

1 Title

2 Human ostensive signals do not enhance gaze following in chimpanzees but do
3 enhance object-oriented attention

4
5 Fumihiro Kano* (1)

6 Richard Moore (2)

7 Christopher Krupenye (3, 4)

8 Satoshi Hirata (1)

9 Masaki Tomonaga (5)

10 Josep Call (3, 4)

11
12 Affiliations:

- 13 1. Kumamoto Sanctuary, Wildlife Research Center, Kyoto University, Kumamoto, Japan
14 2. Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Germany
15 3. Department of Developmental and Comparative Psychology, Max Planck Institute for
16 Evolutionary Anthropology, Leipzig, Germany
17 4. School of Psychology and Neuroscience, University of St. Andrews, St. Andrews, UK
18 5. Primate Research Institute, Kyoto University, Inuyama, Japan

19 *Correspondence: fkanou@gmail.com, Kumamoto Sanctuary, Kyoto University, 990 Misumi, Uki,
20 Kumamoto, 8693201, Japan

21 Abstract

22 Previous studies have shown that human infants and domestic dogs follow the gaze of a human agent only
23 when the agent has addressed them ostensively – e.g. by making eye contact, or calling their name. This
24 evidence is interpreted as showing that they expect ostensive signals to precede referential information.
25 The present study tested chimpanzees, one of the closest relatives to humans, in a series of eye-tracking
26 experiments using an experimental design adapted from these previous studies. In the ostension
27 conditions, a human actor made eye contact, called the participant’s name, and then looked at one of two
28 objects. In the control conditions, a salient cue, which differed in each experiment (a colorful object, the
29 actor’s nodding, or an eating action), attracted participants’ attention to the actor’s face, and then the actor
30 looked at the object. Overall, chimpanzees followed the actor’s gaze to the cued object in both ostension
31 and control conditions, and the ostensive signals did not enhance gaze following more than the control
32 attention-getters. However, the ostensive signals enhanced subsequent attention to both target and
33 distractor objects (but not to the actor’s face) more strongly than the control attention-getters—especially
34 in the chimpanzees who had a close relationship with human caregivers. We interpret this as showing that
35 chimpanzees have a simple form of communicative expectations on the basis of ostensive signals, but
36 unlike human infants and dogs, they do not subsequently use the experimenter’s gaze to infer the intended
37 referent. These results may reflect a limitation of non-domesticated species for interpreting humans’
38 ostensive signals in inter-species communication.

39

40 Key words: Domestication; Gaze following; Great ape; Ostensive signals; Referential communication

41

42 Introduction

43 Humans use ‘ostensive’ signals such as making eye contact and calling an intended recipient’s name as a
44 way of indicating their communicative intention to others. These signals function to alert an intended
45 recipient to the possibility that a signaler has a message to convey (Csibra 2010; Moore 2016; Moore
46 2017; Scott-Phillips 2015a; Scott-Phillips 2015b; Sperber and Wilson 1995), and so provide the recipient
47 with evidence that they should devote their cognitive resources to figuring out the content of that message.
48 The ability to determine, on the basis of non-verbal cues, when others are acting with communicative
49 intention has been argued to play a fundamentally important role in both language acquisition and cultural
50 learning more generally (e.g. Gergely and Csibra 2006; Tomasello 1999; Tomasello 2006).

51 Csibra and Gergely have suggested that humans possess an adaptation for ‘natural pedagogy’,
52 which explains how humans efficiently transmit generic knowledge between individuals (Csibra 2010;
53 Csibra and Gergely 2009; Gergely and Csibra 2006; Gergely et al. 2007). According to their proposal,
54 human infants have a set of perceptual and cognitive biases that make them interpret ostensive signals as
55 indicating that an agent is trying to deliver generic information. These biases are: (1) Preferential attention
56 for the source of ostensive signals: Human infants are highly sensitive to the presence of signals (e.g. eye
57 contact and infant-directed speech) that indicate that they are being addressed by a communicating agent.
58 (2) Referential expectation: Following ostensive signals, human infants expect to find the intended
59 referent of the communicating agent’s message – that is, the entity about which the agent is
60 communicating. (3) Generality: Infants take ostensive communication to provide them with generic
61 information about objects like those to which the agent referred – that is, information that is generalizable
62 to objects in other situations.

63 Although the nature of the mechanisms at work in Csibra and Gergely’s hypothesis remain
64 controversial (Heyes 2016; Hoicka 2016; Moore et al. 2013; Moore et al. 2015b), its functional aspects
65 are supported by compelling evidence. Related to the first feature of their proposal, it is well established
66 that human infants have early-developing preference for attending to faces (Farroni et al. 2002) and

67 infant-directed speech (Cooper and Aslin 1990). Additionally, in relation to the second feature, Senju and
68 Csibra (2008) found that 6-month-old humans followed an experimenter's gaze to an intended referent
69 when it was preceded by either directed speech or ostensive eye contact, but not when a similarly salient
70 animation was used to solicit their attention. This suggests that ostensive signals may help infants to
71 identify an agent's referential goals, and thereby better understand referential communication. Finally,
72 related to the third feature, Topál et al. (2009) found that 9-month-old humans (and domestic dogs, but
73 not wolves) made more frequent search errors in the A-not-B task when an agent ostensively hid the
74 objects compared to when the agent just hid the objects without any ostensive signals. That is, in the
75 ostension condition, infants (and dogs) persistently searched for a hidden object at its initial hiding place
76 even after observing it being hidden at another location. This finding is interpreted as showing that, on the
77 basis of ostensive signals, the infants (and dogs) had formed general expectations about where the objects
78 would be hidden, and these expectations trumped their own experience of seeing them hidden (but see
79 Vorms (2012) for criticism of this interpretation).

80 While the adaptation for natural pedagogy was initially proposed to be uniquely human, more
81 recent studies suggest that domesticated dogs are attuned to humans' ostensive signals in ways that are
82 similar to human infants. Dogs spontaneously attend to human faces in a variety of communicative
83 contexts (Topál et al. 2014), and they use both eye contact and (to a lesser extent) name-calling to identify
84 that an experimenter is communicating with them (Kaminski et al. 2012). Additionally, Téglás et al.
85 (2012) showed that, using the experimental paradigm developed for human infants (Senju and Csibra
86 2008), dogs followed an experimenter's gaze to a referent only if it had been preceded by ostensive eye
87 contact and directed speech. Thus, similar to human infants, dogs may have the expectation that humans'
88 ostensive signals precede referential information. These results support the domestication hypothesis,
89 which postulates that, likely as a consequence of their domestication, domestic dogs have evolved
90 adaptive responses to human referential communication, in a manner similar to human infants. This
91 domestication hypothesis gains support particularly from comparative studies with human infants,

92 domestic dogs and non-domesticated species such as wolves and great apes. Domestic dogs, even from a
93 few weeks old, outperform chimpanzees and wolves at a number of tasks in which they must read human
94 communicative signals such as gazing and pointing to locate hidden food (Hare et al. 2002; Riedel et al.
95 2008; Topál et al. 2014). Virányi et al. (2008) compared hand-raised four-month-old wolf and dog
96 puppies, finding that the dogs were both more willing to maintain eye-contact with experimenters, and
97 better able to use the experimenters' points to find hidden food. While wolves were able to learn to
98 respond to ostensive signals after training, the results suggest that dogs possess an early-developing
99 responsiveness to human communication that wolves do not.

100 If human infants and domestic dogs follow human gaze with the expectation that ostensive
101 signals precede referential information (Senju and Csibra 2008; Téglás et al. 2012), an interesting
102 question is whether nonhuman great apes would do so as well. Although great apes (chimpanzee, bonobo,
103 gorilla and orangutan) have not been domesticated by humans, they are humans' closest living relatives
104 and thus are much more closely related to humans phylogenetically than are dogs. In addition, they are
105 equipped with a variety of key skills that might enable them to act like human infants and dogs. Like
106 humans (and many other animals, Bugnyar et al. 2004; Emery 2000), great apes spontaneously follow
107 others' gaze. They do not just co-orient with others but take others' visual perspectives into consideration
108 when following their line of sight (Bräuer et al. 2005; Okamoto-Barth et al. 2007). Moreover, as in
109 humans, eye contact plays an important role in the natural repertoire of communicative behaviors in great
110 apes (Gomez 1996; Liebal and Call 2012). For example, when chimpanzees and gorillas attempt to
111 reconcile with conspecifics after fighting, they first establish eye contact before approaching their
112 counterparts (De Waal 1990; Yamagiwa 1992). When tension arises among individuals, bonobos regulate
113 it by making eye contact and engaging in sexual activities (De Waal 1988). Some apes even use their eye
114 contact ostensively when requesting food from human experimenters in a laboratory (Gomez 1996).

115 In light of the pervasive role of eye contact in great ape communication, great apes should satisfy,
116 at least at a functional level, the first feature of natural pedagogy hypothesis (Csibra and Gergely 2009):

117 the preference for the potential source of communicative signals. Previous studies have also shown that
118 great apes spontaneously attend to both conspecific and human faces (Kano et al. 2012; Kano and
119 Tomonaga 2009). Like human infants (Farroni et al. 2002), chimpanzee infants preferentially attend to
120 human faces with direct gaze, rather than those with averted gaze (Myowa-Yamakoshi et al. 2003).
121 Chimpanzees are less accurate in distinguishing between the faces of humans and conspecifics when
122 those faces are presented upside-down compared to when they are presented upright (Parr et al. 1998;
123 Tomonaga 1999). Finally, for apes living in a typical zoo and research environment, human caregivers
124 and experimenters regularly call apes' individual names (and make eye contact with them) when
125 attempting to communicate with them. A previous study measuring a chimpanzee's event-related
126 potentials showed that the chimpanzee became attentive immediately after they heard their own names
127 called by a human experimenter (Ueno et al. 2010).

128 However, it remains unclear whether great apes have referential expectations following ostensive
129 signals, or whether they understand others' intentions to communicate about specific referents. One study
130 reported that, when chimpanzees saw that another individual is requesting a particular item, they could
131 infer the item that the other was requesting, even on the basis of ambiguous gestures (Yamamoto et al.
132 2012). Moreover, while it was once doubted that great apes could understand others' intentions at all,
133 recent evidence challenges this view. Great apes understand others' goals and intentions and utilize that
134 knowledge in various social contexts (Call and Tomasello 2008; Kano and Call 2014b). A recent eye-
135 tracking study even showed that, when apes are viewing an agent and an antagonist competing for an
136 object, they anticipate the agent's actions according to the agent's false beliefs (Kano et al. 2017;
137 Krupenye et al. 2016).

138 Nonetheless, in general, great apes do not appear to be as sensitive to the referential aspects of
139 human communication as human infants and domestic dogs. For example, in studies where human
140 experimenters try to inform them of the location of hidden food with referential gestures such as gaze and
141 pointing, both human infants (Behne et al. 2005; Behne et al. 2012) and domestic dogs (Hare and

142 Tomasello 1999; Miklósi et al. 1998) excel at using such experimenters' cues to locate food. However,
143 great apes perform comparatively poorly in similar paradigms (Hare and Tomasello 2004; Herrmann and
144 Tomasello 2006), although enculturated apes (those reared by humans in human environments) generally
145 perform better than unenculturated apes (Call and Tomasello 1994; Call and Tomasello 1996; Lyn et al.
146 2010). Previously only one study (Moore et al. 2015a) has addressed whether orangutans respond
147 differently to pointing gestures produced with and without ostensive signals. In this study, no effect of
148 ostension was found. However, since comprehension was poor in all conditions, this finding is difficult to
149 interpret.

150 A number of studies have tested great apes' gaze following behavior. In several previous
151 experiments (Call et al. 2000; Tomasello et al. 2007), the human experimenter called the apes' names and
152 made eye contact before providing a gaze cue. Such cues did seem to work, at least in drawing the apes'
153 attention to the experimenter's face (before the gaze cue was provided). In an eye-tracking experiment
154 presenting still images to chimpanzees, Hattori et al. (2010) showed that chimpanzees followed the gaze
155 of a conspecific agent but not that of a human agent. In another eye-tracking experiment (Kano and Call
156 2014a), while bonobos, orangutans and human adults all followed the gaze of both conspecific and
157 allospecific agents, human infants and chimpanzees only followed the gaze of conspecific agents. These
158 findings indicate that at least chimpanzees may not be receptive to following a human experimenter's
159 gaze when watching still pictures and movies without clear ostensive cues. None of these studies directly
160 compared the effect of humans' ostensive signals with that of control attention-getters on gaze following
161 in great apes.

162 In this study, we examined whether great apes, particularly chimpanzees, would exhibit enhanced
163 gaze following in response to a human actor establishing ostensive eye contact and calling the
164 participant's name. We designed our experiment based on previous eye-tracking studies of human infants
165 and dogs (Senju and Csibra 2008; Téglás et al. 2012). According to the domestication hypothesis and
166 previous evidence that apes are not good at comprehending human referential communication, they may

167 not understand a human actor's ostensive signals in the same way as human infants and dogs do. On the
168 other hand, based on previous evidence that apes make eye contact with conspecifics and humans, and are
169 familiar with humans calling their names, such signals should have some effect on apes' behavior. While
170 it is possible that humans' ostensive signals would enhance gaze following in great apes just as in human
171 infants and dogs, it is also possible that ostensive signals function in a more limited way – for example,
172 by enhancing attention to the actor's face or to all objects in front of the actor (i.e. to both cued and non-
173 cued objects).

174 The question we address here is whether chimpanzees would be more likely to follow the gaze of
175 a human actor after the actor addressed them ostensively as compared with after a control attention-getter
176 attracted their attention to the actor's face. We specifically tested 1) whether the actor's ostensive signals
177 attracted apes' attention to the actor's face in the cueing phase (when the actor was addressing either
178 ostensively or non-ostensively) as strongly as the control cues in the cueing phase, and 2) whether the
179 ostensive signals enhanced the chimpanzees' gaze following (i.e. their looking at the cued object),
180 attention to the objects (i.e. their looking at both cued and non-cued objects in front of the actor), and/or
181 attention to the actor's face more strongly than the control cues in the looking phase (when the actor was
182 looking at one of the objects).

183 We tested chimpanzees from different rearing backgrounds: zoo-reared and institute-reared
184 individuals. The institute-reared individuals (who were similar but technically not identical to
185 'enculturated' individuals; see Method) had richer early experiences of interacting with human caregivers
186 than the zoo-reared individuals. It is well established that such individuals perform better than the other
187 individuals in tests in which they need to locate hidden foods based on an experimenter's referential cues
188 (Call and Tomasello 1994; Call and Tomasello 1996; Lyn et al. 2010). Also, a previous eye-tracking
189 study found that the institute-reared chimpanzees paid more attention to the objects manipulated by
190 conspecifics than the zoo-reared chimpanzees (Kano et al. 2018). It was thus expected that the institute-
191 reared chimpanzees would show greater sensitivity to human ostension than the zoo-reared chimpanzees

192 in our test. We also tested two other closely related great ape species, bonobos and orangutans, in this
193 study, but mainly focus on the results for the chimpanzees in this article. We do so primarily for the
194 simplicity of analyses and reports: although we found that bonobos and orangutans were similar to
195 chimpanzees in terms of the key results, there were differences between the three species in their basic
196 responses to the human agent's gaze cues (but not to the ostensive *and* gaze cues). We discuss the bonobo
197 and orangutan results briefly in the main text and report them in greater detail in the Supplemental
198 Materials.

199 We used the actor's eye contact and calling of the participant ape's name as ostensive cues. One
200 potential methodological issue is that it remains unclear from previous studies what control cues are
201 appropriate for great apes. In particular, it remains unclear to what extent different control cues attract
202 apes' attention in comparison to the ostensive cues. Previous studies with human infants and dogs (Senju
203 and Csibra 2008; Téglás et al. 2012) used a visually salient object (presented on the top of the actor's
204 head) as a control for eye contact and a low-pitched, adult-directed voice as a control for a high-pitched,
205 infant/dog-directed voice. We could not use a low-pitched voice as a control because in their daily lives
206 the apes we tested often hear their names in a low pitch. We were also uncertain which visual cues could
207 be used as a control because of debates over the extent to which different control stimuli attracted covert
208 and overt attention to the actor's face in previous studies (see Szufnarowska et al. 2014 and also the
209 comments by Senju and Csibra in the same journal). We thus examined the effect of control cues using
210 several different attention-getters in Experiment 1-3 in order to explore to what extent the use of different
211 control cues altered the chimpanzees' responses to the actor's ostensive and gaze cues. Experiments 1-3
212 used the same design but differed in the types of control cue used. Experiment 1 used the actor's head
213 gesture with a voice (unrelated to the participant ape) as a control (partly following Szufnarowska et al.
214 2014). Experiment 2 used a visually salient object with an artificial sound on the actor's face as control
215 (following Senju and Csibra 2008; Téglás et al. 2012). Experiment 3 used the actor's eating action with a
216 crunch sound and a voice as control (a control used for the first time in this study). As a result, we could

217 ensure that chimpanzees were similarly attentive to the human actor's ostensive and gaze cues across the
218 experiments.

219

220 Method

221 Participants

222 Experiment 1 tested 15 chimpanzees (*Pan troglodytes*) from the Wolfgang Koehler Primate Research
223 Center (WKPRC) in Leipzig, Germany. Experiment 2 tested 12 chimpanzees from the Kumamoto
224 Sanctuary (KS) in Kumamoto, Japan, and the Primate Research Institute (PRI) in Inuyama, Japan.
225 Experiment 3 tested the same 19 chimpanzees (5 from KS and 14 from WKPRC). We also tested 7
226 bonobos (*Pan paniscus*) and 7 orangutans (*Pan troglodytes*) from WKPRC in Experiment 1, 6 bonobos
227 from KS in Experiment 2, and the same 13 bonobos and 6 orangutans in Experiment 3 (see the
228 Supplemental Material for the results from these species). We did not exclude any apes in this study. All
229 apes were reared in captivity and lived with conspecifics in enriched naturalistic environments at WKPRC,
230 KS, and PRI. All apes had some experience watching naturalistic movies for enrichment and in
231 experiments, although they were never explicitly trained for their gaze behavior. The chimpanzees from
232 KS (recently moved from Hayashibara Great Ape Research Institute, Okayama, Japan) and PRI had
233 participated in numerous cognitive experiments since their youth. Consequently, they had more human
234 interaction experience than the chimpanzees from WKPRC (they are similar to, but technically not
235 ‘enculturated’ chimpanzees, as enculturated chimpanzees are typically defined as those reared by humans
236 in human environments in literature; our chimpanzees were reared by their biological mothers or mostly
237 by their conspecific peers in a chimpanzee group. See Table S1 for more details about each participant).

238 Ethics statement

239 All participants were tested in the testing rooms prepared for each species, and their daily participation in
240 this study was voluntary. They were given regular feedings, daily enrichment, and had ad libitum access
241 to water. Animal husbandry and research protocol complied with international standards (the Weatherall
242 report “The use of non-human primates in research”) and institutional guidelines (KS: Wildlife Research
243 Center “Guide for the Animal Research Ethics”; PRI: Primate Research Institute 2002 version of “The
244 Guidelines for the Care and Use of Laboratory Primates”; WKPRC: “EAZA Minimum Standards for the

245 Accommodation and Care of Animals in Zoos and Aquaria”, “WAZA Ethical Guidelines for the Conduct
246 of Research on Animals by Zoos and Aquariums”, “Guidelines for the Treatment of Animals in
247 Behavioral Research and Teaching” of the Association for the Study of Animal Behavior (ASAB))”.

248 Apparatus

249 An infrared head-free eye-tracker recorded their eye movements [60 Hz; X120 in WKPRC and X300 in
250 KS and PRI; Tobii Technology AB, Stockholm, Sweden]. The eye-tracker and monitor were installed
251 outside of the testing room. Apes watched the movies on the monitor through a transparent acrylic panel
252 (1-2 cm in thickness); we previously confirmed that this transparent acrylic panel does not interfere with
253 recording of eye movements (Kano et al. 2011). Apes were allowed to sip diluted grape juice via a
254 custom-made juice dispenser attached to the transparent acrylic panel (irrespective of their gaze behavior).
255 In both facilities, the movies were presented at a viewing distance of 70 cm with a resolution of 1280×720
256 pixel on a 23-inch LCD monitor (43×24 degree) with Tobii Studio software (version 3.2.1). Two-point
257 automated calibration was conducted for the apes by presenting a small object or movie clip on each
258 reference point. Each time before the recording session, we manually checked the accuracy and repeated
259 the calibration if necessary. Calibration errors were typically within a degree (Hirata et al. 2010; Kano et
260 al. 2011).

261 Stimuli and procedure

262 Each experiment had two conditions: an ostension condition and a control condition (Figure 1).
263 Experiment 1-3 used the same test (ostension) condition but differed in the types of control cues. The test
264 and control conditions of Experiment 1-2 were designed based on the previous studies with human infants
265 and domestic dogs (Senju and Csibra 2008; Szufnarowska et al. 2014; Téglás et al. 2012). In all
266 experiments, in the ostension condition, a human actor faced the participant and initially looked down,
267 with two identical objects (“still phase”), one on each side of him. After 2 seconds, the actor looked up,
268 made eye contact, and called the participant ape’s name twice (the actor opened his mouth twice, and
269 each participant’s unique name was dubbed into the mouth movements). This “cueing phase” lasted for

270 2.5 seconds. The actor then turned his head to one of the two same objects (“target;” the other object is
271 called “distractor”) and kept still for the remaining time (“looking phase;” 5 seconds). The objects were
272 supported by the actor’s hand while they rested on a table in Experiment 1-2, while they were mounted on
273 tripods to each side of the actor at eye level in Experiment 3 (see Results and Discussion for the reason
274 for these changes). The control condition was the same as the ostension condition except that the cueing
275 phase presented a non-communicative attention-getter instead of the actor’s communicative cue. As a
276 control attention-getter, in Experiment 1, the actor nodded 3 times (following Szufnarowska et al. 2014)
277 and made a “hmm” sound (said as if to himself), during this action (Figure 1A). The actor repeated this
278 action and sound twice during the cueing phase. In Experiment 2, a circle with red-white color patterns
279 (Figure 1B) rolled 360 degrees with a chime sound twice, once clockwise, and then counter-clockwise
280 (following Senju and Csibra 2008). In Experiment 3, the actor ate an apple with crunching and “hmm”
281 sounds (Figure 1C).

282 Each participant was tested in both conditions; each participant first completed all trials in one
283 condition over several days, before completing the other condition on subsequent days (i.e. within-subject
284 design; one concern regarding this design is a carry-over effect, but note that a supporting analysis on the
285 first 6 trials—which mimicked a between-subject design—yielded the same results; see Supplemental
286 Material). The order of presentation of the conditions was counterbalanced between participants. Each
287 condition had 6 trials; in 3 trials the actor looked at the left object in the looking phase and in the other 3
288 trials, the actor looked at the right object. The presentation order of the cued side was pseudo-randomized
289 so that no more than 2 successive trials cued the same side. Each trial presented one movie file. Each trial
290 presented a unique object as both target and distractor. Each participant received typically 2 trials a day
291 (max. 4 trials depending on the motivation of the participant). We did not exclude any trials based on apes’
292 performance or any other criteria.

293 Data analysis

294 The apes' eye-movement responses to each scene feature were coded automatically in the Tobii Studio
295 software based on Area-Of-Interest (AOI). Apes' eye movement was filtered (fixations were extracted)
296 using Tobii Fixation Filter in the same software. The AOIs were defined for the actor's face, the target
297 object the actor gazed at and the distractor object the actor did not gaze at. The AOIs were drawn about
298 20% larger than the actual size of the face/object to avoid any fixation error (see Figure S1 for the defined
299 AOIs). To check if the actor's ostensive signals equally attracted apes' attention, we examined the
300 viewing times for the actor's face during each phase (still, cueing, and looking) in each condition. We
301 then examined apes' responses to the actor's gaze cues during the looking phase. We examined their first
302 looks (i.e. their initial responses) either to the target or distractor object (the number of trials in which
303 they first looked either to the target or distractor object) and their total viewing times (i.e. their overall
304 responses in 5 seconds) to the target and distractor object (the total viewing times for the target and
305 distractor objects). Following previous studies with great apes (Hattori et al. 2010; Kano and Call 2014a),
306 but unlike in the original studies with human infants, we used raw scores of looking to the objects, instead
307 of the number of saccades from the face to the object (because apes typically do not attend to the actor's
308 face as much as human infants do in the gaze-following context); however, preliminary analyses
309 confirmed that the results were the same using either of these measures.

310 It should be noted that the original studies used difference scores for both the first-look and
311 viewing-time measures (the response to the target minus the response to the distractor, divided by the sum
312 of these), while this study used raw scores of first-look and viewing time to the target and distractor. We
313 used raw scores and a repeated-measures ANOVA (instead of difference scores) because this method can
314 analyze participants' overall level of attention to *both* objects (target and distractor). This means that the
315 main effect of AOI (target, distractor) indicates the presence of gaze following (or more looks to the
316 target than distractor), and the interaction effect of AOI and Condition (ostension, control) indicates the
317 difference in gaze-following responses between conditions. The main effect of Condition indicates more

318 looks to *both* target and distractor (which was not measured in the difference-score analysis). Thus, the
319 only difference between our analysis and those based on difference-scores is that ours did not involve
320 dividing the target-distractor differences by the target-distractor sum (because we were also interested in
321 the difference in the target-distractor sum between conditions). Our preliminary analyses, however,
322 confirmed that this method of calculation did not critically change our results (with regard to the presence
323 of gaze following and the difference in gaze following between conditions).

324

325 Results and Discussion for Experiment 1-3

326 Experiment 1 (WKPRC group).

327 To check if the actor's ostensive signals and control attention-getter similarly attracted chimpanzees'
328 attention to the actor's face (Figure 2A), we conducted a repeated-measures ANOVA with Condition
329 (Ostensive, Control) and Phase (Still, Cueing, Looking) on the chimpanzees' viewing times for the actor's
330 face. We found a significant main effect of Phase ($F(2,28) = 33.92, p < 0.001, \eta^2 = 0.71$) but no other
331 significant main or interaction effects. This result indicates that chimpanzees attended more strongly to
332 the actor's face during the cueing phase than the other phases. The results also indicate that chimpanzees
333 attended equally strongly to the actor's face during all conditions (and more strongly during the cueing
334 than the still and looking phases). We then examined the chimpanzees' first-look and viewing-time
335 responses toward the target and distractor objects during the looking phase (Figure 2A). A repeated-
336 measures ANOVA with Condition (Ostensive, Control) and AOI (Target, Distractor) revealed no
337 significant main or interaction effects in the first-look response (Condition: $F(1,14) = 3.06, p = 0.10, \eta^2 =$
338 0.18 ; AOI: $F(1,14) = 2.89, p = 0.11, \eta^2 = 0.17$; Interaction effect; $F(1,14) = 0.02, p = 0.90, \eta^2 = 0.001$) or
339 in the viewing-time response (Condition: $F(1,14) = 3.98, p = 0.07, \eta^2 = 0.22$; AOI: $F(1,14) = 2.15, p =$
340 $0.10, \eta^2 = 0.13$; Interaction effect; $F(1,14) = 0.094, p = 0.76, \eta^2 = 0.007$).

341 Experiment 2 (KS-PRI group).

342 In Experiment 1, chimpanzees responded to the actor's ostensive and gaze cues only weakly. This result
343 was not surprising given that this group of chimpanzees (zoo-reared: WKPRC) didn't follow a human
344 actor's gaze in a similar eye-tracking set-up (Kano and Call 2014a). In Experiment 2, we thus tested
345 another group of chimpanzee participants (institute-reared: KS-PRI) who had richer experiences of
346 interacting with human experimenters/caregivers since youth. In addition, in the results, we noticed that
347 the control attention getter in Experiment 1 elicited apes' attention during the cueing phase slightly more
348 weakly than the actor's ostensive cues. Consequently, we used another type of control cue that has been
349 implemented in the previous studies with infants and dogs (Senju and Csibra 2008; Téglás et al. 2012).

350 As in Experiment 1, we first checked if the actor's ostensive signals and control attention-getter
351 similarly attracted chimpanzees' attention to the actor's face (Figure 2B); a repeated-measures ANOVA
352 with Condition (Ostensive, Control) and Phase (Still, Cueing, Looking) on the chimpanzees' viewing
353 times for the actor's face revealed a significant main effect of Phase ($F(2,22) = 26.43, p < 0.001, \eta^2 =$
354 0.71), but no other significant main or interaction effects; consistent with Experiment 1. We then
355 examined the chimpanzees' first-look and viewing-time responses during the looking phase (Figure 2B).
356 For the first-look responses, a repeated-measures ANOVA with Condition (Ostensive, Control) and AOI
357 (Target, Distractor) revealed a significant main effect of AOI ($F(1,11) = 7.29, p = 0.02, \eta^2 = 0.40$), but no
358 significant effect of Condition ($F(1,11) = 2.17, p = 0.17, \eta^2 = 0.17$) or interaction between AOI and
359 Condition ($F(1,11) = 0.007, p = 0.93, \eta^2 = 0.001$). That is, chimpanzees made more first looks to the
360 target object than the distractor in both conditions. For the viewing-time responses, a repeated-measures
361 ANOVA with Condition (Ostensive, Control) and AOI (Target, Distractor) revealed a significant main
362 effect of Condition ($F(1,11) = 15.18, p = 0.002, \eta^2 = 0.58$), but no significant main effect of AOI ($F(1,11)$
363 $= 2.29, p = 0.16, \eta^2 = 0.17$) or interaction effect ($F(1,11) = 0.33, p = 0.58, \eta^2 = 0.029$). That is,
364 chimpanzees spent more time looking at both the target and distractor objects following the ostensive cue
365 than following the non-ostensive cue. These results suggest that, 1) ostensive cues did not enhance those
366 chimpanzees' gaze-following responses more than control cues, 2) although the actor's gaze cues did
367 guide chimpanzees' looks to the target object (i.e., chimpanzees looked first to the target in both
368 conditions). However, 3) the actor's ostensive cues elicited greater looking to both objects (but not to the
369 actor's face per se) as compared with control cues.

370 [An additional analysis for Experiment 1 and 2.](#)

371 Experiment 1 and 2 consistently did not find significant interaction effects between Condition and AOI
372 (i.e. a differential looking to the target and distractor between conditions), and the effect sizes were also
373 small. In these experiments, however, the statistical results for the main effects of AOI/Condition were
374 mixed; specifically, Experiment 1 found neither significant main effect of AOI nor that of Condition,

375 while Experiment 2 found both (but in different eye-movement measures). Yet the result trends were
376 similar between the experiments (Figure 2). It thus remains unclear if the lack of significant effects
377 simply reflects insufficient power of the statistical tests (e.g. small sample sizes) or specific differences
378 between the experiments, such as the control cues used (the actor's shaking head vs. the animation on the
379 actor's head) or the chimpanzee groups tested (zoo-reared vs. institute-reared).

380 To test this, we conducted a combined analysis on the results from Experiment 1-2 with an
381 addition of the factor Group (WKPRC, KS-PRI). We first checked if the actor's ostensive signals and
382 control attention-getter similarly attracted chimpanzees' attention to the actor's face; a repeated-measures
383 ANOVA with Condition (Ostensive, Control), Phase (Still, Cueing, Looking), and Group (WK PRC, KS-
384 PRI) on the chimpanzee's viewing times for the actor's face revealed a significant main effect of Phase
385 ($F(2,48) = 55.85, p < 0.001, \eta^2 = 0.67$), but no other significant main or interaction effects. We then
386 examined the chimpanzees' first-look and viewing-time responses during the looking phase. For the first-
387 look responses, a repeated-measures ANOVA with Condition (Ostensive, Control), AOI (Target,
388 Distractor), and Group (WK PRC, KS-PRI) revealed a significant main effect of AOI ($F(1,25) = 10.73, p$
389 $= 0.003, \eta^2 = 0.30$) and Condition ($F(1,25) = 5.15, p = 0.032, \eta^2 = 0.17$), but no significant interaction
390 effect ($F(1,11) < 0.001, p = 0.99, \eta^2 < 0.001$). The effect of Group (either main or interaction effect) was
391 not significant. These results indicate that chimpanzees followed the actor's gaze and looked first to the
392 target in both conditions across experiments but also looked more often (in more trials) at both the target
393 and distractor following the ostensive cue. For the viewing-time responses, a repeated-measures ANOVA
394 with Condition (Ostensive, Control), AOI (Target, Distractor), and Group (WK PRC, KS-PRI) revealed
395 significant main effects of AOI ($F(1,25) = 4.56, p = 0.043, \eta^2 = 0.15$) and Condition ($F(1,25) = 20.57, p <$
396 $0.001, \eta^2 = 0.45$), but no significant interaction between AOI and Condition ($F(1,25) = 0.18, p = 0.68, \eta^2$
397 $= 0.007$). These results indicate that chimpanzees spent more time looking at the target across conditions
398 but showed more looking to both the target and distractor following the ostensive cue. The main effect of
399 Group ($F(1,25) = 7.95, p = 0.009, \eta^2 = 0.24$) and the interaction effect of Group and Condition ($F(1,25) =$

400 5.43, $p = 0.028$, $\eta^2 = 0.18$) were also significant (the main effect of Condition was more evident in
401 Experiment 2/KS-PRI group). Conducting the analysis separately for each condition (on the viewing-time
402 responses) revealed that those two groups differed from one another in the ostensive condition ($F(1,25) =$
403 9.31 , $p = 0.005$, $\eta^2 = 0.27$), rather than in the control condition ($F(1,25) = 2.61$, $p = 0.12$, $\eta^2 = 0.095$).
404 These results indicate that the KS-PRI chimpanzees responded to the actor's ostensive cues more strongly,
405 looking more to both objects, than WKPRC chimpanzees (at least in the viewing-time responses). Overall,
406 these combined analyses consolidated the findings from Experiment 1 and 2.

407 Experiment 3

408 Experiment 3 was conducted to further consolidate the findings from Experiment 1 and 2 with a different
409 control attention-getter and minor changes in the stimuli. In Experiment 1 and 2, there was still a concern
410 that those control attention-getters might have been slightly weaker than the actor's ostensive signals to
411 attract apes' attention during the cueing phase. In Experiment 3, we used the actor's eating action as a
412 control attention-getter because a previous eye-tracking study confirmed that eating action strongly
413 catches apes' attention (Kano et al. 2018). Moreover, in Experiment 1-2, there might be a concern that the
414 actor holding both objects in his hands might confound the effect of gaze following (to the target objects)
415 and that of attention to manual actions (to both target and distractor objects). In Experiment 3, we thus
416 made minor changes in the configuration of the scenes in the movies so that the objects were mounted on
417 tripods to each side of the actor at eye level instead of supported by the actor's hand while they rested on
418 a table. Finally, in Experiment 3, we tested the chimpanzees from both groups (WKPRC, KS-PRI) and
419 thus included the factor Group into the analyses.

420 As in Experiment 1 and 2, we first checked if the actor's ostensive signals and control attention-
421 getter similarly attracted chimpanzees' attention to the actor's face (Figure 2C); a repeated-measures
422 ANOVA with Condition (Ostensive, Control), Phase (Still, Cueing, Looking), and Group (WKPRC, KS-
423 PRI) on the chimpanzees' viewing times for the actor's face revealed a significant main effect of Phase
424 ($F(2,34) = 29.5$, $p < 0.001$, $\eta^2 = 0.64$), but no other significant main or interaction effects; consistent with

425 Experiment 1 and 2. We then examined chimpanzees' first-look and viewing-time responses to the target
426 and distractor objects during the looking phase (Figure 2C). For the first-look responses, a repeated-
427 measures ANOVA with Condition (Ostensive, Control), AOI (Target, Distractor), and Group (WKPRC,
428 KS-PRI) revealed a significant main effect of Condition ($F(1,17) = 6.34, p = 0.022, \eta^2 = 0.27$), but no
429 significant main effect of AOI ($F(1,17) = 1.43, p = 0.25, \eta^2 = 0.08$) or interaction effect ($F(1,17) = 0.054,$
430 $p = 0.82, \eta^2 = 0.003$). The effect of Group (either main or interaction effect) was not significant. The main
431 effect of Condition indicates that chimpanzees looked more often (in more trials) at both the target and
432 distractor in the ostension than control condition. For the viewing-time responses, a repeated-measures
433 ANOVA with Condition (Ostensive, Control), AOI (Target, Distractor), and Group (WKPRC, KS-PRI)
434 revealed a significant main effect of Condition ($F(1,17) = 24.89, p < 0.001, \eta^2 = 0.59$), but no significant
435 main effect of AOI ($F(1,17) = 0.069, p = 0.80, \eta^2 = 0.004$) or interaction effect ($F(1,17) = 0.071, p = 0.79,$
436 $\eta^2 = 0.004$). These results indicate that chimpanzees spent more time looking at both objects in the
437 ostensive condition than in the non-ostensive condition. The main effect of Group ($F(1,17) = 7.45, p =$
438 $0.014, \eta^2 = 0.31$) and the interaction effect of Group and Condition ($F(1,17) = 22.94, p < 0.001, \eta^2 = 0.57$)
439 were also significant; the KS-PRI chimpanzees looked more to both objects than WKPRC chimpanzees.
440 Overall, the results from Experiment 3 were consistent with Experiment 1 and 2 (although we did not find
441 the main effect of AOI, but this is likely due to weaker power in this analysis compared with the
442 combined analysis, rather than due to the changes in the stimulus; note that bonobos and orangutans
443 showed the main effect of AOI; see Supplemental Material).

444 [Summary of the results for the other species \(bonobos and orangutans\)](#)

445 We also tested bonobos in Experiment 1-3 and orangutans in Experiments 1 and 3 using the same
446 stimuli and procedure. See Figure S2 and S3 for the summary of the results from bonobos and orangutans,
447 respectively. Like chimpanzees, neither species followed the actor's gaze more sensitively in the
448 ostension than control condition (i.e. no significant interaction between Condition and AOI). Interestingly,
449 while orangutans were somewhat similar to chimpanzees in that they viewed both target and distractor

450 objects longer in the ostensive than control condition (i.e. a significant main effect of Condition at least in
451 Experiment 1), bonobos were not; in all three experiments, they spent similar time looking at the objects
452 in both conditions. Moreover, bonobos viewed the actor's face longer in the ostension than control
453 condition during the cueing phase. Presumably, such behaviors were driven by bonobos' reflexivity in
454 following others' gaze (Kano and Call 2014a) and their general sensitivity to eye contact with other
455 individuals (Kano et al. 2015).

456

457

458 General Discussion

459 This study tested whether humans' ostensive signals enhance gaze following in great apes, particularly in
460 chimpanzees. We found that, although chimpanzees did follow the actor's gaze (i.e., looked first to the
461 target object following gaze cueing), unlike infants and domestic dogs, human ostensive signals did not
462 enhance gaze following more strongly than control attention-getters for chimpanzees (nor for bonobos or
463 orangutans; see supplementary materials). However, chimpanzees did distinguish between the ostensive
464 signals and the control attention-getters (as did orangutans to some extent, although not bonobos). In the
465 ostension condition, they spent more time attending to both the target object (the actor's intended
466 referent) and the distractor than in the control condition. Importantly, they did so even though they paid
467 an equal level of attention to the actor's face across conditions during both the cueing and looking phases.
468 Thus, these results showed that the ostensive signals increased apes' attention specifically to the objects
469 but not to the actor following the actor's ostensive cues. Overall, therefore, chimpanzees seemed to expect
470 that the actor's ostensive signals would precede information specifically about the objects (rather than
471 about the actor). Nonetheless, this expectation seems more functionally limited than in human infants and
472 domestic dogs, because chimpanzees did not subsequently focus their attention on the intended referent of
473 the actor's communicative act.

474 The finding that humans' ostensive signals do not enhance gaze following in great apes is consistent
475 with the idea that human infants and domestic dogs are better at understanding humans' referential signals
476 than great apes and wolves (Hare et al. 2002; Topál et al. 2009). It thus suggests that non-domesticated
477 species such as great apes lack one of the skills (or perceptual/cognitive biases) that would help them to
478 understand or respond appropriately to human referential communication, while human infants and
479 domestic dogs have acquired such skills through ontogeny and evolution (Csibra and Gergely 2009; Senju
480 and Csibra 2008; Téglás et al. 2012). However, our results do suggest that apes understand, at least partly,
481 that humans' ostensive signals precede referential information, as they searched the environment longer
482 after witnessing the ostensive signals compared to equally attention-grabbing non-ostensive signals. What

483 they clearly didn't do is attempt to specify the intended referent further, on the basis of the actor's gaze
484 behavior after seeing the ostensive signals. This might suggest that, although they understood a basic role
485 of humans' ostensive signals (i.e. the agent is trying to communicate something), they have failed to
486 understand the function of ostensive and gaze signals combined (i.e. the agent is trying to communicate
487 something about the cued object). Such reliance on environmental cues in apes may be observed not only
488 in communicative contexts but also in the context of apes' social referencing behavior in general – and
489 particularly in the context of research on chimpanzee imitation and emulation. Previous studies
490 consistently suggest that when apes watch others using tools, they preferentially attend to the features of
491 an environment that permit certain sorts of causal affordances, while being relatively inattentive to the
492 particular techniques used by those whom they observe (Tennie et al. 2009; Tomasello 1999). By virtue
493 of attending to environmental affordances, apes can learn to use tools by watching others. However,
494 because they are inattentive to the particular techniques produced by agents, they are unable to reproduce
495 any arbitrary features of actions (this is sometimes described as a preference for emulation over imitation).
496 In contrast, human infants tend to show an opposite preference; they sometimes even over-imitate others
497 (Whiten et al. 2009). It may be that, when apes are seeking to gain information about their environment,
498 they have a stronger tendency than human infants to attend to the environment rather than to social cues –
499 even when attending to social cues might prove particularly helpful.

500 The differences between our results with chimpanzees and previous results with human infants and
501 dogs are not due to differences in the particular control stimuli used in this and previous studies.
502 Experiment 2 presented a similar control attention-getter (a salient pattern on the actor's head) to that
503 used in previous studies (Senju and Csibra 2008; Téglás et al. 2012). Although the artificial nature of this
504 control attention-getter was considered a potential problem in other studies (Gredebäck et al. 2018;
505 Szufnarowska et al. 2014), in our study, this control condition produced similar results to the more natural
506 actions performed by the human actor in Experiments 1 and 3. Critically, while human infants and dogs in
507 previous studies (Senju and Csibra 2008; Téglás et al. 2012) followed the actor's gaze –looked more

508 frequently and longer at the cued object respectively in the first-look and viewing-time measures—only in
509 the ostension conditions, apes in this study did not show such a pattern in either measure. Instead, we
510 found that chimpanzees followed the actor’s gaze across conditions (i.e., looked first to the cued target
511 object) but continued to search longer (i.e., looked longer at both cued and non-cued objects) in the
512 ostension conditions than the controls.

513 Interestingly, the differences between our results and previous ones with chimpanzees highlight the
514 potential importance of attention-getters in eliciting reliable gaze following. One notable difference
515 between this and previous ape eye-tracking (gaze-following) studies is that chimpanzees followed the
516 human actor’s gaze in this study; in previous studies that lacked ostensive or attention-getting signals,
517 chimpanzees did not follow human gaze (Hattori et al. 2010; Kano and Call 2014a). Thus, this study
518 builds on previous work by providing evidence that the general presence of attention-getters, including
519 both ostensive signals and non-ostensive attention-getters, may help chimpanzees to follow human gaze
520 in this setting. This may explain why chimpanzees reliably follow the gaze of a human experimenter in a
521 live setup where the human experimenter typically ensures the chimpanzees’ attention to the face (or at
522 least to the body) before providing the gaze cue (e.g. Call et al. 2000; Tomasello et al. 2007).
523 Interestingly, this same argument may apply, to some extent, to human infants. That is, while human
524 infants (and domestic dogs) showed limited gaze-following responses after seeing non-ostensive control
525 attention-getters in the two earliest studies (Senju and Csibra 2008; Téglás et al. 2012), in two more
526 recent studies (Gredebäck et al. 2018; Szufnarowska et al. 2014), they followed an actor’s gaze after both
527 ostensive and non-ostensive actions (e.g. head shivering; but see Senju and Csibra’s (2014) commentary
528 on the potential problems in Szefnarowska’s article). It should be also noted that, unlike chimpanzees,
529 human infants and dogs, human adults should follow the actor’s gaze in any condition because the task
530 demands in this and previous studies seem too easy for them. Our preliminary tests with human adults
531 indeed showed that human adults strongly followed the actor’s gaze in both ostention and control
532 conditions (Figure S5). Thus, although the results of our three studies consistently show that apes do not

533 gaze follow more robustly in response to ostensive cues than various non-ostensive cues, if humans also
534 do not distinguish between those stimuli as efficiently as previously assumed, our conclusion about a
535 species difference between humans and the other apes must be tentative; further studies are necessary to
536 examine to what extent and in what circumstances humans distinguish between an actor's ostensive
537 signals and control actions in potentially communicative contexts.

538 It is noteworthy that chimpanzees who had richer early experiences interacting with human caregivers
539 paid more attention to the objects (but not to the actor's face) following the actor's ostensive cues than did
540 other chimpanzees. This suggests that chimpanzees with more experience of human interaction are more
541 sensitive to humans' ostensive signals than are the less experienced chimpanzees. This is consistent with
542 previous findings that enculturated chimpanzees show improved performances in tests in which they need
543 to locate hidden foods based on an experimenter's referential cues (Call and Tomasello 1994; Call and
544 Tomasello 1996; Lyn et al. 2010). Nonetheless, in our study, the ostensive signals did not enhance gaze
545 following in either group of chimpanzees. These results suggest that, although chimpanzees with richer
546 experiences with humans are generally more sensitive to human signals, even those experienced
547 chimpanzees do not interpret humans' ostensive signals in the same way as human infants and domestic
548 dogs do.

549 An interesting question that can be addressed in future studies is whether the current findings can be
550 extended to situations in which a conspecific ape actor, instead of a human actor, addresses participants
551 ostensively. In general, chimpanzees (but not bonobos and orangutans) are more likely to follow
552 conspecific than human gaze (Hattori et al. 2010; Kano and Call 2014a). While such differences may be
553 explained by the fact that chimpanzees are more attentive to conspecific faces than human faces in the
554 context of gaze following (Kano and Call 2014a), it remains untested whether chimpanzees would be
555 more likely to follow the gaze of a conspecific after seeing a combination of species-typical attention-
556 getters (hand clapping, spattering) and indexical cues (e.g. gazing and extending arms).

557 In conclusion, we confirmed that, unlike human infants and domestic dogs, humans' ostensive signals
558 do not enhance gaze following in great apes. However, we also found that such signals do enhance
559 subsequent object-related attention or search behaviors in apes, at least in chimpanzee participants who
560 had richer early experiences of interacting with humans. Thus, they may, at least in part, expect ostensive
561 signals to precede referential information. However, instead of fully relying on referential cues, apes may
562 search for additional environmental cues to interpret communication. This may be a limitation (or a lack
563 of human-like perceptual/cognitive bias) of non-domesticated species for interpreting humans' ostensive
564 signals in inter-species communication.

565 Data availability

566 All data generated or analyzed during this study are included in this published article (and its
567 supplementary information files).

568 Acknowledgements

569 We thank the staff at Kumamoto Sanctuary, Primate Research Institute, and Wolfgang Kohler Primate
570 Research Center for their assistance in conducting the series of experiments. We also thank two
571 anonymous reviewers for their helpful comments. Financial support came from Ministry of Education,
572 Culture, Sports, Science and Technology [K-CONNEX to FK], Japan Society for Promotion of Science
573 [KAKENHI 26885040, 16K21108, and 18H05072 to FK, 26245069, 16H06301, and 16H06283, SH,
574 15H05709, 16H06238, and JSPS-CCSN to MT, and JSPS-LGP-U04 and Great Ape Information Network
575 to SH and MT, and the European Research Council [SOMICS 609819 to JC].

576

577 References

578 Behne T, Carpenter M, Tomasello M (2005) One-year-olds comprehend the communicative intentions
579 behind gestures in a hiding game Dev Sci 8:492-499 doi:10.1111/j.1467-7687.2005.00440.x

580 Behne T, Liszkowski U, Carpenter M, Tomasello M (2012) Twelve-month-olds' comprehension and
581 production of pointing *British Journal of Developmental Psychology* 30:359-375
582 doi:10.1111/j.2044-835X.2011.02043.x

583 Bräuer J, Call J, Tomasello M (2005) All great ape species follow gaze to distant locations and around
584 barriers *J Comp Psychol* 119:145-154 doi:10.1037/0735-7036.119.2.145

585 Bugnyar T, Stöwe M, Heinrich B (2004) Ravens, *Corvus corax*, follow gaze direction of humans around
586 obstacles *Proceedings of the Royal Society of London Series B: Biological Sciences* 271:1331-
587 1336 doi:10.1098/rspb.2004.2738

588 Call J, Agnetta B, Tomasello M (2000) Cues that chimpanzees do and do not use to find hidden objects
589 *Anim Cogn* 3:23-34 doi:10.1007/s100710050047

590 Call J, Tomasello M (1994) Production and comprehension of referential pointing by orangutans (*Pongo*
591 *pygmaeus*) *J Comp Psychol* 108:307 doi:10.1037//0735-7036.108.4.307

592 Call J, Tomasello M (1996) The effect of humans on the cognitive development of apes. In: Russon AE,
593 Bard KA, Parker ST (eds) *Reaching into thought: The minds of the great apes*. Cambridge
594 University Press, New York, pp 371-403

595 Call J, Tomasello M (2008) Does the chimpanzee have a theory of mind? 30 years later *Trends Cogn Sci*
596 12:187-192 doi:10.1016/j.tics.2008.02.010

597 Cooper RP, Aslin RN (1990) Preference for infant directed speech in the first month after birth *Child Dev*
598 61:1584-1595 doi:10.1111/j.1467-8624.1990.tb02885.x

599 Csibra G (2010) Recognizing communicative intentions in infancy *Mind & Language* 25:141-168
600 doi:10.1111/j.1468-0017.2009.01384.x

601 Csibra G, Gergely G (2009) Natural pedagogy *Trends Cogn Sci* 13:148-153
602 doi:10.1016/j.tics.2009.01.005

603 De Waal FB (1988) The communicative repertoire of captive bonobos (*Pan paniscus*), compared to that
604 of chimpanzees *Behaviour* 106:183-251 doi:10.1163/156853988X00269

605 De Waal FBM (1990) *Peacemaking among primates*. Harvard University Press, Cambridge

606 Emery NJ (2000) The eyes have it: The neuroethology, function and evolution of social gaze *Neurosci*
607 *Biobehav Rev* 24:581-604 doi:10.1016/S0149-7634(00)00025-7

608 Farroni T, Csibra G, Simion F, Johnson MH (2002) Eye contact detection in humans from birth *Proc Nat*
609 *Acad Sci* 99:9602 doi:10.1073/pnas.152159999

610 Gergely G, Csibra G (2006) Sylvia's recipe: The role of imitation and pedagogy in the transmission of
611 cultural knowledge. In: Enfield N, Levenson SC (eds) *Roots of human sociality: Culture,*
612 *cognition, and human interaction.* Berg Publishers, Oxford, pp 229-255

613 Gergely G, Egyed K, Király I (2007) On pedagogy *Dev Sci* 10:139-146 doi:10.1111/j.1467-
614 7687.2007.00576.x

615 Gomez JC (1996) Ostensive behavior in great apes: The role of eye contact. In: Russon AE, Bard KA,
616 Parker ST (eds) *Reaching into thought: The minds of the great apes.* Cambridge University Press,
617 New York, pp 131-151

618 Gredebäck G, Astor K, Fawcett C (2018) Gaze Following Is Not Dependent on Ostensive Cues: A
619 Critical Test of Natural Pedagogy *Child Dev*:in press doi:doi:10.1111/cdev.13026

620 Hare B, Brown M, Williamson C, Tomasello M (2002) The domestication of social cognition in dogs
621 *Science* 298:1634-1636 doi:10.1126/science.1072702

622 Hare B, Tomasello M (1999) Domestic dogs (*Canis familiaris*) use human and conspecific social cues to
623 locate hidden food *J Comp Psychol* 113:173 doi:10.1037//0735-7036.113.2.173

624 Hare B, Tomasello M (2004) Chimpanzees are more skilful in competitive than in cooperative cognitive
625 tasks *Anim Behav* 68:571-581 doi:10.1016/j.anbehav.2003.11.011

626 Hattori Y, Kano F, Tomonaga M (2010) Differential sensitivity to conspecific and allospecific cues in
627 chimpanzees and humans: A comparative eye-tracking study *Biology Lett* 6:610-613
628 doi:10.1098/rsbl.2010.0120

629 Herrmann E, Tomasello M (2006) Apes' and children's understanding of cooperative and competitive
630 motives in a communicative situation *Dev Sci* 9:518-529 doi:10.1111/j.1467-7687.2006.00519.x

631 Heyes C (2016) Born pupils? Natural pedagogy and cultural pedagogy *Perspectives on Psychological*
632 *Science* 11:280-295 doi:10.1177/1745691615621276

633 Hirata S, Fuwa K, Sugama K, Kusunoki K, Fujita S (2010) Facial perception of conspecifics:
634 Chimpanzees (*Pan troglodytes*) preferentially attend to proper orientation and open eyes *Anim*
635 *Cogn* 13:679-688 doi:10.1007/s10071-010-0316-y

636 Hoicka E (2016) Parents and toddlers distinguish joke, pretend and literal intentional contexts through
637 communicative and referential cues *Journal of Pragmatics* 95:137-155
638 doi:10.1016/j.pragma.2015.10.010

639 Kaminski J, Schulz L, Tomasello M (2012) How dogs know when communication is intended for them
640 *Dev Sci* 15:222-232 doi:10.1111/j.1467-7687.2011.01120.x

641 Kano F, Call J (2014a) Cross-species variation of gaze following and conspecific preference among great
642 apes, human infants and adults *Anim Behav* 91:137-150 doi:10.1016/j.anbehav.2014.03.011

643 Kano F, Call J (2014b) Great apes generate goal-based action predictions: An eye-tracking study *Psychol*
644 *Sci* 25:1691-1698 doi:10.1177/0956797614536402

645 Kano F, Call J, Tomonaga M (2012) Face and eye scanning in gorillas, orangutans, and humans: Unique
646 eye-viewing patterns in humans among hominids *J Comp Psychol* 126:388-398
647 doi:10.1037/a0029615.

648 Kano F, Hirata S, Call J (2015) Social attention in *Pan*: bonobos exhibit more eye contacts than
649 chimpanzees *PLoS One* 10:e0129684 doi:10.1371/journal.pone.0129684

650 Kano F, Hirata S, Call J, Tomonaga M (2011) The visual strategy specific to humans among hominids: A
651 study using the gap-overlap paradigm *Vision Res* 51:2348-2355 doi:10.1016/j.visres.2011.09.006

652 Kano F, Krupenye C, Hirata S, Call J (2017) Eye tracking uncovered great apes' ability to anticipate that
653 other individuals will act according to false beliefs *Communicative & Integrative Biology*
654 10:e1299836 doi:10.1080/19420889.2017.1299836

655 Kano F, Shepherd SV, Hirata S, Call J (2018) Primate social attention: Species differences and effects of
656 individual experience in humans, great apes, and macaques PLOS ONE 13:e0193283
657 doi:10.1371/journal.pone.0193283

658 Kano F, Tomonaga M (2009) How chimpanzees look at pictures: a comparative eye-tracking study Proc
659 Roy Soc B 276:1949-1955 doi:10.1098/rspb.2008.1811

660 Krupenye C, Kano F, Hirata S, Call J, Tomasello M (2016) Great apes anticipate that other individuals
661 will act according to false beliefs Science 354:110-114 doi:10.1126/science.aaf8110

662 Liebal K, Call J (2012) The origins of non-human primates' manual gestures Philosophical Transactions
663 of the Royal Society of London B: Biological Sciences 367:118-128 doi:10.1098/rstb.2011.0044

664 Lyn H, Russell JL, Hopkins WD (2010) The impact of environment on the comprehension of declarative
665 communication in apes Psychol Sci 21:360-365 doi:10.1177/0956797610362218

666 Miklósi Á, Polgárdi R, Topál J, Csányi V (1998) Use of experimenter-given cues in dogs Anim Cogn
667 1:113-121 doi:10.1007/s100710050016

668 Moore R (2016) Meaning and ostension in great ape gestural communication Anim Cogn 19:223-231
669 doi:10.1007/s10071-015-0905-x

670 Moore R (2017) Gricean communication and cognitive development The Philosophical Quarterly 67:303-
671 326 doi:10.1093/pq/pqw049

672 Moore R, Call J, Tomasello M (2015a) Production and comprehension of gestures between orang-utans
673 (*Pongo pygmaeus*) in a referential communication game PloS One 10:e0129726
674 doi:10.1371/journal.pone.0129726

675 Moore R, Liebal K, Tomasello M (2013) Three-year-olds understand communicative intentions without
676 language, gestures, or gaze Interaction Studies 14:62-80 doi:10.1075/is.14.1.05moo

677 Moore R, Mueller B, Kaminski J, Tomasello M (2015b) Two-year-old children but not domestic dogs
678 understand communicative intentions without language, gestures, or gaze Dev Sci 18:232-242
679 doi:10.1111/desc.12206

680 Myowa-Yamakoshi M, Tomonaga M, Tanaka M, Matsuzawa T (2003) Preference for human direct gaze
681 in infant chimpanzees (*Pan troglodytes*) *Cognition* 89:113-124 doi:10.1016/S0010-
682 0277(03)00071-4

683 Okamoto-Barth S, Call J, Tomasello M (2007) Great apes' understanding of other individuals' line of
684 sight *Psychol Sci* 18:462-468 doi:10.1111/j.1467-9280.2007.01922.x

685 Parr LA, Dove T, Hopkins WD (1998) Why faces may be special: evidence of the inversion effect in
686 chimpanzees *J Cogn Neurosci* 10:615-622 doi:10.1162/089892998563013

687 Riedel J, Schumann K, Kaminski J, Call J, Tomasello M (2008) The early ontogeny of human–dog
688 communication *Anim Behav* 75:1003-1014 doi:10.1016/j.anbehav.2007.08.010

689 Scott-Phillips TC (2015a) Meaning in animal and human communication *Anim Cogn* 18:801-805
690 doi:10.1007/s10071-015-0845-5

691 Scott-Phillips TC (2015b) Nonhuman primate communication, pragmatics, and the origins of language
692 *Current Anthropology* 56:56-80 doi:10.1086/679674

693 Senju A, Csibra G (2008) Gaze following in human infants depends on communicative signals *Curr Biol*
694 18:668-671 doi:10.1016/j.cub.2008.03.059

695 Senju A, Csibra G (2014) A commentary by Csibra, Senju, et al. on gaze following. Retrieved from
696 [http://cognitionandculture.net/blog/icci-blog/for-the-record-a-commentary-by-csibra-senju-et-al-](http://cognitionandculture.net/blog/icci-blog/for-the-record-a-commentary-by-csibra-senju-et-al-on-gaze-following)
697 [on-gaze-following](http://cognitionandculture.net/blog/icci-blog/for-the-record-a-commentary-by-csibra-senju-et-al-on-gaze-following).

698 Sperber D, Wilson D (1995) *Relevance: Communication and cognition* (2nd edition). Blackwell, Oxford

699 Szufnarowska J, Rohlfing KJ, Fawcett C, Gredebäck G (2014) Is ostension any more than attention?
700 *Scientific Reports* 4:5304 doi:10.1038/srep05304

701 Téglás E, Gergely A, Kupán K, Miklósi Á, Topál J (2012) Dogs' gaze following is tuned to human
702 communicative signals *Curr Biol* 22:209-212 doi:10.1016/j.cub.2011.12.018

703 Tennie C, Call J, Tomasello M (2009) Ratcheting up the ratchet: on the evolution of cumulative culture
704 *Philosophical Transactions of the Royal Society B: Biological Sciences* 364:2405-2415
705 doi:10.1098/rstb.2009.0052

706 Tomasello M (1999) The cultural origins of human cognition. Harvard University Press, Cambridge, MA

707 Tomasello M (2006) Why don't apes point? In: Enfield N, Levinson SC (eds) Roots of human sociality:
708 Culture, cognition, and interaction, vol 197. Berg, London, pp 506-524

709 Tomasello M, Hare B, Lehmann H, Call J (2007) Reliance on head versus eyes in the gaze following of
710 great apes and human infants: The cooperative eye hypothesis *J Hum Evol* 52:314-320
711 doi:10.1016/j.jhevol.2006.10.001

712 Tomonaga M (1999) Inversion effect in perception of human faces in a chimpanzee (*Pan troglodytes*)
713 *Primates* 40:417-438 doi:10.1007/BF02557579

714 Topál J, Gergely G, Erdőhegyi Á, Csibra G, Miklósi Á (2009) Differential sensitivity to human
715 communication in dogs, wolves, and human infants *Science* 325:1269-1272
716 doi:10.1126/science.1176960

717 Topál J, Kis A, Oláh K (2014) Dogs' sensitivity to human ostensive cues: a unique adaptation. In:
718 Kaminski J, Marshall-Pescini S (eds) *The social dog: Behavior and cognition*. Elsevier, pp 319-
719 346

720 Ueno A et al. (2010) Brain activity in an awake chimpanzee in response to the sound of her own name
721 *Biology Lett* 6:311-313 doi:10.1098/rsbl.2009.0864

722 Virányi Z, Gácsi M, Kubinyi E, Topál J, Belényi B, Ujfalussy D, Miklósi Á (2008) Comprehension of
723 human pointing gestures in young human-reared wolves (*Canis lupus*) and dogs (*Canis familiaris*)
724 *Anim Cogn* 11:373 doi:10.1007/s10071-007-0127-y

725 Vorms M (2012) A-not-B errors: testing the limits of Natural Pedagogy theory *Review of Philosophy and*
726 *Psychology* 3:525-545 doi:10.1007/s13164-012-0113-4

727 Whiten A, McGuigan N, Marshall-Pescini S, Hopper LM (2009) Emulation, imitation, over-imitation and
728 the scope of culture for child and chimpanzee *Philosophical Transactions of the Royal Society of*
729 *London B: Biological Sciences* 364:2417-2428 doi:10.1016/j.anbehav.2007.08.010

730 Yamagiwa J (1992) Functional analysis of social staring behavior in an all-male group of mountain
731 gorillas *Primates* 33:523-544 doi:10.1007/BF02381153

732 Yamamoto S, Humle T, Tanaka M (2012) Chimpanzees' flexible targeted helping based on an
733 understanding of conspecifics' goals Proc Nat Acad Sci 109:3588-3592
734 doi:10.1073/pnas.1108517109

735

736 **Figure caption**

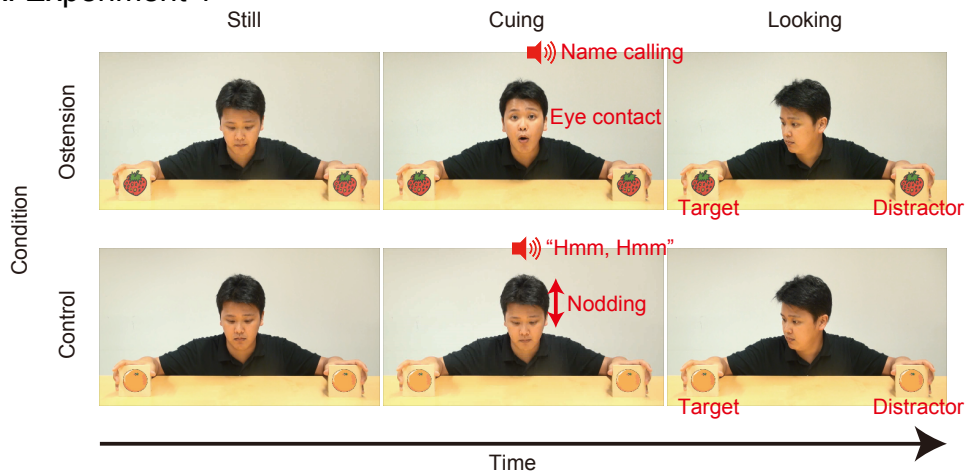
737

738 Figure 1. Movie stimuli used in Experiment 1-3 (A-C)

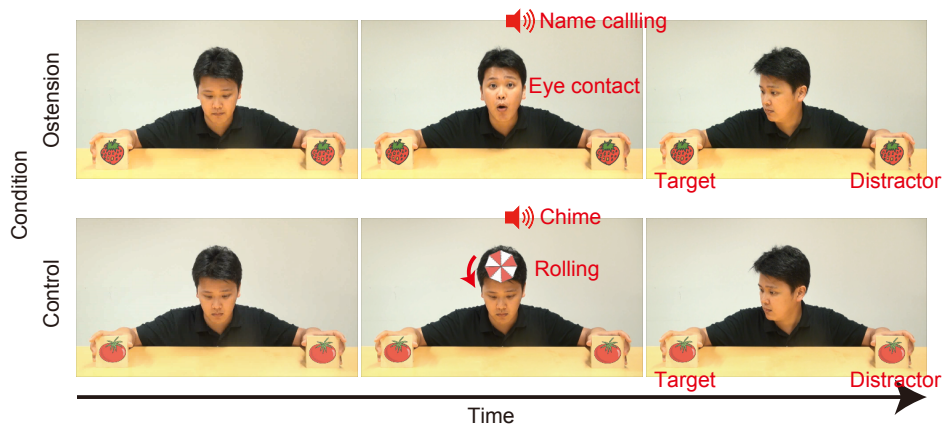
739

740 Figure 2. Results from Experiment 1-3 (A-C). Mean viewing times for the actor's face in each phase of
741 each condition (in 2.5 seconds), and number of first looks (in 6 trials) to the target or distractor object, and
742 mean viewing times for the target or distractor object (in 5 seconds). Asterisks indicate the significance of
743 the main effects (AOI-target/distractor, Condition-ostension/control). * $p < 0.05$, ** $p < 0.01$, *** $p <$
744 0.001

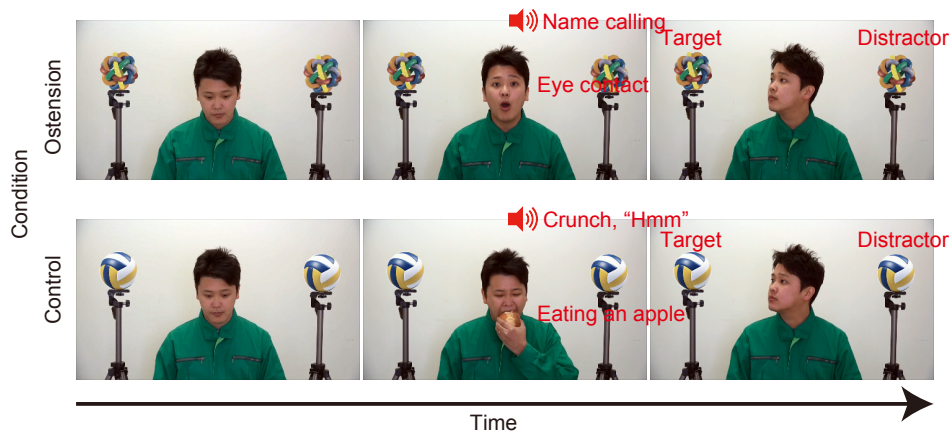
A. Experiment 1



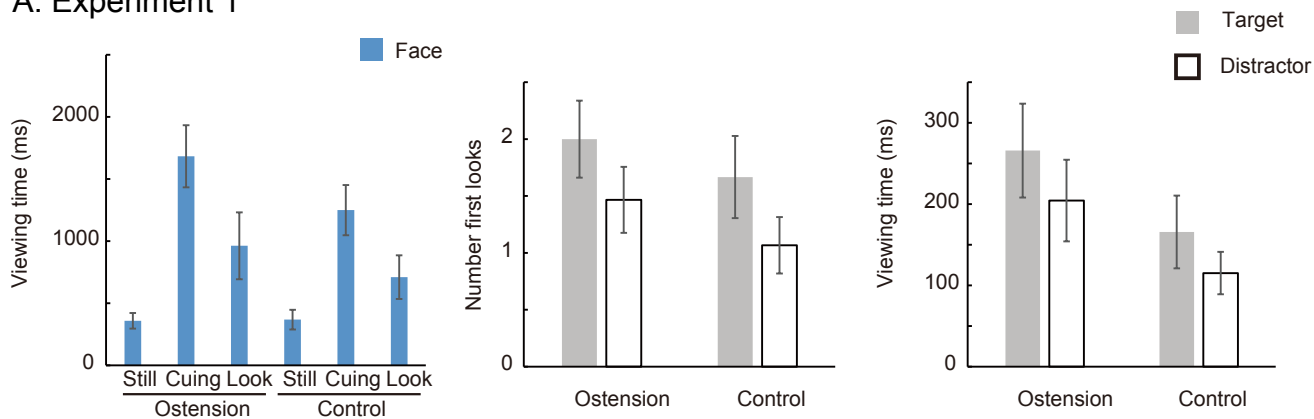
B. Experiment 2



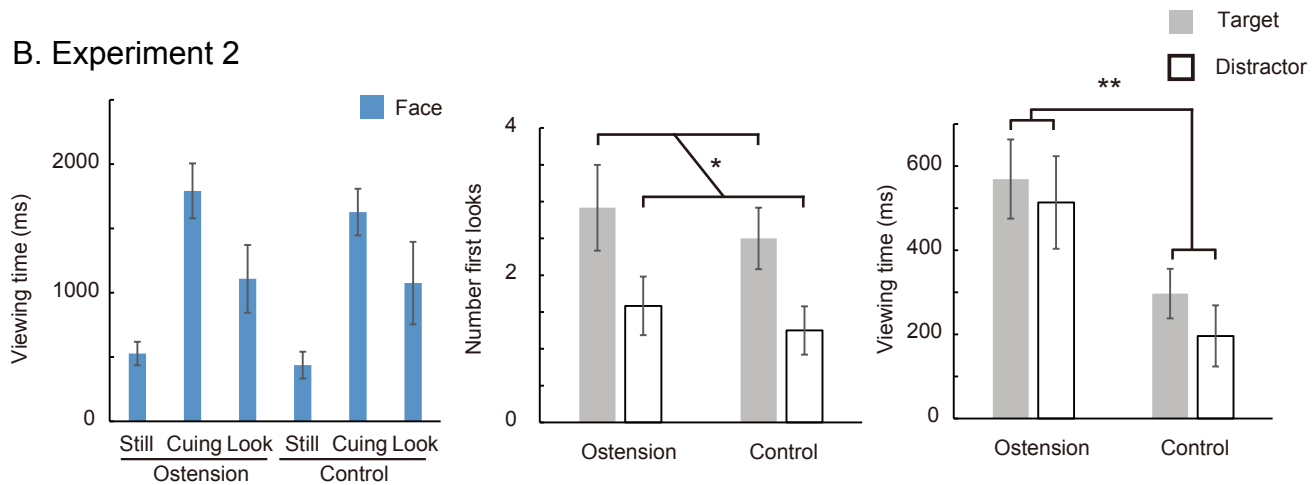
C. Experiment 3



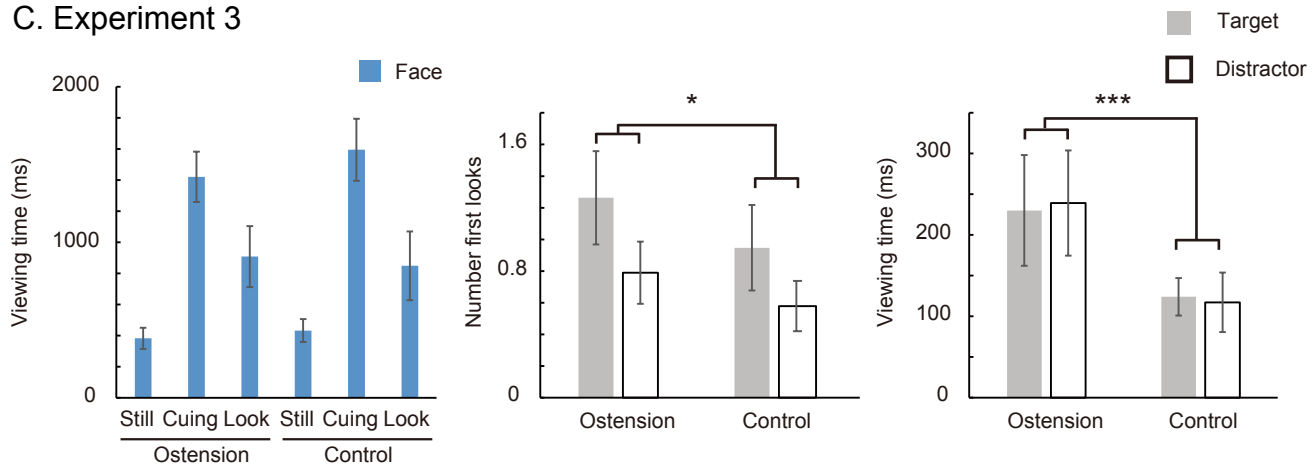
A. Experiment 1



B. Experiment 2



C. Experiment 3



Supplemental Texts

“Within-subject” vs. “between-subject”

Due to sample size limitations, we used a within-subject design in this study, while previous studies with human infants and dogs used a between-subject design. One concern is that there might be a carry-over effect across conditions in this within-subject design, thereby enhancing chimpanzees' gaze-following responses particularly in the control condition. We therefore analyzed the data using the first 6 trials and removing the remaining 6 trials, mimicking a between-subject design (note that each participant first completed one condition in the first 6 trials, before completing the other condition in the remaining 6 trials). We found very similar results in this additional analysis compared with the original analysis (although the results were statistically weaker due to the reduced sample sizes; most importantly here, we did not find the interaction effect between Condition and AOI in either first-look or viewing-time responses).

Specifically, in all experiments, a repeated-measures ANOVA with Condition and Phase on the viewing times for the actor's face confirmed a significant main effect of Phase (Experiment 1: $F(2,26) = 37.02$, $p < 0.001$, $\eta^2 = 0.74$; Experiment 2: $F(2,20) = 19.76$, $p < 0.001$, $\eta^2 = 0.66$; Experiment 3: $F(2,30) = 18.89$, $p < 0.001$, $\eta^2 = 0.56$), but no other significant main/interaction effects, confirming that chimpanzees viewed the actor's face during each phase similarly between conditions. We then conducted a repeated-measures ANOVA with Condition and AOI (and Group in some of the analyses) on the first-look and viewing-time responses to the target/distractor objects. In Experiment 1, we found no significant effect in either first-look or viewing-time responses to the target/distractor. In Experiment 2, we found no significant effects in the first-look response. In the viewing-time response, we found a significant main effect of Condition ($F(1,10) = 6.58$, $p = 0.028$, $\eta^2 = 0.40$), but not a significant main effect of AOI or interaction effect. In the combined analysis for Experiment 1-2, we found a significant main effect of Condition ($F(1,23) = 5.78$, $p = 0.025$, $\eta^2 = 0.20$), but no significant main effect of AOI or Group, or interaction effect, in the first-look response. In the viewing time responses, we found a significant main effect of Condition ($F(1,23) = 8.84$, $p = 0.007$, $\eta^2 = 0.28$) and Group ($F(1,23) = 4.31$, $p = 0.049$, $\eta^2 = 0.16$), but no significant main effect of AOI or interaction effect. In Experiment 3, we found no significant effect in the first-look response. In the viewing time responses, we found a significant main effect of Condition ($F(1,15) = 13.59$, $p = 0.002$, $\eta^2 = 0.48$) and Group ($F(1,15) = 7.65$, $p = 0.014$, $\eta^2 = 0.34$) and the interaction effect between Condition and Group ($F(1,15) = 7.35$, $p = 0.016$, $\eta^2 = 0.33$; the effect was stronger in the KS-PRI group than in the WKPRC group), but no significant main effect of AOI or interaction effect.

Results for bonobos and orangutans

Figure S2 and S3 visualizes the results for bonobos and orangutans, respectively. See the main text for a summary of these results.

Bonobo

In Experiment 1, we tested 7 bonobos from WKPRC with the same stimuli and procedure as described in the main text (Figure 1A). A repeated-measures ANOVA with Condition and Phase on the viewing times for the actor's face confirmed a significant interaction effect between Condition and Phase ($F(2,12) = 11.53$, $p = 0.002$, $\eta^2 = 0.66$) as well as a significant main effect of Condition, $F(1,6) = 6.23$, $p = 0.048$, $\eta^2 = 0.51$, and Phase, $F(2,12) = 41.29$, $p < 0.001$, $\eta^2 = 0.66$). Bonobos viewed the actor's face longer in the ostensive than control condition during the cueing phase (paired t-test, $t(6) = 3.33$, $p = 0.016$). In the looking phase, a repeated-measures ANOVA with Condition and AOI on the first-look responses to the target/distractor objects revealed no significant effects. A repeated-measures ANOVA with Condition and AOI on the viewing-time responses to the target/distractor objects revealed a significant main effect of AOI ($F(1,6) = 6.81$, $p = 0.040$, $\eta^2 = 0.53$), but no significant main effect of Condition or interaction effect. Bonobos spent more time looking at the target than the distractor object across conditions.

In Experiment 2, we tested 6 bonobos from KS with the same stimuli and procedure as described in the main text (Figure 1B). A repeated-measures ANOVA with Condition and Phase on the viewing times for the actor's face confirmed a significant interaction effect between Condition and Phase ($F(2,10) = 5.94$, $p = 0.020$, $\eta^2 = 0.54$); also a significant main effect of Phase, $F(2,10) = 14.93$, $p = 0.001$, $\eta^2 = 0.75$). Bonobos tended to view the actor's face longer in the ostensive than control condition during the cueing phase (paired t-test, $t(5) = 2.51$, $p = 0.054$). In the looking phase, a repeated-measures ANOVA with Condition and AOI on the first-look responses to the target/distractor objects revealed no significant effects. A repeated-measures ANOVA with Condition and AOI on the viewing-time responses to the target/distractor objects revealed a significant main effect of AOI ($F(1,5) = 6.93$, $p = 0.046$, $\eta^2 = 0.58$), but no significant main effect of Condition or interaction effect. Again, bonobos spent more time looking at the target than the distractor object across conditions.

In Experiment 3, we tested the same 13 bonobos from WKPRC and KS with the same stimuli and procedure as described in the main text (Figure 1C). A repeated-measures ANOVA with Condition and Phase on the viewing times for the actor's face confirmed a significant main effect of Phase ($F(2,24) = 53.97$, $p < 0.001$, $\eta^2 = 0.82$), but no other significant main/interaction effects. In the looking phase, a repeated-measures ANOVA with Condition and AOI on the first-look responses to the target/distractor objects revealed a significant main effect of AOI ($F(1,12) = 5.98$, $p = 0.031$, $\eta^2 = 0.33$), but no significant

main effect of Condition or interaction effect. A repeated-measures ANOVA with Condition and AOI on the viewing-time responses to the target/distractor objects revealed a significant main effect of AOI ($F(1,12) = 6.71, p = 0.024, \eta^2 = 0.36$), but no significant main effect of Condition or interaction effect. Bonobos looked first and looked longer to the target than the distractor object across conditions.

Orangutan

In Experiment 1, we tested 7 orangutans from WKPRC with the same stimuli and procedure as described in the main text (Figure 1A). A repeated-measures ANOVA with Condition and Phase on the viewing times for the actor's face confirmed a significant main effect of Phase ($F(2,12) = 33.61, p < 0.001, \eta^2 = 0.85$), but no other significant main/interaction effects. Like the other species, orangutans attended to the actor's face more in the cueing phase than in the other phases. In the looking phase, a repeated-measures ANOVA with Condition and AOI on the first-look responses to the target/distractor objects revealed no significant effects. A repeated-measures ANOVA with Condition and AOI on the viewing-time responses to the target/distractor objects revealed a significant main effect of Condition ($F(1,6) = 16.41, p = 0.007, \eta^2 = 0.73$), but not a significant main effect of AOI or interaction effect. Like chimpanzees, orangutans spent more time looking at the objects following the ostensive cue than the non-ostensive cue.

Experiment 2 was skipped with this species due to the unavailability of the participants at that time.

In Experiment 3, we tested the same 6 orangutans from WKPRC with the same stimuli and procedure as described in the main text (Figure 1C). A repeated-measures ANOVA with Condition and Phase on the viewing times for the actor's face confirmed a significant main effect of Phase ($F(2,10) = 50.87, p < 0.001, \eta^2 = 0.91$), but not the other significant main/interaction effects. Orangutans attended to the actor's face more in the cueing phase than in the other phases. In the looking phase, a repeated-measures ANOVA with Condition and AOI on the first-look responses to the target/distractor objects revealed a significant main effect of AOI ($F(1,5) = 7.35, p = 0.042, \eta^2 = 0.60$), but no significant main effect of Condition or interaction effect. A repeated-measures ANOVA with Condition and AOI on the viewing-time responses to the target/distractor objects revealed a significant main effect of AOI ($F(1,5) = 11.20, p = 0.020, \eta^2 = 0.69$), but no significant main effect of Condition or interaction effect. Like bonobos, orangutans looked first and looked longer to the target than the distractor object across conditions.

Follow-up experiments

We conducted two follow-up experiments in response to the reviewers' comments. The first comment was increasing the number of distractors, and/or presenting different (rather than identical) objects as the distractors may change the results. To address this comment, we tested the 8 PRI chimpanzees with stimuli identical to those used in experiment 3 except we used 3 different objects as distractors (Figure S4A). We found the similar results (Figure S4B, see its caption for the statistical results) regardless of those changes.

The second comment was regarding how human adults would perform in the same task. That is, the previous human studies have tested only young infants but not human adults. Human adults may perform differently from human infants, and the results with human adults may not be different from those with great apes. To address this issue, we tested 20 human adults (13 females, all Japanese) using the stimuli similar to those used with infants (Figure S5A). We tested the participants in a standard office and they provided written informed consent before participating in the study. We found that human adults indeed performed differently from human infants tested in previous studies, namely they followed the actor's gaze in both conditions (Figure S5B). No difference was observed between conditions (except the significant tendency in the first-look measure; see Figure S5 caption for statistical results). However, unlike apes, their performance seemed to reach ceiling levels in both conditions. Our chimpanzee participants did not respond to the actor's gaze as sensitively as did human adults; thus, the results were not equivalent between chimpanzees and humans. The institute-reared chimpanzees or bonobos were more sensitive than the zoo-reared chimpanzees to the actor's gaze. Nonetheless, we found no evidence that the ostensive signals enhanced the gaze-following responses in any of those ape groups.

Thus, although we confirmed that human adults responded differently from human infants and apes, the ceiling effects are problematic, potentially precluding the appearance of any potential differences between conditions. Also, it is possible that, as pointed out by Szufnarowska et al. (2014) and Gredebäck et al. (2018), obtaining attention to the actor's face during the cuing phase may be sufficient to elicit gaze following in humans. Addressing this issues is beyond the scope of this study but should be addressed in future comparative and developmental studies. One idea for future research is to use cinema-like stimuli in which the actor's gaze cues are embedded within a series of actions in a complex scenario that mimics the natural contexts where human adults typically see others' ostensive signals.

Supplemental Tables and Figures

Table S1. Participant information.

Species	Participated	Name	Facility	Age Class	Rearing history
Chimpanzee	Exp. 1	Kara	WKPRC	Adult	Mother
Chimpanzee	Exp. 1	Lome	WKPRC	Adult	Mother
Chimpanzee	Exp. 1	Natasha	WKPRC	Adult	Mother
Chimpanzee	Exp. 1 & 3	Alex	WKPRC	Adult	Nursery/peer
Chimpanzee	Exp. 1 & 3	Bangolo	WKPRC	Juvenile	Mother
Chimpanzee	Exp. 1 & 3	Daza	WKPRC	Adult	Mother
Chimpanzee	Exp. 1 & 3	Frederike	WKPRC	Adult	Mother
Chimpanzee	Exp. 1 & 3	Fraukje	WKPRC	Adult	Nursery/peer
Chimpanzee	Exp. 1 & 3	Jahaga	WKPRC	Adult	Mother
Chimpanzee	Exp. 1 & 3	Jeudi	WKPRC	Adult	Mother
Chimpanzee	Exp. 1 & 3	Kofi	WKPRC	Adult	Mother
Chimpanzee	Exp. 1 & 3	Lobo	WKPRC	Adult	Mother
Chimpanzee	Exp. 1 & 3	Riet	WKPRC	Adult	Nursery/peer
Chimpanzee	Exp. 1 & 3	Robert	WKPRC	Adult	Mother
Chimpanzee	Exp. 1 & 3	Sandra	WKPRC	Adult	Mother
Chimpanzee	Exp. 2	Ai	PRI	Adult	Wildborn/peer
Chimpanzee	Exp. 2	Akira	PRI	Adult	Wildborn/peer
Chimpanzee	Exp. 2	Chloe	PRI	Adult	Mother
Chimpanzee	Exp. 2	Cleo	PRI	Adult	Mother
Chimpanzee	Exp. 2	Mari	PRI	Adult	Wildborn/peer
Chimpanzee	Exp. 2	Pal	PRI	Adult	Mother
Chimpanzee	Exp. 2	Pan	PRI	Adult	Nursery/peer
Chimpanzee	Exp. 2 & 3	Misaki	KS	Adult	Mother
Chimpanzee	Exp. 2 & 3	Hatsuka	KS	Juvenile	Nursery/peer
Chimpanzee	Exp. 2 & 3	Iroha	KS	Juvenile	Mother
Chimpanzee	Exp. 2 & 3	Mizuki	KS	Adult	Nursery/peer
Chimpanzee	Exp. 2 & 3	Natsuki	KS	Adult	Mother
Chimpanzee	Exp. 3	Frodo	WKPRC	Adult	Mother

Chimpanzee	Exp. 3	Tai	WKPRC	Adult	Mother
Orangutan	Exp. 1	Batak	WKPRC	Juvenile	Mother
Orangutan	Exp. 1 & 3	Dokana	WKPRC	Adult	Mother
Orangutan	Exp. 1 & 3	Padana	WKPRC	Adult	Mother
Orangutan	Exp. 1 & 3	Pini	WKPRC	Adult	Nursery/peer
Orangutan	Exp. 1 & 3	Raja	WKPRC	Adult	Mother
Orangutan	Exp. 1 & 3	Suaq	WKPRC	Juvenile	Mother
Orangutan	Exp. 1 & 3	Tanah	WKPRC	Juvenile	Mother
Bonobo	Exp. 1	Jasongo	WKPRC	Adult	Mother
Bonobo	Exp. 1	Joey	WKPRC	Adult	Nursery/peer
Bonobo	Exp. 1	Luiza	WKPRC	Adult	Mother
Bonobo	Exp. 1 & 3	Fimi	WKPRC	Juvenile	Mother
Bonobo	Exp. 1 & 3	Gemena	WKPRC	Adult	Mother
Bonobo	Exp. 1 & 3	Kuno	WKPRC	Adult	Nursery/peer
Bonobo	Exp. 1 & 3	Yasa	WKPRC	Adult	Mother
Bonobo	Exp. 2 & 3	Ikela	KS	Adult	Nursery/peer
Bonobo	Exp. 2 & 3	Junior	KS	Adult	Mother
Bonobo	Exp. 2 & 3	Connie-Lenore	KS	Adult	Mother
Bonobo	Exp. 2 & 3	Lolita	KS	Adult	Nursery/peer
Bonobo	Exp. 2 & 3	Louise	KS	Adult	Nursery/peer
Bonobo	Exp. 2 & 3	Vijay	KS	Adult	Nursery/peer
Bonobo	Exp. 3	Kasai	WKPRC	Juvenile	Mother
Bonobo	Exp. 3	Lexi	WKPRC	Adult	Nursery/peer
Bonobo	Exp. 3	Yaro	WKPRC	Juvenile	Mother

* Rearing history: whether the ape participants were mainly reared by their biological mothers (mother) or human caregivers and nonrelated conspecifics (nursery/peer), or wild-born (then reared by human caregivers and nonrelated conspecifics).

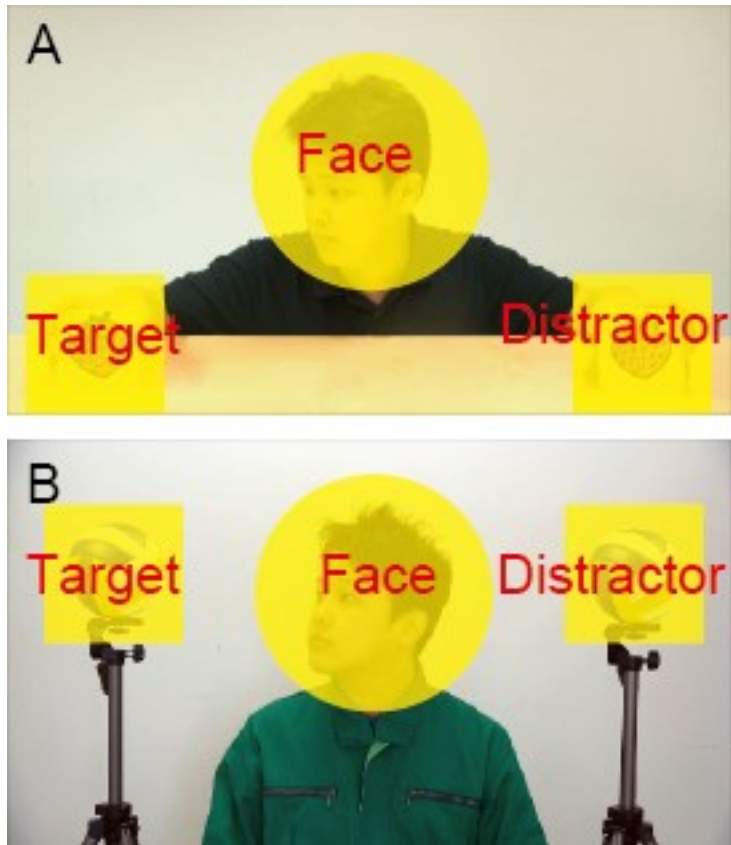


Figure S1. A. Areas-Of-Interest defined for the actor's face, the looking target and the distractor in Experiments 1 and 2 and B. in Experiment 3.

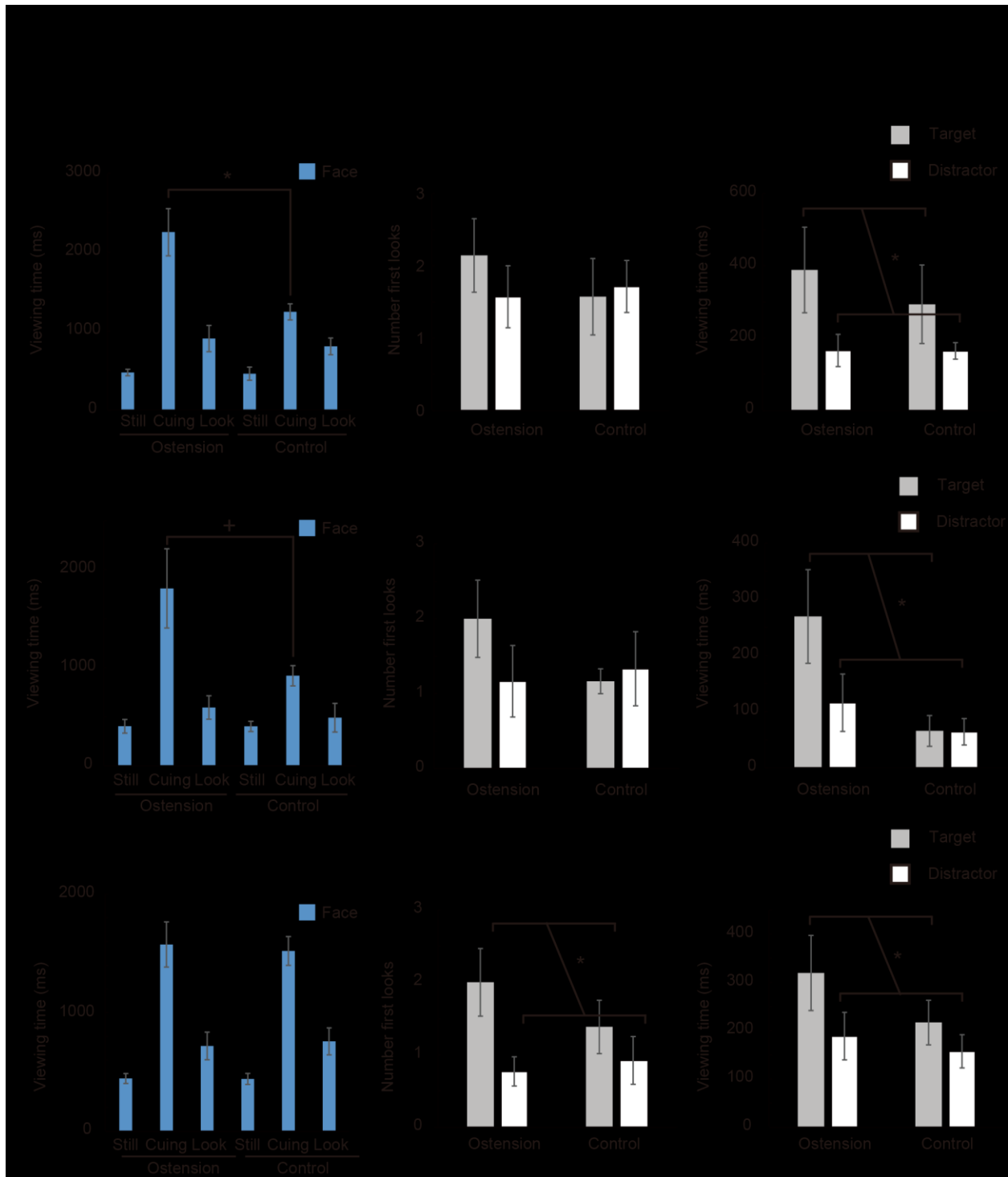


Figure S2. Results for bonobos from Experiment 1-3 (A-C). Mean viewing times for the actor's face in each phase of each condition (in 2.5 seconds), and number of first looks (in 6 trials) to the target or distractor object, and mean viewing times for the target or distractor object (in 5 seconds). + $p < 0.1$, * $p < 0.05$.

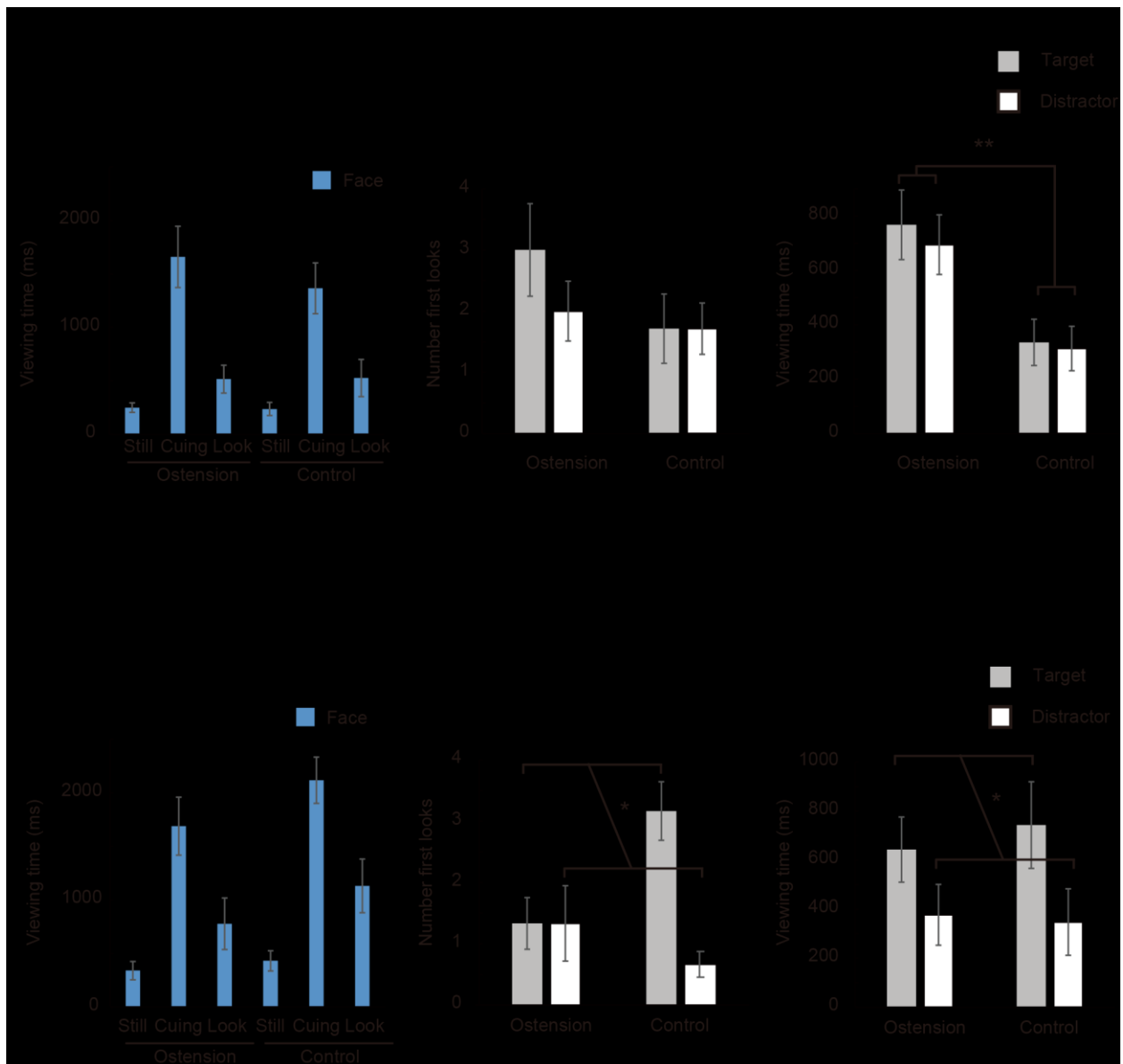


Figure S3. Results for orangutans from Experiment 1 and 3 (A-B). Mean viewing times for the actor's face in each phase of each condition (in 2.5 seconds), and number of first looks (in 6 trials) to the target or distractor object, and mean viewing times for the target or distractor object (in 5 seconds). * $p < 0.05$, ** $p < 0.01$.



Figure S4. Results for a follow-up test with the 8 PRI chimpanzees (institute-reared). The stimuli were identical to experiment 3's except that different objects were used as the distractors and the number of distractors were increased to 3. A repeated-measures ANOVA with Phase (Still, Cuing, Look) and Condition (Ostension, Control) on the viewing times for the actor's face revealed a significant main effect of Phase ($F(2,14) = 32.95, p < 0.001, \eta^2 = 0.83$). A repeated-measures ANOVA with AOI (Target, Distractor1, Distractor2, Distractor3) and Condition (Ostension, Control) on the viewing times for the objects revealed a significant main effect of Condition ($F(1,7) = 14.80, p = 0.006, \eta^2 = 0.68$). The other main/interaction effects were not significant. Therefore, these results were similar to those from Experiment 3.

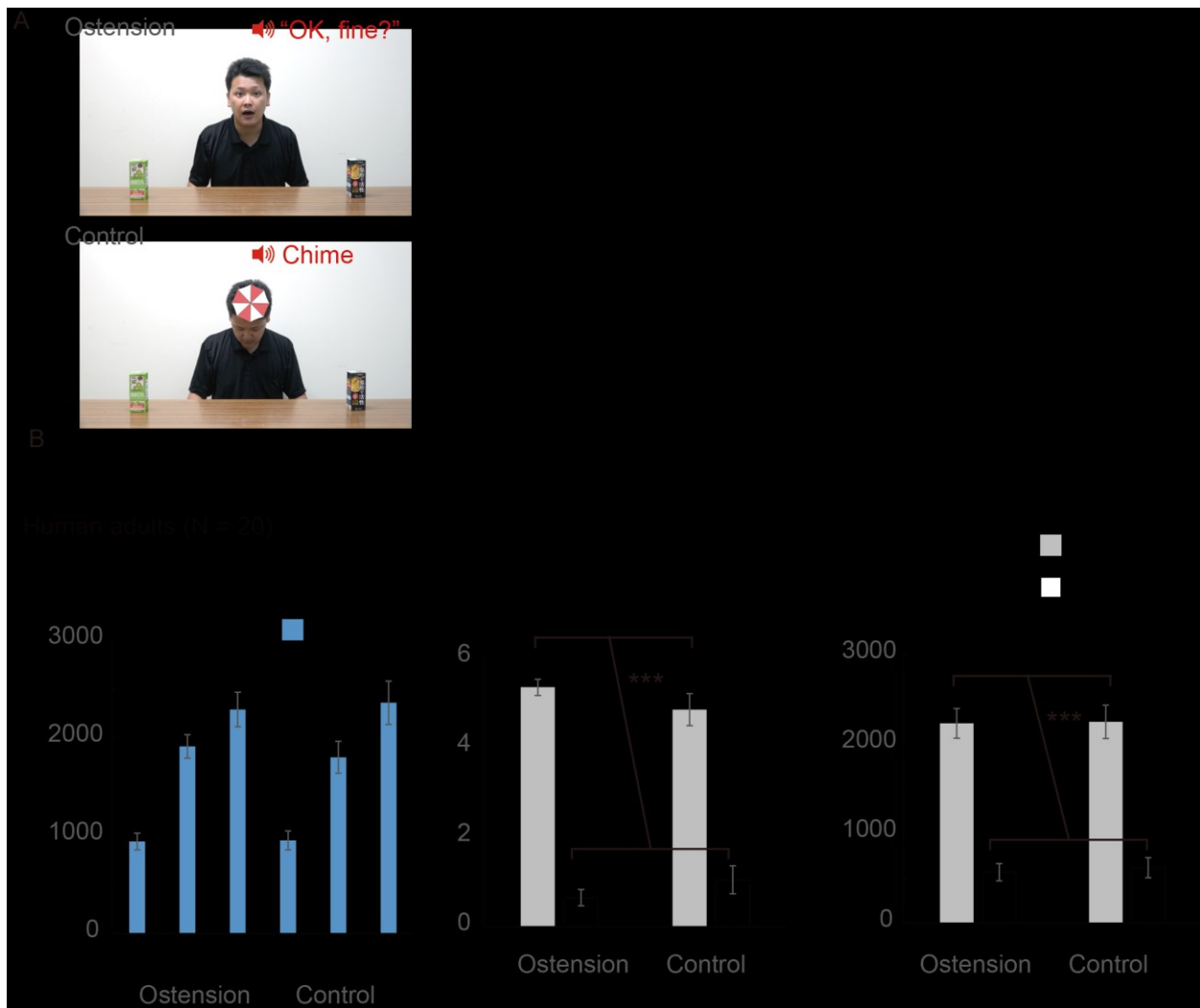


Figure S5. Results for a follow-up test with 20 human adults. The stimuli were the same as those used in Experiment 2 (and therefore similar to those used in the original infant study). The actor made eye contact and said non-directional communicative words (“OK, fine?” in Japanese) in the ostension condition. A visually-salient object and a chime attracted the participants’ attention in the control condition. A repeated-measures ANOVA with Phase (Still, Cuing, Look) and Condition (Ostension, Control) on the viewing times for the actor’s face revealed a significant main effect of Phase ($F(2,38) = 43.81, p < 0.001, \eta^2 = 0.70$). A repeated-measures ANOVA with AOI (Target, Distractor) and Condition (Ostension, Control) on the first looks to the objects revealed a significant main effect of AOI ($F(1,19) = 81.96, p < 0.001, \eta^2 = 0.70$) and a marginally significant interaction effect of Condition and AOI ($F(1,19) = 3.08, p = 0.095, \eta^2 = 0.14$). A repeated-measures ANOVA with AOI (Target, Distractor) and Condition (Ostension, Control) on the viewing times for the objects revealed a significant main effect of AOI ($F(1,19) = 77.32, p < 0.001, \eta^2 = 0.80$). The other main/interaction effects were not significant. We also

analyzed the same data with the difference-score analysis (used in the original study with infants) and found similar results [first look: ostension (mean \pm SE), 0.76 ± 0.06 , control, 0.66 ± 0.09 , ostension vs. control, $t(19) = 1.91$, $p = 0.07$, total looks: ostension, 0.75 ± 0.08 , control, 0.56 ± 0.13 , ostension vs. control, $t(19) = 1.27$, $p = 0.22$].