1

2

3

4

5

6 Research article

7 Great apes generate goal-based action predictions by eye movements

8 Fumihiro Kano (1, 2)

9 Josep Call (1, 3)

10

11 1. Department of Developmental and Comparative Psychology, Max-Planck Institute

12 for Evolutionary Anthropology, Leipzig, Germany

13 2. Japan Society for Promotion of Science, Tokyo, Japan

14 3. School of Psychology and Neuroscience, University of St Andrews, UK

15

16 Corresponding Author's address

17 fkanou@gmail.com

18 Max-Planck Institute for Evolutionary Anthropology, Leipzig, Germany

19

20

21 **Abstract**

22 To examine great apes' online prediction of other's actions, this study used an eye-tracking

23 technique and an experimental paradigm previously used to test human infants. Twenty-two

24 great apes including bonobos, chimpanzees, and orangutans were familiarized to movie

25 clips of a human hand reaching to grasp one of two objects. Then object locations were

26 swapped, and the hand made an incomplete reach between the objects. In a control

27 condition, a mechanical claw performed the same action. Apes predictively looked at the

28 familiarized goal object rather than the familiarized location when viewing the hand action.

29 However, they did make no prediction when viewing the claw action. These results were

30 similar to those reported previously in human infants, and there was no species difference

31 among great apes. Thus, great apes make online goal-based prediction about the other's

32 action, which is not a unique skill among humans but is shared more widely among

33 primates.

34 Key words: action prediction; eye-tracking; nonhuman primates; proactive goal-directed eye

35 movements

36

Introduction

Our eyes move proactively rather than reactively when we perform manual action as well as when we observe manual action performed by the other individual (Falck-Ytter, Gredebäck, & Von Hofsten, 2006; Flanagan & Johansson, 2003; Johansson, Westling, Bäckström, & Flanagan, 2001; Land & Furneaux, 1997) That is, our eyes spontaneously seek for the goal object which is slightly ahead of our and the other's action. This online spontaneous prediction about the other's action should have immense functions in our dynamic social world. For example, it helps us to perform smooth coordination with others and also to have an advantage in competition with others. Also, by similarly attending to the other's and own action, we may encode the other's action in a similar way with our own action and thereby enhance the understandings of the other's action.

Ontogenetically, this proactive goal-directed eye movement during action observation emerges around the 6 month of age in humans (Falck-Ytter, et al., 2006; Kanakogi & Itakura, 2011). Importantly, by the first year of life, those online predictions are performed in a non-reflexive, goal-based manner. For example, in Cannon and Woodward (2012), when 11-month-old infants had seen a person reach for two objects, they looked predictively toward the object for which the person had previously reached, even after the objects' locations had changed (an adaptation of Woodward (1998) paradigm for an eye-tracking task). In Southgate, Senju, & Csibra (2007), when 2-year-old children had seen a person approach two containers, they looked predictively toward the one in which the person had previously seen a toy hidden, even though the toy was no longer there. Therefore, during the course of development, humans acquire a sophisticated skill in which they visually predict the other's actions based on the underlying goals and intentions.

Evidence of such proactive goal-directed eye movements in nonhuman species is

61 relatively sparse (Kano & Tomonaga, 2013; Myowa-Yamakoshi, Scola, & Hirata, 2012).

62 Kano and Tomonaga (2013) examined the eye movements of a chimpanzee while she was

63 performing a manual imitation task with an experimenter in a real life setting. They found

64 that the chimpanzee predictively looked at the reward when an experimenter was reaching

65 for it, while she reactively looked at the goal of reaching when the experimenter was

66 performing the task gestures. Myowa-Yamakoshi et al. (2012) found that chimpanzees,

67 12-month-old human infants and adults predictively looked at the goals of various actions

68 presented in the movies, while 8-month-old human infants reactively looked at them. Thus,

69 chimpanzees seem to predict the goals of some actions spontaneously as humans do.

70 However, it remains unclear whether such eye movements are goal-based; i.e. not simply

71 oriented toward the direction of hand/object movements.

72   There are numerous behavioral studies showing that nonhuman primates make

73 predictions about the other's actions based on the perception and knowledge. For example,

74 when a subordinate and a dominant chimpanzee were competing for the two foods, the

75 subordinate avoided the food which the dominant can see and instead chose the food which

76 the dominant cannot see (Hare, Call, & Tomasello, 2000). In addition, chimpanzees helped

77 a human or conspecific partner by taking the object which the partner was trying to reach

78 (Warneken & Tomasello, 2006), possibly by identifying the partner's current needs

79 (Yamamoto, Humle, & Tanaka, 2012). In a study using a looking time measure, when

80 macaque monkeys were habituated to a human experimenter reaching for an object behind

81 a barrier, they were more surprised to see the same yet inefficient movements of arms than

82 *vice versa* after the barrier was removed (Rochat, Serra, Fadiga, & Gallese, 2008); also see

83 Uller (2004). For most of the studies mentioned above, similar results were obtained in

84 human infants. Therefore, like humans, nonhuman primates seem to understand the other's

85 actions not just in terms of surface behaviors but also in terms of the underlying goals and

86 intentions (Call & Tomasello, 2008).

87 However, an important outstanding question is whether nonhuman primates make

88 online goal-based prediction by eye movements when observing the other's action. The aim

89 of the current study was to investigate this aspect in three great ape species (chimpanzees,

90 bonobos, and orangutans) with different degrees of phylogenetic distance to humans.   All

91 great ape species show basic similarities in eye movements and scene scanning (Kano, Call,

92 & Tomonaga, 2012; Kano, Hirata, Call, & Tomonaga, 2011). Since previous studies have

93 reported evidence of goal attribution in chimpanzees, we anticipated that chimpanzees

94 would produce goal-based action prediction. Much less is known about the other species,

95 with one study reporting positive results for orangutans, but not for bonobos (Buttelmann,

96 Carpenter, Call, & Tomasello, 2007, 2008). Therefore, it is unclear whether these other

97 species would spontaneously produce goal-based prediction when observing the other's

98 action, which seems to be a basic skill for general goal attribution.

99 We used the eye-tracking technique and adopted an experimental paradigm

100 previously used with human infants (Cannon & Woodward, 2012). Apes were familiarized to

101 a repeated reaching action directed to one of two objects by a human agent. The locations

102 of the objects were then swapped, and apes' predictive looks to the prior location versus the

103 prior goal were examined as the agent made an incomplete reach between the objects (Fig.

104 1). If apes make goal-based predictions, then we expected that they would look predictively

105 toward the prior goal, rather than to the prior location, like human infants. We also included a

106 control condition in which a mechanical claw moved in a similar manner to the hand. The

107 previous study with infants (Cannon & Woodward, 2012) found that infants did not predict

108 the goal of claw actions during the test event, although the grasping action of both claw and

109   hand caught their attention strongly during familiarization. Thus, this control allowed us to

110   rule out the possibility that participants looked at the prior goal during the test event simply

111   because they attended to a salient grasping action and thereby formed an association

112   between the hand/claw and object during familiarization.

113
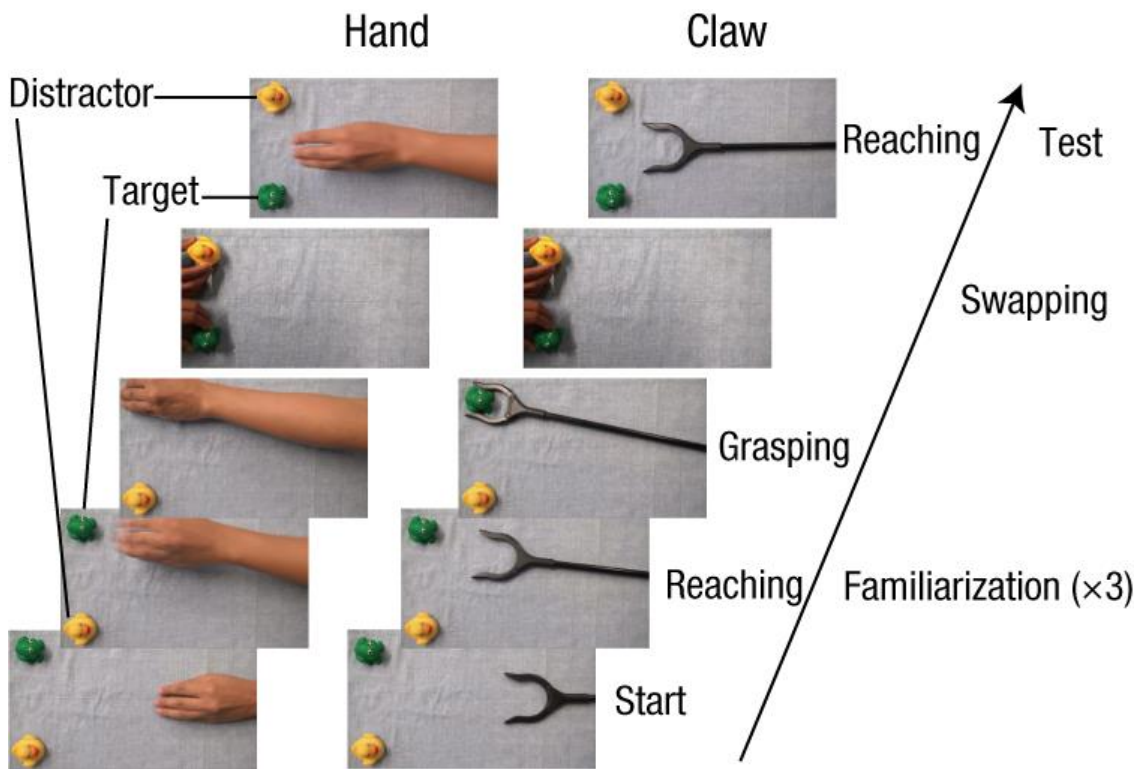
Figure 1. Procedure of a trial. After the three familiarization events in which the hand or claw

reaches for and grasps the target, the locations of objects were swapped, and then the hand

or claw reaches strait between the two objects and pauses in front of, and equidistantly from

the two objects.

121 **Methods**

122 **Participants**. Twenty-two great apes participated in this study (4 bonobos, *Pan paniscus,*

123 12 chimpanzees, *Pan troglodytes,* 6 orangutans, *Pongo abelii;* 12 females, 10 males; mean

124 16.0 ± S.D. 11.5 years of age; for the details, see Table S1). Five additional apes were

125 tested but excluded from the analysis because they attended insufficiently to the stimuli

126 (excessive off-screen fixations (n = 4); excessive eccentric eye movements (n = 1)). All apes

127 live with their conspecifics in semi-natural indoor and outdoor enclosures at the Wolfgang

128 Köhler Primate Research Center (WKPRC). All apes were tested in testing rooms at the

129 WKPRC, and their daily participations of experiments were voluntary. They were given

130 regular feedings, daily enrichment and water ad lib. Animal husbandry and research

131 complied with the EAZA Minimum Standards for the Accommodation and Care of Animals in

132 Zoos and Aquaria and the WAZA Ethical Guidelines for the Conduct of Research on Animals

133 by Zoos and Aquariums, respectively.

134

135 **Apparatus**. The eye movements of apes were non-invasively recorded with an infrared

136 eye-tracker (60 Hz; Tobii X120, Tobii Technology AB, Stockholm, Sweden). Stimuli were

137 presented using Tobii Studio software in a 22-inch LCD monitor (1366×768 pixel) at a 60-cm

138 viewing distance (1 degree of gaze angle corresponded to approximately 1 cm on the

139 monitor). We tested them unrestrained but separated from the experimenter and eye-tracker

140 with a transparent acrylic panel. However, in order to keep their heads relatively still, we

141 implemented a nozzle and tube attached to the acrylic panels, which produce a dripping of

142 grape juice, and let the apes suck the nozzle during the recording (see Fig. S1). No explicit

143 training was conducted for the apes.

144

145   **Calibration**. Two-point automated calibration was conducted by presenting a small object

146   on each reference point. Relatively small numbers of reference points were used in this

147   study because the apes tended to view those reference points only shortly. However, we

148   manually checked the accuracy at five points after the initial calibration and repeated the

149   calibration if necessary. As a result, we obtained the comparable accuracy between ape and

150   human participants; a validation session with more than 10 apes and 10 humans under

151   standard calibration results confirmed that the positional error values estimated at the five

152   spots on the screen were less than 1-2 degree. Before every test session, we checked the

153   accuracy manually and started the session when we confirmed the error value less than 1-2

154   degree around center on the screen.

155

156   **Stimuli and procedure**. We created our videos by slightly modifying those used in the

157   previous study with infants (Cannon & Woodward, 2012). Overall, we slightly shortened the

158   length of video (speeding up the movements) to match the apes' relatively rapid shift of

159   attention (Kano, et al., 2011). On each trial, we presented a 13160 millisecond (ms) video

160   (resolution 1280×720 pixel) on the center of monitor, which contained a rubber toy duck

161   (yellow) and frog (green) in the scene (Fig. 1). The video comprised three familiarization

162   events (2230 ms each), one swap event (2550 ms), and one test event (2920 ms). During

163   the familiarization event, either a human left hand (hand condition) or a plastic rod with claw

164   (claw condition) appeared in the mid-right corner of scene (0-680 ms), reached to (reaching

165   phase; 681-1210 ms), and grasped (grasping phase; 1211-2230 ms) one of the objects

166   (target). The hand or claw moved by following a curvilinear path from the start to the object.

167   Grasping the frog and duck accompanied a croak and quack sound, respectively. During the

168   swap event, the other two hands (left and right hands) grasped the two objects

169 simultaneously and swapped the locations of them. Finally, during the test event, either a

170 hand or a claw appeared (0-1020 ms), reached straight, and paused in front of, and

171 equidistantly from the two objects (1021-2919 ms).

172 Each ape viewed one video (one trial) per day, for a total of 4 videos of hand

173 condition and 4 videos of claw condition (i.e. within-subject design) (8 trials, over 8 days).

174 Each ape competed either hand or claw condition first and then completed the other

175 condition, and its order was counterbalanced across apes. Also, the object location (up or

176 down) or the object type (duck or frog) during the familiarization was counterbalanced

177 across apes. To prevent the hand orientation (the thumb downward/upward) from serving as

178 a cue to predict the hand action during the test event, we used the same hand during the

179 familiarization and test events. Therefore, if the participants learned from the familiarization

180 event to respond to thumb orientation, their action prediction during the test events should

181 be based on the location but not on the object (after swapping the locations of objects). The

182 initial hand orientation in relation to the target was counterbalanced across apes.

183 The experimenter initiated the presentation of each video when apes were

184 attending to the monitor. In those cases where apes moved away from the monitor during

185 the recording (thereby severely disrupted the eye-tracking signals), we repeated the same

186 video on the next day. This occurred only once in 4 apes. All apes completed all trials.

187

188 **Data analysis**. For the familiarization events, the apes' predictive look was estimated by (1)

189 the viewing time during the reaching phase (from the onset of hand/claw movement to the

190 touch to the target). We also examined (2) the viewing time for each object during the

191 grasping phase (from the touch to the target to the end of each familiarization event) in order

192 to examine whether the participants similarly attended to the target and distractor (i.e. the

193 other object) in both hand and claw conditions. For the test events, the apes' predictive look

194 was estimated by (1) the viewing time during reaching (from the onset of hand/claw

195 movements to the end of video). As there was a slight variation of phase durations among

196 stimulus videos (familiarization reaching phase; 506 ms ± 110; familiarization grasping

197 phase; 1023 ms ± 107 test reaching phase; 1899 ms ± 51, mean duration ± SD), the viewing

198 times were rescaled to the values proportional to the mean duration of each phase. Apes

199 viewed off-screen areas for 22 % ± 10.3 (mean ± SD) of total video duration. When the ape

200 fixated off-screen areas for the entire duration of each phase, we coded those fixations as

201 missing values (4.5 % of all data). The apes' predictive look was additionally estimated by

202 (2) the proportion of trials in which apes looked at the target first (vs. the distractor first)

203 during the test event. We analyzed only the trials in which apes viewed either target or

204 distractor (55 % of trials). Two apes and one ape did not view the objects in the hand and

205 claw condition across 4 trials, respectively, and we excluded them from the analysis (i.e. we

206 analyzed 19 apes for this measure).

207 A square-shaped AOI, 200×200 pixels in size, was defined for each of the target

208 and distractor (both approximately 150×150 pixels in size). In addition, an AOI was defined

209 for the trajectory area in which the hand or claw moved (Fig. S2). The fixation filtering was

210 conducted using the Tobii fixation filter (version 3.2.1). The above-mentioned measurements

211 were calculated using the Tobii Studio and Matlab (MathWorks, USA). Statistical analyses

212 were conducted in SPSS version 20.

213

214

Results

Familiarization. We conducted our initial analysis for the viewing time, using a repeated-measures ANOVA with Condition (hand, claw), Object (target, distractor) as within-subject factors and Species (bonobo, chimpanzee, orangutan) as between-subject factors, respectively for reaching and grasping phase during the familiarization (averaged over three events; Fig. 2a). We confirmed that the object location (up or down) or the object type (duck or frog) which was included into the ANOVA did not significantly influence the apes' viewing time during the familiarization, neither the main effect nor interactions with the other factors, neither during the reaching nor grasping phase ($ps > 0.23$). We therefore did not include this counterbalanced factor in the subsequent analyses.

We found a significant interaction between Condition and Object in the viewing time when the hand or claw was reaching for the target during the familiarization [$F(1,19) = 8.42$, $p = 0.009$, $\eta^2 = 0.30$]. We also found the main effect of Species [$F(2,19) = 3.87$, $p = 0.039$, $\eta^2 = 0.29$]. The other main effects or interactions were not significant. Post-hoc tests revealed that apes viewed the target for a longer time than the distractor in hand condition [$t(21) = 2.58$, $p = 0.017$, Cohen's $d = 0.57$] but not in the claw condition [$t(21) = 0.51$, $p = 0.61$, Cohen's $d = 0.16$] during this period. Also, there was a significant difference between conditions in the viewing time for the target [$t(21) = 2.64$, $p = 0.015$, Cohen's $d = 0.56$] but not for the distractor [$t(21) = 0.03$, $p = 0.97$, Cohen's $d = 0.007$]. Thus, these results revealed that apes predicted the target of a hand action but not that of a claw action during the familiarization.

To examine the presence of predictive saccades to the target, we measured the timing of gaze arrival at the target relative to the start of grasping (the end of reach) in the first of three familiarization events. On average, apes viewed the target 12 ms ($SD = 411$)

12

239 before and 213 ms (SD = 771) after the start of first grasping, in the hand and claw condition,

240 respectively. Compared with the ordinary saccadic reaction time in these species (Kano, et

241 al., 2011), minimum 200 ms, they viewed the hand action proactively [$t(21)$ = 2.41, p = 0.025,

242 Cohen's d = 0.51], while they did the claw action reactively [$t(21)$ = 0.08, p = 0.93, Cohen's d

243 = 0.01].

244   In contrast to the reaching phase, there was no interaction between Condition and

245 Goal in the viewing time when the hand or claw was grasping the target during the

246 familiarization [$F(1,19)$ = 1.17, p = 0.29, η^2 = 0.058]. Thus, apes attended to the target

247 versus distractor with the claw and hand similarly during this period. We also found the

248 interaction between Object and Species [$F(2, 19)$ = 5.39, p = 0.014, η^2 = 0.36] and the main

249 effects of Object [$F(1,19)$ = 143.01, p < 0.001, η^2 = 0.88] and Species [$F(2, 19)$ = 6.31, p =

250 0.008, η^2 = 0.39]. The other main effects or interactions were not significant.

251

252 **Test.** Crucially, during the test event, we found a significant interaction between Condition

253 and Object in the viewing time when the hand or claw was reaching between the two objects

254 [$F(1,19)$ = 7.48, p = 0.013, η^2 = 0.28]. We also found the main effect of Species [$F(2,19)$ =

255 7.93, p = 0.003, η^2 = 0.45]. The other main effects or interactions were not significant.

256 Post-hoc tests revealed that apes viewed the target for a longer time than the distractor in

257 the hand condition [$t(21)$ = 2.50, p = 0.020, Cohen's d = 0.53] but not in the claw condition

258 [$t(21)$ = 1.46, p = 0.15, Cohen's d = 0.27] during this period. Also, there was a significant

259 difference between conditions in the viewing time for the target [$t(21)$ = 3.75, p = 0.001,

260 Cohen's d = 0.80] but not for the distractor [$t(21)$ = 0.68, p = 0.50, Cohen's d = 0.14]. The

261 order of presentation (hand or claw) had no significant effect when we included it into the

262 ANOVA (neither a main effect nor interactions with the other factors; ps > 0.22). Thus, these

263 results show that apes predicted the target of hand action based on the object but not on the

264 location.

265   Including the factor Trial (1st-4th) in the analysis did not change the results. In

266 addition, although apes viewed the claw trajectory generally longer than the hand trajectory,

267 their viewing times for the objects (i.e. target plus distractor) were similar between the claw

268 and hand conditions (Fig. S3).

269   To examine the initial responses to the target/distractor when the hand or claw

270 starts reaching, we examined the proportion of trials in which apes looked at the target first

271 (vs. the distractor first) (Fig. 2b). We found that apes' first look to the target was more

272 frequent in the hand than claw condition [$t(18) = 2.19$, $p = 0.041$]. Post-hoc tests revealed

273 that the first look to the target was more frequent than chance (0.5) in the hand condition

274 [$t(20) = 3.47$, $p = 0.002$] but not in the claw condition [$t(19) = 0.15$, $p = 0.88$].

275

276 **Species difference.** For the analyses mentioned above, species differences emerged in the

277 strength of viewing the objects (i.e. the main effect of Species in the above analyses) but not

278 in the pattern of predictive look (i.e. the interaction between Species and Condition/Object).

279 In general, orangutans looked at the objects longer (thus looked at the trajectory area

280 shorter) than the other two species (see Fig. S4). Including sex (male, female) and age

281 (young ≤ 9, adult > 9 years of age) into the analyses did not change the results.

282
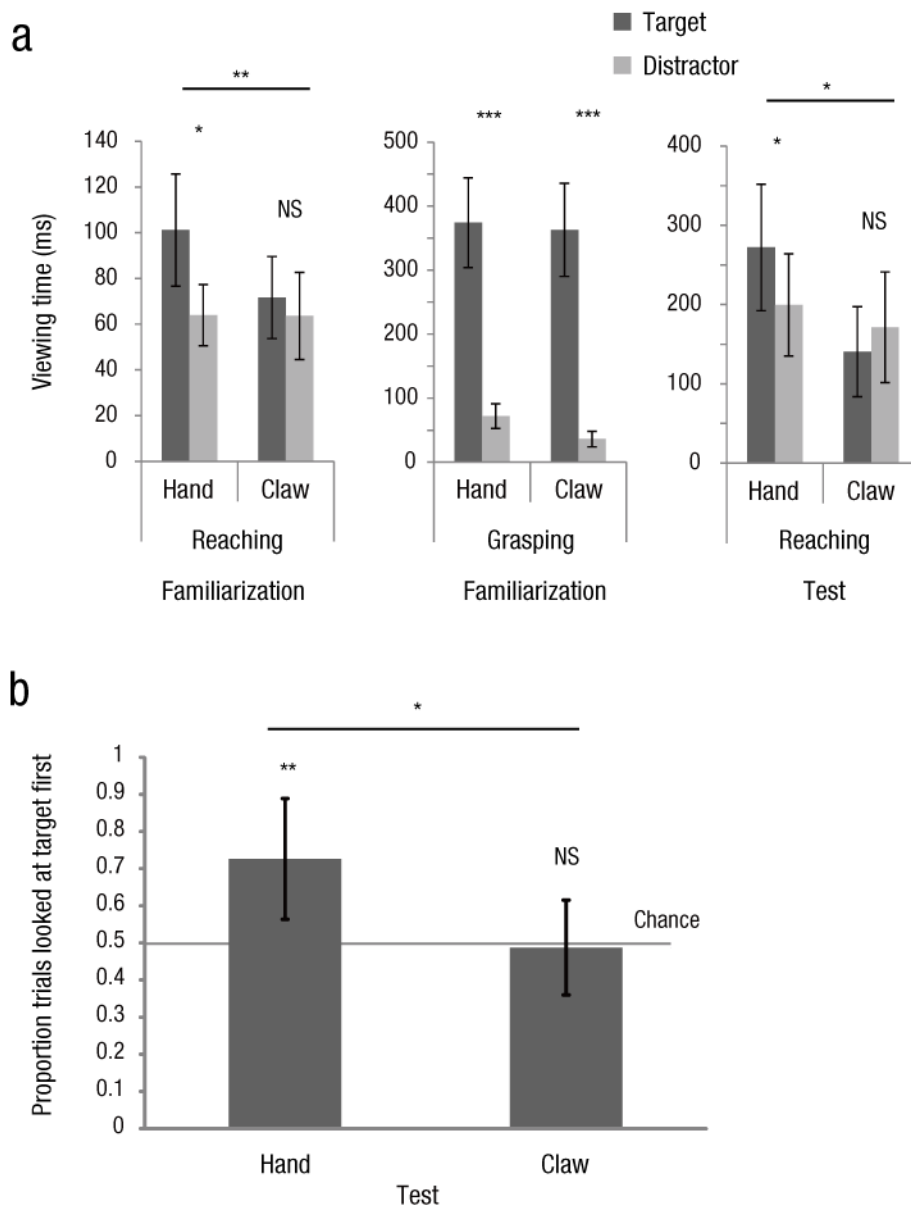
Figure 2. (a) Viewing time (ms) when the hand or claw was reaching for the target (duration: 506 ms) and grasping the target (dur. 1023 ms) during the familiarization or was reaching between the two objects (dur. 1899 ms) during the test events. (b) Proportion of trials in which the apes looked at the target first (vs. the distractor first) during the test events. Error bars denote 95 % confidence intervals. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

290

291

292 **Discussion**

293 Great apes' eye movements were proactive when viewing the reaches of a person but

294 reactive when viewing the reaches of a mechanical claw. Such proactive eye movements

295 were goal-directed, and not simply oriented toward the direction of movements. That is, after

296 apes viewed a person reaching to grasp one of the two target (familiarization event), and

297 then saw that the objects' locations were swapped (swap event), they predicted that the

298 person's subsequent reaches would be directed to the prior goal (test event). In contrast to

299 this hand condition, they did not make any prediction when viewing the same actions of claw

300 during the test event.

301   Great apes' goal-based prediction about hand action is unlikely to result from

302 simple learning of the entire action sequence because apes never viewed completed hand

303 actions once the target had been moved during the test event. It is also unlikely to result

304 from a simple association of grasping action and object during the familiarization because

305 they similarly and strongly attended to the target being grasped in both hand and claw

306 conditions. Finally, the absence of goal-based prediction about claw action is unlikely to

307 result from their excessive attention to the unfamiliar claw stimulus (i.e. the failure to

308 disengage attention from the claw itself) because apes viewed both target and distractor in

309 both hand and claw conditions equally and, furthermore differentially viewed the target

310 versus distractor in the former but not in the latter condition (i.e. Object × Condition). Thus,

311 apes' goal-based action prediction seemed to depend on the familiarity of the agents'

312 goal-directed behaviors rather than the agents' saliency or movement *per se.*

313   The current findings are strikingly similar to those of human infants reported

314 previously (Cannon & Woodward, 2012). Also, in this study, although ape species (bonobos,

315 chimpanzees, and orangutans) differed in their overall viewing patterns (the strength of

316   viewing the objects vs. the agent), their pattern of predictive looks did not differ between the

317   species. Thus, we conclude that online goal-based action prediction is not uniquely human,

318   but are shared more widely among hominoids. The only potential difference between this

319   and previous studies with infants is that, when viewing the claw actions during the test event,

320   infants made location-based action predictions (i.e. more frequent first look to the distractor

321   than to the target), while apes made no significant prediction (i.e. chance level). However,

322   although not significant but consistent with a location-based prediction, apes tended to view

323   the distractor longer than target object in the claw condition. Nevertheless, both

324   location-based and chance-level prediction are theoretically plausible given that human

325   infants did not distinguish the outcomes of the claw action in the looking-time paradigm

326   (Woodward, 1998). Future studies are needed to confirm whether this potential difference

327   reflects merely methodological issues or the species difference. In any case, the most

328   important result of this and previous studies is that apes and infants made the goal-based

329   prediction only in the hand condition, although the claw action similarly captured their

330   attention when the claw/hand was grasping the target.

331   Why then do great apes seem to predict the goal of a hand but not a claw action?

332   One possibility is that action familiarity shapes action understanding in great apes, as shown

333   in many previous studies with macaque monkeys and humans (Falck-Ytter, et al., 2006;

334   Flanagan & Johansson, 2003; Rochat, et al., 2008). For example, in human infants, the

335   first-person motor experience (rather than the visual experience) of an unfamiliar tool-use

336   action helped infants to understand the goal of that action (Sommerville, Hildebrand, &

337   Crane, 2008). The possible mechanism underlying these phenomena is a direct matching

338   process, in which observed actions are mapped onto motor representations of that action

339   (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fogassi, & Gallese, 2001). Thus,

340   this same mechanism may be shared widely in phylogeny (Bonini & Ferrari, 2011; Hecht et

341   al., 2013), facilitating the processing of familiar over unfamiliar actions in these species.

342   However, the previous studies also suggest other possibilities. For example, providing

343   infants with additional abstract movement cues helped them to understand the goal of

344   mechanical object (Biro & Leslie, 2007; Luo & Baillargeon, 2005). Also, after seeing or

345   interacting with a human agent operating the mechanical claw, infants understood the goal

346   of mechanical claw (Gerson & Woodward, 2012; Hofer, Hauf, & Aschersleben, 2005). Future

347   studies should consider these possibilities to further examine the underlying mechanisms for

348   action understanding in great apes and the other nonhuman primates.

349   Does the current finding, the goal-based action prediction in great apes, reflect

350   their understandings of mental or intentional state of others?   Recent evidence and

351   theories suggest that the answer is not necessarily yes (e.g. teleological stance theory;

352   Gergely & Csibra, 2003). That is, they suggest that humans may develop two modes of

353   action interpretations ontogenetically, in which the attribution of goal states to others may

354   precede the attribution of mental states to others. These two modes of action interpretations

355   can be distinguished from one another by examining whether the interpreted actions are

356   based on the true belief or false belief. Thus, it is conceivable that human 11-month-old

357   infants who showed the goal attribution in the current test (Cannon & Woodward, 2012)

358   might fail to pass the nonverbal false-belief task (Southgate, et al., 2007). As the previous

359   studies with great apes have consistently failed to produce positive evidence for nonverbal

360   false-belief attribution (Call & Tomasello, 2008; Kaminski, Call, & Tomasello, 2008), future

361   studies should examine whether they also fail to make online prediction about the other's

362   actions based on false beliefs.

363   In conclusion, by adopting an eye-tracking technique with great apes and the same

364   method from the previous study with infants, we found evidence that great apes also make

365   online goal-based predictions about the other's actions. In line with the other comparative

366   studies, we suggest that humans are not the only hominoids who are sensitive to the other's

367   goal and spontaneously predict the other's action.

368

369   Acknowledgement

373

374   References

375   Biro, S., & Leslie, A. M. (2007). Infants' perception of goal‐directed actions: development

376   through cue‐based bootstrapping. *Developmental Science, 10*(3), 379-398.

377   Bonini, L., & Ferrari, P. F. (2011). Evolution of mirror systems: a simple mechanism for complex

378   cognitive functions. *Annals of the New York Academy of Sciences, 1225*(1), 166-175.

379   Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2007). Enculturated chimpanzees

380   imitate rationally. *Developmental Science, 10*(4), F31-F38.

381   Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2008). Rational tool use and tool choice

382   in human infants and great apes. *Child Development, 79*(3), 609-626.

383   Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later.

384   *Trends in Cognitive Sciences, 12*(5), 187-192.

385   Cannon, E. N., & Woodward, A. L. (2012). Infants generate goal-based action predictions.

386   *Developmental Science, 15*(2), 292-298.

387   Falck-Ytter, T., Gredebäck, G., & Von Hofsten, C. (2006). Infants predict other people's action

388   goals. *Nature neuroscience, 9*(7), 878-879.

389   Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature,*

390   *424*(6950), 769-771.

391   Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor

392   cortex. *Brain, 119*(2), 593-609.

393   Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational

394   action. *Trends in Cognitive Sciences, 7*(7), 287-292.

395   Gerson, S. A., & Woodward, A. L. (2012). A claw is like my hand: Comparison supports goal

396   analysis in infants. *Cognition, 122*(2), 181-192.

397   Hare, B., Call, J., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not

398   see. *Animal Behaviour, 59*, 771-785.

399   Hecht, E., Murphy, L., Gutman, D., Votaw, J., Schuster, D., Preuss, T. M., Orban, G. A., Stout, D.,

400   & Parr, L. (2013). Differences in neural activation for object-directed grasping in

401   chimpanzees and humans. *Journal of Neuroscience, 33*(35), 14117-14134.

402   Hofer, T., Hauf, P., & Aschersleben, G. (2005). Infant's perception of goal-directed actions

403   performed by a mechanical device. *Infant Behavior and Development, 28*(4), 466-480.

404   Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye-hand coordination

405   in object manipulation. *Journal of Neuroscience, 21*(17), 6917-6932.

406   Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not

407   what they believe. *Cognition, 109*(2), 224-234.

408   Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction

409   and motor ability in early infancy. *Nature Communications, 2*, 341.

410   Kano, F., Call, J., & Tomonaga, M. (2012). Face and eye scanning in gorillas, orangutans, and

411   humans: Unique eye-viewing patterns in humans among hominids. *Journal of*

412   *Comparative Psychology, 126*(4), 388-398.

413   Kano, F., Hirata, S., Call, J., & Tomonaga, M. (2011). The visual strategy specific to humans

414   among hominids: A study using the gap-overlap paradigm. *Vision Research, 51*(23),

415   2348-2355.

416   Kano, F., & Tomonaga, M. (2013). Head-Mounted Eye Tracking of a Chimpanzee under

417   Naturalistic Conditions. *PLoS One, 8*(3), e59785.

418   Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system.

419   *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,*

420   *352*(1358), 1231-1239.

421   Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning

422   in 5-month-old infants. *Psychological Science, 16*(8), 601-608.

423   Myowa-Yamakoshi, M., Scola, C., & Hirata, S. (2012). Humans and chimpanzees attend

424   differently to goal-directed actions. *Nature Communications, 3*, 693.

425   Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the

426   understanding and imitation of action. *Nature Reviews Neuroscience, 2*(9), 661-670.

427   Rochat, M. J., Serra, E., Fadiga, L., & Gallese, V. (2008). The evolution of social cognition: goal

428   familiarity shapes monkeys' action understanding. *Current Biology, 18*(3), 227-232.

429   Sommerville, J. A., Hildebrand, E. A., & Crane, C. C. (2008). Experience matters: the impact of

430   doing versus watching on infants' subsequent perception of tool-use events.

431   *Developmental Psychology, 44*(5), 1249.

432   Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief

433   by 2-year-olds. *Psychological Science, 18*(7), 587-592.

434   Uller, C. (2004). Disposition to recognize goals in infant chimpanzees. *Animal Cognition, 7*(3),

435   154-161.

436 Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young

437   chimpanzees. *Science, 311*(5765), 1301-1303.

438 Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition,*

439   *69*(1), 1-34.

440 Yamamoto, S., Humle, T., & Tanaka, M. (2012). Chimpanzees' flexible targeted helping based on

441   an understanding of conspecifics' goals. *Proceedings of the National Academy of*

442   *Sciences, 109*(9), 3588-3592.

443

444