005) posited a Referential Problem Space to account  
86 for the very sparse observations of pointing in wild chimpanzees, compared to captive apes,  
87 who very frequently point; they suggested that pointing emerges in circumstances in which  
88 pointers are forced to rely on the manipulation of others to obtain objects of interest, a  
89 situation that characterises both human infants, who are often restrained, and captive apes,  
90 whose prehension is often blocked by cage mesh.

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

## HUMAN REFERENTIAL PROBLEM SPACE

97 (Butterworth, 2001; Butterworth & Morrisette, 1996) in an environment rich in human  
98 interactions. It is this developmental process that Bard and Leavens (2014) argued is essential  
99 to consider when making cross-species comparisons in socio-cognitive tasks, and they  
100 highlight the contemporary prevalence of basing phylogenetic theories on the performance of  
101 subjects unmatched for developmental experience. Indeed, Lyn (2010) argued that pre-  
102 experimental exposure to humans can differentially affect an individual's performance on the  
103 OCT, and evidence is accumulating to support this argument. Lyn, Russell and Hopkins  
104 (2010) found enculturated nonhuman primates to be successful at following human gestural  
105 cues and a growing body of work shows that domestic dogs with less exposure to humans  
106 perform significantly worse than their pet dog counterparts (e.g., D'Aniello et al., 2017;  
107 Duranton & Gaunet, 2016; Lazarowski & Dorman, 2015; Udell, Dorey & Wynne, 2008),  
108 whereas other canids raised in environments rich in human interactions perform well on the  
109 task (Barrera & Bentosela, 2016; Udell, Dorey, Spencer & Wynne, 2012; Udell et al., 2008).

110         Leavens et al. (2019) argued that a further inconsistency in the testing protocols  
111 adopted with different species is that of the presence of a barrier between the subject and the  
112 baited container in the testing environment. Due to safety issues surrounding the testing of  
113 nonhuman primates, species from this taxonomic group are tested from outside their cages,  
114 therefore imposing a barrier between the subject and the experimenter and test apparatus.  
115 Working with individuals from other taxonomic groups, domestic dogs or human infants, for  
116 example, does not necessitate the use of such safety precautions, and, as such, there is an  
117 absence of this barrier in the testing environment with these species; this constitutes a  
118 confound between experimental protocol and species classification in a significant number of  
119 contemporary studies. In a review of 71 published nonhuman primate and dog OCT studies,  
120 Clark et al. (2019) found that 99% of nonhuman primates were tested with a barrier present in  
121 the testing environment, compared with less than 1% of dogs. They therefore argued that this

## HUMAN REFERENTIAL PROBLEM SPACE

122 inconsistency in the test set-ups used across different species represents an experimental  
123 confound that may affect individuals' performances. This renders conclusions that group  
124 differences are attributable to differences in evolutionary histories implausible.

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

143         In order to investigate the possible confounding effects of this systematic difference in  
144 experimental protocols administered to representatives of different species, we tested children  
145 with and without a barrier on an OCT. Children from 14 months of age have been shown to  
146 reliably follow pointing cues on the OCT (Behne, Lizkowski, Carpenter & Tomasello, 2012)

147 and so we tested children aged 18 months and 36 months in order to ensure that any  
148 differences in performance or behavioral responses between the two conditions were as a  
149 result of our experimental manipulation, rather than the lack of emergence of these skills.

150

### 151 **Study 1: Barrier vs. No Barrier**

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

159

### 160 **Method**

#### 161 **Participants**

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

#### 171 **Procedure**

## HUMAN REFERENTIAL PROBLEM SPACE

172           On arrival at the testing suite, participants and their parents were given time to  
173 become familiar with the surroundings, with participants playing freely in the playroom and  
174 interacting with the experimenter. When parents judged their child to be settled and  
175 comfortable, the experimenter, participant, and parent moved to the testing room, where the  
176 experimenter demonstrated a “ball run” toy to the child, and then encouraged the child to play  
177 with the toy. The test room was set up as shown in Figure 1. The experimenter then informed  
178 the child that they were going to play a fun hiding game with the balls, and that if the child  
179 found the balls they could put them in the ball run. The experimenter then asked the child to  
180 sit on the playmat with their parent and explained that she would hide the ball under one of  
181 two cups, and then give the participant a “clue” to see if they could find it. In the *barrier*  
182 condition, a child’s playpen was set up, such that the playmat was inside the pen, and the ball  
183 run was outside, but accessible to the child. In the *no barrier* condition, the playmat and ball  
184 run were in the same positions, but the playpen was not in place. The experimenter hid the  
185 ball under one of two cups, behind a cardboard occluder, then made eye contact with the  
186 participant, asking “are you ready for your clue?” The experimenter then presented an  
187 ipsilateral, dynamic, index-finger pointing cue, whilst alternating her gaze between the  
188 container and the experimenter. A dynamic point, according to Miklósi and Soproni’s (2006)  
189 definition of the different point types used on the OCT, is one in which the pointing gesture is  
190 carried out in front of the participant and remains in place until the participant makes a  
191 choice. The distance between the experimenter’s fingertip and the container was  
192 approximately 5cm. The experimenter maintained this position until the participant made a  
193 choice. If the participant was unresponsive, the experimenter encouraged the participant to  
194 make a choice by giving verbal encouragement such as “can you find that ball?” If the  
195 participant failed to respond after approximately 2 minutes, or was fussy (for example, trying  
196 to get out of the playpen), then the trial was terminated, and the experimenter attempted to



## HUMAN REFERENTIAL PROBLEM SPACE

197 increase motivation by again demonstrating the ball in the ball run. If the participant made a  
198 correct choice, they were given the ball and encouraged to put it in the ball run. If the  
199 participant made an incorrect choice, the experimenter lifted the incorrect cup and said, for  
200 example, “oh no! It’s not in that one! Let’s see if it was in the other one!” and then lifted the  
201 correct cup, showed the child the ball, and said “Never mind! Let’s hide it again!”  
202 Participants received 8 trials in the *barrier* condition and 8 trials in the *no barrier* condition.  
203 Order of administration was counterbalanced across participants, and in between conditions,  
204 participants left the test room with their parent and were engaged in another task, such as  
205 looking at wall stickers of animals. The baited container was on the right or left an equal  
206 number of times in each condition, and the order was counterbalanced, such that the reward  
207 was never on the same side for more than two consecutive trials.

### 208 **Materials**

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

### 219 **Data Scoring**

220 Test sessions were video-recorded and coded at a later date. For each trial, data were  
221 coded for whether or not the choice made was correct, latency of response (from maximum

## HUMAN REFERENTIAL PROBLEM SPACE

222 extension of the index-finger by the experimenter to the participant choosing a cup), type of  
223 response, the direction of the participant's gaze whilst giving the response, and whether the  
224 response was accompanied by a vocalisation. Response types were categorized according to  
225 the following scheme:

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

228         *Whole-hand point*: An indicative gesture categorized by outstretched arm and  
229 extended fingers, which is not a direct attempt to obtain the container (Leavens & Hopkins,  
230 1999).

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

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

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

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

### 241         **Analyses**

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

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

### 248 **Reliability**

249 An independent coder who was blind to the purpose of the study coded 20% of the  
250 videos (five participants, with sixteen trials each, for a total of 80 trials). For correct choices,  
251 There was complete agreement between the two coders for whether participants chose the  
252 correct cup on each trial, Cohen's kappa,  $\kappa = 1.00$ ,  $p < .001$ . There was excellent agreement  
253 for response latency,  $r_s = 0.8769$ ,  $p < .001$ , and substantial agreement for the type of response  
254 elicited from participants on each trial response type,  $\kappa = 0.66$ ,  $p < .001$  (Landis & Koch,  
255 1977). In the event of disagreement between coders on a specific trial, the original coding  
256 was maintained.

257

## 258 **Results**

### 259 **Correct Choices**

#### 260 **Eighteen-month-olds**

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

#### 267 **Thirty-six-month-olds**

268 The 36-month-olds also performed above chance as a group in both the *barrier*  
269 (binomial test,  $p < .001$ ) and the *no barrier* (binomial test,  $p < .001$ ) conditions. There was no  
270 significant difference in the proportion of correct trials between the *barrier* ( $Mdn = 1.00$ ) and  
271 the *no barrier* ( $Mdn = 1.00$ ) conditions,  $Z = -1.36$ ,  $p = .175$ ,  $r = -.08$ . This shows the older

272 children were also able to use the pointing cue despite the presence of a barrier.

### 273 **Age comparisons**

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

### 280 **Response Latency**

#### 281 **Eighteen-month-olds**

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

#### 286 **Thirty-six-month-olds**

287 There was no significant difference in response latency between the two conditions  
288 for the 36-month-olds (*barrier*  $Mdn = 2.88s$ , *no barrier*  $Mdn = 3.13s$ ),  $Z = -0.22$ ,  $p = .825$ ,  $r =$   
289  $-.01$ . This shows that the older children's response times were unaffected by the barrier.

### 290 **Age comparisons**

291 In the *barrier* condition, the 18-month-olds ( $Mdn = 13.00s$ ) were significantly slower  
292 to respond than the 36-month-olds ( $Mdn = 2.88s$ ), Mann-Whitney  $U = 18.0$ ,  $p < .001$ ,  $r =$   
293  $.27$ . The 18-month-olds ( $Mdn = 4.75s$ ) were also significantly slower to respond than the 36-  
294 month-olds ( $Mdn = 3.13s$ ) in the *no barrier* condition, Mann-Whitney  $U = 61.5$ ,  $p = .005$ ,  $r =$   
295  $-.18$ . This shows that the 18-month-olds were generally slower to respond than the older  
296 children, and these response times were further increased by the presence of a barrier in the

297 testing environment.

298 **Response Type**

299 **Eighteen-month-olds**

300 In the barrier condition, there was a significant difference in the proportion of the  
 301 types of responses elicited, Friedman's  $\chi^2(3) = 25.41, p < .001$ . Participants used significantly  
 302 more *index-finger points* than grabs,  $Z = -3.19, p = .001$ , and reaches,  $Z = -2.90, p = .005$ , and  
 303 significantly more *whole-hand points* than grabs,  $Z = -2.94, p = .003$ , and reaches,  $Z = -2.58,$   
 304  $p = .010$ . There were no other significant differences. In the no barrier condition, there was a  
 305 significant difference in the proportion of the different response types elicited, Friedman's  $\chi^2$   
 306  $(3) = 38.06, p < .001$ . Participants used significantly more index-finger points than whole-  
 307 hand points,  $Z = -2.55, p = .011$  and reaches,  $Z = -3.30, p = .001$ , and significantly more grabs  
 308 than index-finger points,  $Z = -3.30, p = .001$ , and reaches,  $Z = -3.45, p = .001$ . There were no  
 309 other significant differences.

310 There were a number of differences in the response types elicited from the younger  
 311 group as a function of the presence of a barrier. Eighteen-month-olds used significantly more  
 312 *index-finger points* in the *barrier* ( $Mdn = .43$ ) than in the *no barrier* ( $Mdn = .13$ ) condition,  $Z$   
 313  $= -3.18, p = .001, r = -.39$ , as well as significantly more *whole-hand points* in the *barrier*  
 314 ( $Mdn = .43$ ) than in the *no barrier* condition ( $Mdn = .00$ ),  $Z = -2.94, p = .003, r = -.43$ . They  
 315 grabbed the container significantly less in the *barrier* condition ( $Mdn = .00$ ) than in the *no*  
 316 *barrier* ( $Mdn = .88$ ) condition,  $Z = -3.45, p = .001, r = -.35$ . There was no significant  
 317 difference in 18-month-olds' tendency to reach for the container between the *barrier* ( $Mdn =$   
 318  $.00$ ) and *no barrier* ( $Mdn = .00$ ) conditions,  $Z = -1.34, p = .180, r = -.54$ . This shows that the  
 319 younger group were more likely to respond using a communicative cue such as an index-  
 320 finger point or a whole-hand point when there was a barrier present, and more likely to grab  
 321 the container when there was no barrier present. Analyses were not performed where

## HUMAN REFERENTIAL PROBLEM SPACE

322 responses were categorized as *other indicative gesture* or *other*, as these only constituted  
323 0.36% and 1.47% of the total responses (for both age groups combined), respectively. Figure  
324 2a shows the distribution of the response types for the younger children.

### 325 **Thirty-six-month-olds**

326 There was no significant difference in the proportion of the different response types  
327 elicited in the barrier condition, Friedman's  $\chi^2(3) = 3.08, p = .380$ . In the no barrier  
328 condition, there was a significant difference in the proportion of the different response types  
329 elicited, Friedman's  $\chi^2(3) = 47.24, p < .001$ . Participants used significantly more index-finger  
330 points than whole-hand points,  $Z = -2.11, p = .035$ , and reaches,  $Z = -2.69, p = .007$ , and  
331 significantly more grabs than index-finger points,  $Z = -3.60, p < .001$ , and reaches,  $Z = -$   
332  $3.89, p < .001$ .

333 There was no significant difference in the proportion of responses that were index-  
334 finger points for the 36-month-olds between the *barrier* ( $Mdn = .00$ ) and *no barrier* ( $Mdn =$   
335  $.00$ ) conditions,  $Z = -1.83, p = .066, r = -.24$ , but they did use significantly more whole-hand  
336 points when the barrier was present ( $Mdn = .00$ ) than when it was not ( $Mdn = .00$ ),  $Z = -2.20,$   
337  $p = .028, r = -.44$ . The 36-month-olds grabbed significantly more in the *no barrier* condition  
338 ( $Mdn = 1.00$ ) than in the *barrier* condition ( $Mdn = .38$ ),  $Z = -3.24, p = .001, r = -.24$ . Thirty-  
339 six-month-olds were significantly more likely to reach in the *barrier* ( $Mdn = .00$ ) than in the  
340 *no barrier* condition ( $Mdn = .00$ ),  $Z = -2.69, p = .007, r = -.51$ . This shows that the older  
341 children were also more likely to use some communicative gestures when the barrier was  
342 present and again, more likely to grab, or try to grab, the container when the barrier was  
343 absent. Figure 2b shows the distribution of response types for the older group.

### 344 **Age comparisons**

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

## HUMAN REFERENTIAL PROBLEM SPACE

347 80.5,  $p = .030$ ,  $r = -.23$ , but there were no significant effects of age on use of *index-finger*  
348 *points* in the *no barrier* condition, Mann-Whitney  $U = 125.5$ ,  $p = .560$ ,  $r = -.13$ . This shows  
349 that there were effects of both age and barrier on the use of this type of response, with 18-  
350 month-olds using *index-finger points* to indicate the container in which they thought the  
351 reward was hidden more than the 36-month-olds in the barrier condition.

352 The proportion of 18-month-olds' responses that were *whole-hand points* was  
353 significantly greater than that of 36-month-olds' in the *barrier* condition, Mann-Whitney  $U =$   
354  $81.0$ ,  $p = .022$ ,  $r = -.27$ , but not in the *no barrier* condition, Mann-Whitney  $U = 140.5$ ,  $p =$   
355  $.945$ ,  $r = -.12$ . This shows that children from both age groups were more likely to indicate  
356 their choice using a *whole hand point* when there was a barrier in the testing environment  
357 than when there was not, and that this effect of the barrier was particularly pronounced for  
358 the 18-month-olds.

359 Eighteen-month-olds were significantly less likely to grab in the *barrier* condition  
360 than the 36-month-olds, Mann-Whitney  $U = 64.5$ ,  $p = .006$ ,  $r = -.41$  but not in the *no barrier*  
361 condition, Mann-Whitney  $U = 133.0$ ,  $p = .758$ ,  $r = -.02$ . This shows that both age groups  
362 tended to grab the container in which they thought the reward was hidden more when there  
363 was no barrier present in the testing environment than when there was, but that this effect was  
364 less pronounced for the 36-month-olds. There was no significant effect of age on reaching  
365 behaviors in either the *barrier* condition, Mann-Whitney  $U = 95.5$ ,  $p = .050$ ,  $r = -.34$ .  
366 Participants from neither age group reached in the *no barrier* condition. This shows that the  
367 36-month-olds were more likely to reach for the container in which they thought the reward  
368 was hidden in the barrier condition, however the 18-month-olds were not.

### 369 Discussion

370 There were no differences in performance (correct responding) as a function of either  
371 age or the imposition of a barrier. That both groups of children demonstrated ceiling-level

## HUMAN REFERENTIAL PROBLEM SPACE

372 performance when there was no barrier present was expected; however, it is interesting to  
373 find that the imposition of a barrier did not have a decreasing effect on success levels for the  
374 human children in the way that Kirchhofer et al. (2012) found for domestic dogs. This shows  
375 that human children from 18 months are reliably and flexibly able to follow index-finger  
376 pointing cues, even with a partial visual barrier, although it must be noted that the two studies  
377 differed procedurally in terms of the distances between the containers (1.5m in Kirchhofer et  
378 al.) and the locomotor demands on the participants in retrieving the rewards (dogs in  
379 Kirchhofer et al.'s study were required to retrieve the object, turn around, and locomote to  
380 give it to the experimenter).

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

396         There were differences in the types of response produced by the children as a function



## HUMAN REFERENTIAL PROBLEM SPACE

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

422 (2005), that the children see the bars of the playpen as a barrier between themselves and a  
423 desirable, but out-of-reach object, and thus use a communicative gesture in order to influence  
424 another to retrieve said desirable object. That the object itself was not actually out of reach,  
425 but was instead simply partially obstructed, has interesting implications for the way the  
426 children perceived the barrier, perhaps as a form of psychological restraint.

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

447 experience of, and exposure to, index-finger points, are responding in a more  
448 conventionalized manner than the 18-month-olds, in terms of the shape of the hand when  
449 gesturing (Leavens & Hopkins, 1999).

### 450 **Study 2 Manipulating the distance of the reward**

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

### 462 **Method**

#### 463 **Participants**

464 Participants were seventeen 36-month-olds ( $M = 36$  mos 4days,  $range = 31$  mos 30  
465 days – 39 mos 26 days), comprised of 6 males ( $M = 37$  mos 4 days,  $range = 31$  mos 30 days  
466 – 39 mos 19 days) and 11 females ( $M = 36$  mos 11 days,  $range = 32$  mos 27 days – 39 mos  
467 26 days). Participants were recruited from a participant database, where parents had  
468 registered their interest in participating with their children in cognitive studies, and from  
469 advertisements on social media sites. Data were collected between December 2017 and  
470 January 2018.

#### 471 **Procedure**

## HUMAN REFERENTIAL PROBLEM SPACE

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

### 492           **Data Scoring**

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

### 494           **Analysis**

495           Five participants were excluded from the final analyses due to experimenter error  
496 during testing. Three of these cases were due to the experimenter using a momentary, rather

497 than a dynamic, point, and two were because the cameras were placed such that the  
498 experimenter was not in shot in the videos, and therefore the moment of pointing could not be  
499 ascertained.

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

### 505 Results

#### 506 Correct Choices

507 The data were not normally distributed and so non-parametric tests were used  
508 throughout the analyses. There was no significant difference in the proportion of correct  
509 choices made in the *proximal* ( $Mdn = 1.00$ ) and *distal* ( $Mdn = 1.00$ ) conditions,  $Z = -1.00$ ,  $p =$   
510  $.317$ ,  $r = -.15$ . This shows that the children performed at ceiling level in both conditions, as  
511 expected.

#### 512 Response Latency

513 There was no significant difference in the mean latency to respond between the  
514 *proximal* ( $Mdn = 3.00$  secs) and *distal* ( $Mdn = 3.75$  secs) conditions,  $Z = 0.00$ ,  $p = 1.00$ ,  $r =$   
515  $0$ . This shows that the children were equally quick to respond in both conditions.

#### 516 Type of Response

517 Only two of the possible response types were used by the children, these were index-  
518 finger points and grabs. Figure 4 shows the mean proportion of each type of response used in  
519 the two conditions. The proportion of trials in which the children used an index-finger point  
520 to indicate their choice of container was significantly lower in the *proximal* condition ( $Mdn =$   
521  $1.00$ ) than in the *distal* condition ( $Mdn = 1.00$ ),  $Z = -2.24$ ,  $p = .025$ ,  $r = -.35$ . This shows that

522 when the containers were out of reach, the children were more likely to respond by using an  
523 index-finger point than when the containers were within reach.

524 The proportion of trials in which the children grabbed the container was significantly  
525 higher in the *proximal* ( $Mdn = 1.00$ ) than in the *distal* condition ( $Mdn = .00$ ),  $Z = -2.24$ ,  $p =$   
526  $.025$ ,  $r = -1.00$ . This shows that the children were more likely to indicate their choice by  
527 grabbing the container when the containers were within reach.

### 528 **Order of administration and grabbing behavior**

529 In order to investigate whether the order of presentation had an effect on the response  
530 types used, participants were categorized as ‘grabbers’ (grabbed the container on at least one  
531 trial in the *proximal* condition or ‘non-grabbers’ (did not grab on either trial in the proximal  
532 condition). There was no significant effect of order of administration on the likelihood of  
533 grabbing (*proximal* first  $Mdn = 1.00$ ; *distal* first  $Mdn = .00$ ), Mann-Whitney  $U = 9.00$ ,  $p =$   
534  $.093$ ,  $r = -.25$ . There was a trend, however, such that those tested with the *proximal* condition  
535 first were more likely to grab in the *proximal* condition, whereas those tested with the *distal*  
536 condition first were slightly less likely to grab at all in the *proximal* condition (see Figure 5).

### 537 **Discussion**

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

543 The children in this study used only two response types to indicate the container in  
544 which they thought the reward was hidden, index-finger points and direct grabs of the  
545 container. These were the responses most often utilized by the 3-year-olds in Study 1,  
546 however, it is notable that there was an absence of the use of whole-hand points and reaches

## HUMAN REFERENTIAL PROBLEM SPACE

547 in the current study. The absence of whole-hand points is congruent with Cochet and  
548 Vauclair's (2010) findings that the incidence of whole-hand pointing in a declarative context  
549 decreases with age and that, here, the children were responding to the experimenter by  
550 demonstrating where they believed the reward to be hidden, rather than demanding the cup in  
551 an imperative manner. As in Study 1, the children chose to grab the container to look inside  
552 themselves in the *proximal* condition on a number of trials. That they did not try to reach for  
553 the container on any of the *distal* trials, nor did they exhibit any whole-hand points - which  
554 Cochet and Vauclair (2010) argued could be the result of failed grasping attempts in an  
555 imperative context- demonstrates that 3-year-olds are not categorically 'grabbers', but rather  
556 that their grabbing responses are a result of a spatial evaluation. When the container is not in  
557 reach, they do not try to grab it.

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

572

573

### General Discussion

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

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



## HUMAN REFERENTIAL PROBLEM SPACE

597 communication as a kind of social tool use—the use of an agent in goal-directed activities  
598 (e.g., Bard, 1990; Bates, Camaioni, & Volterra, 1975; Leavens, Hopkins, & Bard, 1996).  
599 Notably, in Study 1, the 18-month-olds displayed a virtually complete shift from a praxic  
600 mode to a communicative mode in the Barrier condition, whereas the 3-year-olds displayed a  
601 mixture of tactics in the Barrier condition, with some still grabbing the containers through the  
602 bars of the playpen. This demonstrates that, in representatives of our own species, life history  
603 stage influences the size of the palette of response forms, with older children displaying a  
604 proportionately lower propensity to use communication in the face of a permeable barrier,  
605 relative to younger children. In our studies, both praxic and communicative responses were  
606 deemed “correct,” but it is commonplace to present adult nonhumans with versions of an  
607 OCT in which only communicative responses or only praxic responses were considered  
608 “correct” (e.g., Call & Tomasello, 1994; Hopkins, Russell, McIntyre, & Leavens, 2013;  
609 Povinelli et al., 1999). Thus, when presenting similar tasks across representatives of different  
610 species, it is important to consider how response requirements or expectations interact with  
611 life history stage and pre-experimental life experience, to ensure that organisms are given the  
612 best opportunity to display their cognitive competencies; for example, it is well-demonstrated  
613 that enculturated apes significantly outperform institutionalised apes in similar experimental  
614 contexts (Lyn et al., 2010; Russell et al., 2011; and see, in a different context, arguments by  
615 Horowitz, 2003, and Thomas, Murphy, Pitt, Rivers, & Leavens, 2008, to the effect that  
616 younger humans are not representative of adult humans in some cognitive assays). Thus,  
617 because previous cross-species comparisons have generally not controlled for life history  
618 stage or task-relevant pre-experimental experience (Clark et al., 2019; Krause et al., 2018;  
619 Leavens et al., 2019), and because the present study shows a developmental shift in human  
620 children toward a reduced reliance on the use of communication in the presence of a barrier,  
621 therefore, we recommend a systematic revision to the OCT that permits both communicative

## HUMAN REFERENTIAL PROBLEM SPACE

622 and praxic responses. This adjustment will foster best performance to be captured and reduce  
623 the existing bias towards false negatives in some testing circumstances.

624         In conclusion, here we add to and extend the arguments put forward in recent reviews  
625 that detail the procedural and methodological flaws in the OCT literature (Lyn, 2010;  
626 Mulcahy & Hedge, 2012), and analyses of ape-human comparisons more generally (Leavens,  
627 2014, 2018; Leavens et al., 2019) and emphasize the necessity of ensuring matched  
628 conditions, as well as selection of correct response patterns that are appropriate to life history  
629 stage, in experimental testing. Furthermore, we demonstrate that these subtle manipulations  
630 of the testing environment can lead to large differences in the behavioral responses of  
631 members of the taxonomic group most experienced in the use of human gestural cues. This  
632 significantly challenges theories that rely on generalizations of the ability of representatives  
633 of a single sample to their entire species in studies which fail to adequately control for testing  
634 environment.

635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659

## 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*

- 660 *Psychology*, 30, 359-375. <https://doi.org/10.1111/j.2044-835X.2011.02043.x>
- 661 Bräuer, J., Kaminski, J., Riedel, J., Call, J. & Tomasello, M. (2006) Making inferences about  
 662 the location of hidden food: Social dog, causal ape. *Journal of Comparative*  
 663 *Psychology*, 120(1), 38-47. <http://dx.doi.org/10.1037/0735-7036.120.1.38>
- 664 Butterworth, G. (2001) Joint visual attention in infancy. In J. G. Bremner & A. Fogel (Eds.),  
 665 *Blackwell handbook of infant development* (pp. 213-240). Hove: Blackwell.
- 666 Butterworth, G. & Morissette, P. (1996) Onset of pointing and the acquisition of language in  
 667 infancy. *Journal of Reproductive and Infant Psychology*, 14(3), 219-231.  
 668 <https://doi.org/10.1080/02646839608404519>
- 669 Call, J., & Tomasello, M. (1994). Production and comprehension of referential pointing by  
 670 orangutans (*Pongo pygmaeus*).” *Journal of Comparative Psychology*, 108, 307-317.
- 671 Clark, H., Elsherif, M., & Leavens, D. A. (2019). Ontogeny vs. phylogeny in Primate/Canid  
 672 comparisons: An Object Choice Task review. *Neuroscience and Biobehavioural*  
 673 *Reviews, Advance online publication*.  
 674 <https://dx.doi.org/10.1016/j.neubiorev.2019.06.001>.
- 675 Cochet, H. & Vauclair, J. (2010) Pointing gestures produced by toddlers from 15 to 30  
 676 months: Different functions, hand shapes and laterality patterns. *Infant Behaviour and*  
 677 *Development*, 33, 431-441. <https://doi.org/10.1016/j.infbeh.2010.04.009>
- 678 D’Aniello, B., Alterisio, A., Scandurra, A., Petremolo, E., Iommelli, M. R. & Aria, M. (2017)  
 679 What’s the point? Golden and Labrador Retrievers living in kennels do not understand  
 680 pointing gestures. *Animal Cognition*, 20, 777-787. [https://doi.org/10.1007/s10071-](https://doi.org/10.1007/s10071-017-1098-2)  
 681 [017-1098-2](https://doi.org/10.1007/s10071-017-1098-2)
- 682 Duranton, C. & Gaunet, F. (2016) Effects of shelter housing on dogs’ sensitivity to human  
 683 social cues. *Journal of Veterinary Behaviour*, 14, 20-27.  
 684 <https://doi.org/10.1016/j.jveb.2016.06.011>.

## HUMAN REFERENTIAL PROBLEM SPACE

- 685 Franco, F. & Butterworth, G. (1996) Pointing and social awareness: Declaring and requesting  
686 in the second year. *Journal of Child Language*, 23, 307-336.  
687 <https://doi.org/10.1017/S0305000900008813>
- 688 Hare, B. & Tomasello, M. (2005) Human-like social skills in dogs? *TRENDS in Cognitive*  
689 *Sciences*, 9(9), 439-444. <https://doi.org/10.1016/j.tics.2005.07.003>
- 690 Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B. & Tomasello, M. (2007) Humans  
691 have evolved specialized skills of social cognition: The cultural intelligence  
692 hypothesis. *Science*, 317, 1360-1367. <https://doi.org/10.1126/science.1146282>
- 693 Hopkins, W. D., Russell, J. L., McIntyre, J., & Leavens, D. A. (2013). Are chimpanzees  
694 really so poor at understanding imperative pointing? Some new data and an  
695 alternative view of canine and ape social cognition. *PLoS ONE*, 8, e79338.  
696 <https://doi.org/10.1371/journal.pone.0079338>
- 697 Horowitz, A. C. (2003). Do humans ape? Or do apes human? Imitation and intention in  
698 humans (*Homo sapiens*) and other animals. *Journal of Comparative Psychology*, 117,  
699 325–336. <https://doi.org/10.1037/0735-7036.117.3.325>
- 700 Kirchhofer, K. C., Zimmermann, F., Kaminski, J. & Tomasello, M. (2012) Dogs (*Canis*  
701 *familiaris*), but not chimpanzees (*Pan troglodytes*), understand imperative pointing.  
702 *PLoS ONE*, 8(2), e30913. <https://doi.org/10.1371/journal.pone.0030913>
- 703 Kishimoto, T. (2013). Do referential problem spaces affect the frequency of imperative  
704 pointing by infants? *Psychology of Language and Communication*, 17, 295-307.  
705 <https://doi.org/10.2478/plc-2013-0019>
- 706 Krause, M. A., Udell, M. A. R., Leavens, D. A., & Skopos, L. (2018). Animal pointing:  
707 Changing trends and findings from 30 years of research. *Journal of Comparative*  
708 *Psychology*, 132, 326-345. <http://dx.doi.org/10.1037/com0000125>
- 709 Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical

- 710 data. *Biometrics*, 33, 159-174. <https://doi.org/10.2307/2529310>
- 711 Lazarowski, L. & Dorman, D. C. (2015) A comparison of pet and purpose-bred research dog  
712 (*Canis familiaris*) performance on human-guided object-choice tasks. *Behavioural*  
713 *Processes*, 110, 60-67. <https://doi.org/10.1016/j.beproc.2014.09.021>
- 714 Leavens, D. A. (2014). The plight of the sense-making ape. In M. Cappuccio & T. Froese  
715 (Eds.), *Enactive cognition at the edge of sense-making* (pp. 81-104). Basingstoke,  
716 U.K.: Palgrave Macmillan. [https://doi.org/10.1057/9781137363367\\_4](https://doi.org/10.1057/9781137363367_4)
- 717 Leavens, D. A. (2018). The cognitive implications of intentional communication: A multi-  
718 faceted mirror. In L. Di Paolo, F. Di Vincenzo, & F. De Petrillo (Eds.), *Evolution of*  
719 *primate social cognition* (pp. 59-77). New York: Springer.  
720 <https://doi.org/10.1007/978-3-319-93776-2>
- 721 Leavens, D. A. & Hopkins, W. D. (1999) The whole-hand point: The structure and function  
722 of pointing from a comparative perspective. *Journal of Comparative Psychology*,  
723 113(4), 417- 425. <http://dx.doi.org/10.1037/0735-7036.113.4.417>
- 724 Leavens, D. A., Bard, K. A., & Hopkins, W. D. (2019). The mismeasure of ape social  
725 cognition. *Animal cognition*, 22, 487-504. <https://doi.org/10.1007/s10071-017-1119-1>
- 726 Leavens, D. A., Hopkins, W. D., & Bard, K. A. (1996). Indexical and referential pointing in  
727 chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 110, 346-353.  
728 <https://doi.org/10.1037/0735-7036.110.4.346>
- 729 Leavens, D. A., Hopkins, W. D. & Bard, K. A. (2005) Understanding the point of  
730 chimpanzee pointing: Epigenesis and ecological validity. *Current Directions in*  
731 *Psychological Science*, 14(4), 185-189. [https://doi.org/10.1111/j.0963-](https://doi.org/10.1111/j.0963-7214.2005.00361.x)  
732 [7214.2005.00361.x](https://doi.org/10.1111/j.0963-7214.2005.00361.x)
- 733 Leung, E. H. L. & Rheingold, H. L. (1981) Development of pointing as a social gesture.  
734 *Developmental Psychology*, 17(2), 215-220. <http://dx.doi.org/10.1037/0012->

735 [1649.17.2.215](https://doi.org/10.1649.17.2.215)

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

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

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

745 Miklósi, A., & Soproni, K. (2006). A comparative analysis of animals' understanding of the  
746 human pointing gesture. *Animal Cognition*, 9(2), 81-93.

747 <https://doi.org/10.1007/s10071-005-0008-1>

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

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

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

756 Russell, J.L., Lyn, H., Schaeffer, J.A., Hopkins, W.D. (2011). The role of socio-  
757 communicative rearing environments in the development of social and physical  
758 cognition in apes. *Developmental Science*, 14, 1459-1470.

759 <https://doi.org/10.1111/j.1467-7687.2011.01090.x>

## HUMAN REFERENTIAL PROBLEM SPACE

- 760 van Schaik, C. P., & Burkart, J. M. (2011). Social learning and evolution: The cultural  
761 intelligence hypothesis. *Philosophical Transactions of the Royal Society, Series B*,  
762 366, 1008-1016. <https://doi.org/10.1098/rstb.2010.0304>
- 763 Thomas, E., Murphy, M., Pitt, R., Rivers, A., & Leavens, D. A. (2008). Understanding of  
764 visual attention by adult humans (*Homo sapiens*): A partial replication of Povinelli,  
765 Bierschwale, and Čech (1999). *Journal of Comparative Psychology*, 122, 428-436.  
766 <https://doi.org/10.1037/0735-7036.122.4.428>
- 767 Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of novel communicative  
768 signs by apes and human children. *Child Development*, 68(6), 1067- 1080.  
769 <https://doi.org/10.1111/j.1467-8624.1997.tb01985.x>
- 770 Udell, M. A. R., Dorey, N. R. & Wynne, C. D. L (2008) Wolves outperform dogs in  
771 following human social cues. *Animal Behaviour*, 76, 1767-1773.  
772 <https://doi.org/10.1016/j.anbehav.2008.07.028>
- 773 Udell, M. A. R., Giglio, R. F., & Wynne, C. D. L. (2008). Domestic dogs (*Canis familiaris*)  
774 use human gestures but not nonhuman tokens to find hidden food. *Journal of*  
775 *Comparative Psychology*, 122, 84–93. <http://dx.doi.org/10.1037/0735-7036.122.1.84>
- 776 Udell, M. A. R., Spencer, J. M., Dorey, N. R. & Wynne, C. D. L. (2012) Human-socialized  
777 wolves follow diverse human gestures... And they may not be alone. *International*  
778 *Journal of Comparative Psychology*, 25, 97-117.  
779 <https://doi.org/10.1016/j.anbehav.2008.07.028>
- 780



781  
782  
783  
784  
785  
786  
787  
788

**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.

789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808

**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.

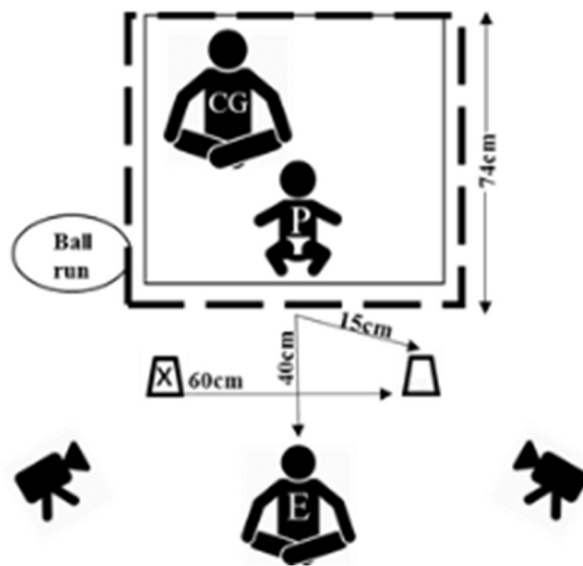
*Figure 4.* The mean proportion of response types, with standard errors, in the *proximal* and *distal* conditions. Means and standard errors are depicted, here, to more clearly display the effects, although nonparametric statistical tests were applied. IFP = index finger point. \*  $p < .05$ .

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

# HUMAN REFERENTIAL PROBLEM SPACE

809 Figure 1

810



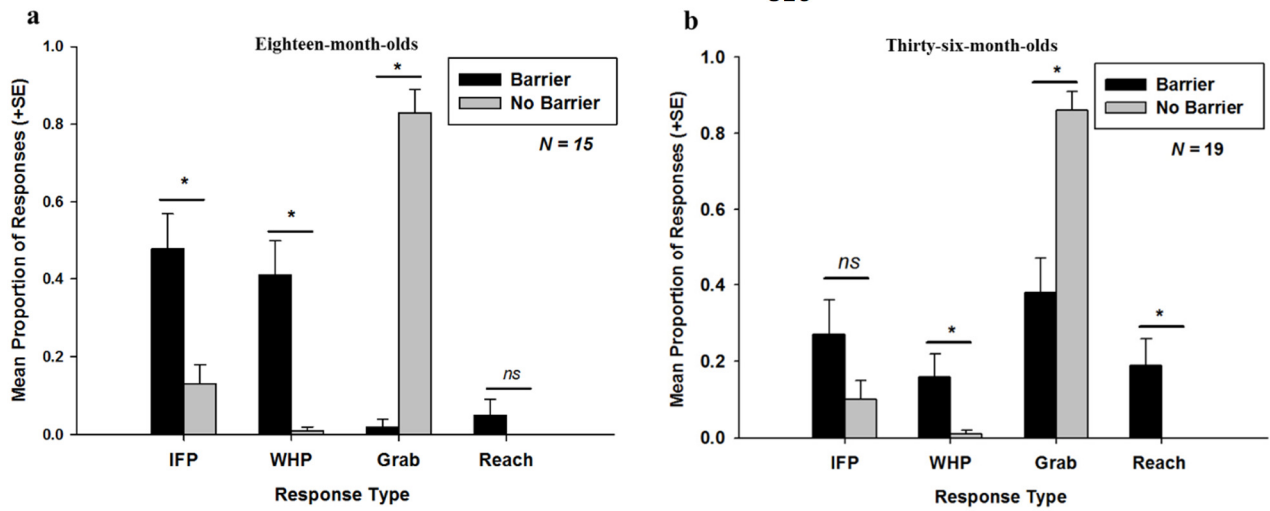
811

812

813

814 Figure 2

815



817

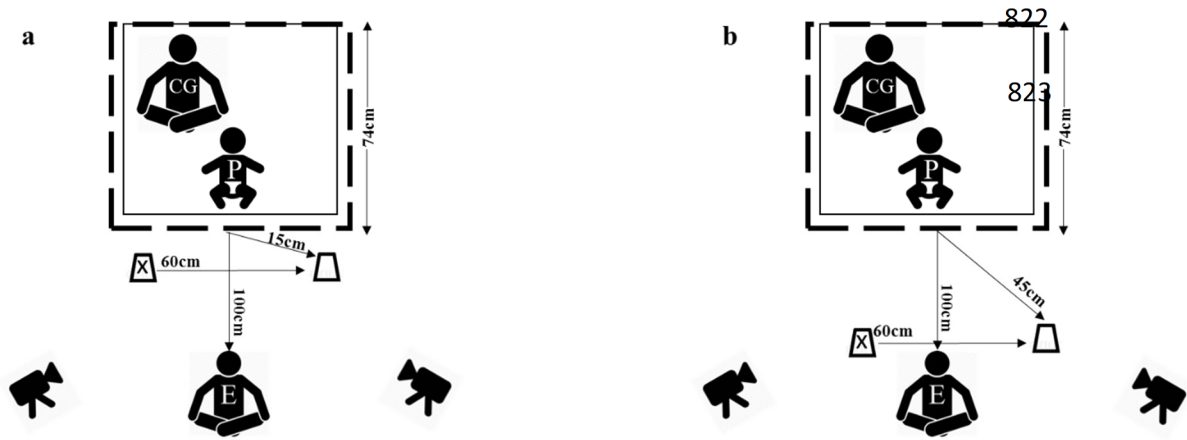
818

819

# HUMAN REFERENTIAL PROBLEM SPACE

820 Figure 3

821

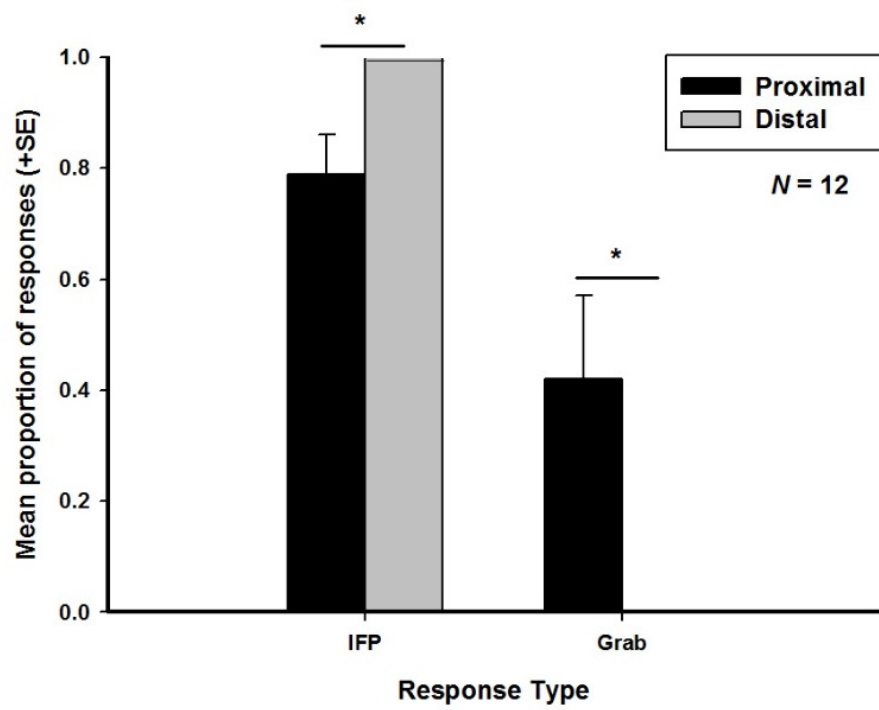


# HUMAN REFERENTIAL PROBLEM SPACE

824 Figure 4

825

826



# HUMAN REFERENTIAL PROBLEM SPACE

827 Figure 5

828

829

