

1 **How do non-human primates represent others' awareness of where objects are**
2 **hidden?**

3 Daniel J. Horschler^{1,2,*}, Laurie R. Santos³, & Evan L. MacLean^{1,2,4,5}

4 ¹School of Anthropology, University of Arizona, Tucson, AZ 85719, USA

5 ²Cognitive Science Program, University of Arizona, Tucson, AZ 85719, USA

6 ³Department of Psychology, Yale University, New Haven, CT 06520, USA

7 ⁴Department of Psychology, University of Arizona, Tucson, AZ 85719, USA

8 ⁵College of Veterinary Medicine, University of Arizona, Tucson, AZ 85719, USA

9 *Correspondence: horschler@email.arizona.edu (D. J. Horschler)

10 ORCID: 0000-0003-0440-8217

11 **Abstract**

12 Although non-human primates (NHPs) generally appear to predict how knowledgeable
13 agents use knowledge to guide their behavior, the cognitive mechanisms that enable this remain
14 poorly understood. We assessed the conditions under which NHPs' representations of an agent's
15 awareness break down. Free-ranging rhesus macaques (*Macaca mulatta*) watched as an agent
16 observed a target object being hidden in one of two boxes. While the agent could no longer see the
17 boxes, the box containing the object flipped open and the object either changed in size/shape
18 (Experiment 1) or color (Experiment 2). Monkeys looked longer when the agent searched for the
19 object incorrectly rather than correctly following the color change (a non-geometric manipulation),
20 but not the size/shape change (a geometric manipulation). Even though the agent maintained
21 knowledge of the object's location in both cases, monkeys no longer expected the agent to search
22 correctly after it had been geometrically (but not non-geometrically) manipulated. Experiment 3
23 confirmed that monkeys were sensitive to the color manipulation used in Experiment 2, making it
24 unlikely that a failure to perceive the color manipulation accounted for our findings. Our results
25 show that NHPs do not always expect that knowledgeable agents will act on their knowledge to
26 obtain their goals, consistent with heuristic-based accounts of how NHPs represent others' mental
27 states. These findings also suggest that geometric changes that occur outside the agent's perceptual
28 access may disrupt attribution of awareness more so than non-geometric changes.

29 **Keywords:** comparative cognition; non-human primates; social cognition; theory of mind;
30 knowledge representation; violation of expectation

31 1. Introduction

32 Theory of mind—the ability to represent the knowledge, ignorance, beliefs, and desires of
33 others (Premack & Woodruff, 1978)—is a critical component of human social cognition. As adult
34 humans, our theory of mind (ToM) enables us to explicitly understand that others have mental
35 states that often differ from our own. For example, by representing a person’s *desire* to obtain an
36 object and her *belief* about its location, we readily predict that she will search for the object in the
37 location where she believes it to be. This situation involves meta-representation of others’ mental
38 states—explicitly reasoning about how concepts like desires and beliefs interact to guide behavior.
39 The developmental and evolutionary origins of such capacities have been studied extensively
40 throughout the last four decades (for reviews, see Arre & Santos, n.d.; Baillargeon, Scott, & He,
41 2010; Call & Tomasello, 2008; Horschler, MacLean, & Santos, 2020; Krupenye & Call, 2019;
42 Rosati, Santos, & Hare, 2010; Saxe, 2013; Scott & Baillargeon, 2017; Slaughter, 2015; Wellman,
43 Cross, & Watson, 2001). Many comparative studies show that non-human animals behave in ways
44 that are consistent with an understanding of at least some mental states. However, whether any
45 non-human animals possess a human-like representational ToM is still debated.

46 Most comparative ToM studies have focused on non-human primates (NHPs). This work
47 has shown that NHPs are skilled at understanding whether other agents are aware of objects they
48 have seen hidden in the past, and specifically that they have correct expectations about how
49 *knowledgeable* agents will behave. For example, both monkeys (Arre, Stumph, & Santos, 2021;
50 Drayton & Santos, 2018; Horschler, Santos, & MacLean, 2019; Marticorena, Ruiz, Mukerji,
51 Goddu, & Santos, 2011) and apes (Hare, Call, & Tomasello, 2001; Kaminski, Call, & Tomasello,
52 2008; Krachun, Carpenter, Call, & Tomasello, 2009) typically expect agents who have previously
53 seen where a desired object was hidden to search for it in this location. In these cases, NHPs must

54 represent some connection between the location of an object and an agent to infer that the agent
55 will search for the object correctly, but the precise contents of these representations remain poorly
56 understood.

57 Scholars have proposed a wide range of mechanisms to account for these findings. Low-
58 level behavior-reading accounts propose that NHPs use rigid behavioral rules associating cues
59 (e.g., an agent’s eyes being visible when food was hidden) with observed outcomes (e.g., the agent
60 correctly reaches for the food) to make predictions about how agents will behave (Penn &
61 Povinelli, 2007; Povinelli & Vonk, 2003, 2012; see also the sub-mentalizing account: Heyes, 2014,
62 2017). Others have explained NHPs’ performance using the same non-mentalistic explanations for
63 successful performance in human infants (e.g., Kovacs, Teglas, & Endress, 2010; Onishi &
64 Baillargeon, 2005; Southgate, Senju, & Csibra, 2007; Woodward, 1999), such as minimal ToM
65 (Apperly & Butterfill, 2009; Butterfill & Apperly, 2013), experiential record-keeping (Perner &
66 Roessler, 2012), and teleological reasoning (Gergely & Csibra, 2003).

67 In contrast, some scholars have proposed more mentalistic hypotheses for NHPs’
68 performance. One such account—the knowledge-ignorance account—hypothesizes that NHPs
69 represent others’ knowledge and ignorance states as such¹ (Call & Santos, 2012; Call & Tomasello,
70 2008; Kaminski et al., 2008; Rosati, Santos, & Hare, 2010; Whiten, 2013). Although many
71 previous findings are consistent with the knowledge-ignorance account, some have argued that
72 this account cannot fully explain NHPs’ performance (Martin & Santos, 2016). For example, the
73 knowledge-ignorance account predicts that NHPs should understand what it means for an agent to
74 be ignorant—that the agent does not know where an object is. However, there is currently no

¹ The knowledge-ignorance account also hypothesizes that NHPs *cannot* represent others’ belief states, such as when an agent believes some information that is not congruent with reality (i.e., has a false belief).

75 experimental evidence that NHPs make positive predictions about how ignorant agents will
76 behave, such as demonstrating an expectation that an ignorant agent is likely to search for
77 information that they lack. Additionally, there is no evidence that NHPs intentionally induce
78 ignorance in others (e.g., by taking an object that an agent is aware of and hiding it in a new
79 location), despite clear evidence for acting in ways that avoid making others aware (e.g., leaving
80 information that an agent is unaware of hidden from them) (Byrne & Whiten, 1991; Hare, Call, &
81 Tomasello, 2006; Karg, Schmelz, Call, & Tomasello, 2015; A Whiten & Byrne, 1988).

82 Therefore, some scholars have argued against the knowledge-ignorance account, claiming
83 that NHPs employ simpler mentalistic representations to predict the actions of knowledgeable
84 agents. One candidate set of mechanisms—termed “awareness relations”—are thought to link true
85 information about the world to the minds of other agents (Martin & Santos, 2016). Under this
86 view, NHPs represent an agent as being “aware” of some true piece of information (e.g., “the agent
87 has awareness that an apple is in the box”). These awareness relations are proposed to have an
88 “on-off” quality, such that NHPs either represent a link between an agent and true information or
89 represent nothing at all about this relationship. As a function of this on-off quality, NHPs are
90 hypothesized to be unable to represent an agent’s state of ignorance, in contrast to the knowledge-
91 ignorance account. Importantly, awareness relations are hypothesized to break down if the
92 information linked to the agent changes while the agent is not aware of it (e.g., an object inside a
93 box moves in and out while an agent is no longer looking).

94 Numerous hypotheses about the representations that guide NHPs’ ToM have thus been
95 proposed, but how should researchers distinguish between these nuanced accounts? One strategy
96 is to explore the circumstances under which these representations break down. This approach has
97 been used effectively to elucidate the contents of different representations in other domains.

98 Consider, for example, the domain of numerical cognition. Initially, nonverbal number
99 discrimination was believed to be wholly dependent on the ratio between numbers being compared
100 (i.e., discrimination obeying Weber's law). However, subsequent work with both humans and
101 NHPs demonstrated surprising scenarios in which this prediction did not hold (Feigenson, Carey,
102 & Hauser, 2002; Hauser & Carey, 2003). Additionally, related studies showed that representations
103 of small sets of objects can break down in both humans and NHPs based on how the objects move
104 (e.g., in cases of discontinuous appearance and disappearance or dispersion and coalescence) or
105 are presented (e.g., in cases where objects are embedded within one another or are connected by
106 lines) (Huntley-Fenner, Carey, & Solimando, 2002; Mitroff, Scholl, & Wynn, 2004; Scholl, 2001;
107 Scholl & Pylyshyn, 1999; VanMarle, Aw, McCrink, & Santos, 2006). Thus, by discovering these
108 limitations, cognitive scientists were able to identify properties of the specific representations
109 underlying performance on diverse number tasks. Here, we apply a similar approach to the study
110 of mental state representation in NHPs. More specifically, we assess the conditions that cause
111 NHPs' representations of an agent's awareness of a hidden object to break down.

112 Previous work has shown that NHPs no longer expect an agent to be aware of an object
113 they saw hidden after the object was arbitrarily moved out of and back into a hiding location while
114 the agent could not see (Horschler et al., 2019; Kaminski et al., 2008). Additionally, one study
115 found that monkeys no longer expected an agent to be aware of the location of an object she saw
116 hidden in one of two differently colored boxes after the boxes rotated into different positions while
117 the agent's view was occluded (Drayton & Santos, 2018). However, monkeys did attribute
118 awareness to the agent when she saw the rotation, suggesting that monkeys maintained
119 representations of the agent's awareness only when the agent witnessed all positional
120 manipulations. Thus, current evidence strongly suggests that positional movement of an object

121 while outside of an agent’s perceptual access is sufficient to disrupt attribution of awareness to the
122 agent. However, the specific factors governing when and why these types of manipulations
123 interfere with knowledge-like attribution remain poorly understood.

124 One possible explanation is that any manipulation to a target object while outside the
125 agent’s perceptual access is sufficient to disrupt attribution of awareness. This hypothesis proposes
126 that subjects form a link between the agent and the object, and when the state of the object is
127 changed in any way while outside of the agent’s awareness, the representation linking the agent
128 and object breaks down. However, an alternative and more nuanced explanation is that in previous
129 studies, spatial movements have had this effect because the tasks specifically assessed NHPs’
130 representations of an agent’s awareness about the *location* of a target object—representations that
131 necessarily contained spatial information about the object. On this account, NHPs’ representations
132 were disrupted not simply because the object was manipulated, but rather because a specific
133 property of the object that was critical to the representation was manipulated. Therefore, this
134 account leaves open the possibility that non-spatial manipulations to an object may not have
135 similar effects, because these manipulations do not alter a property of the object that is relevant to
136 the representation.

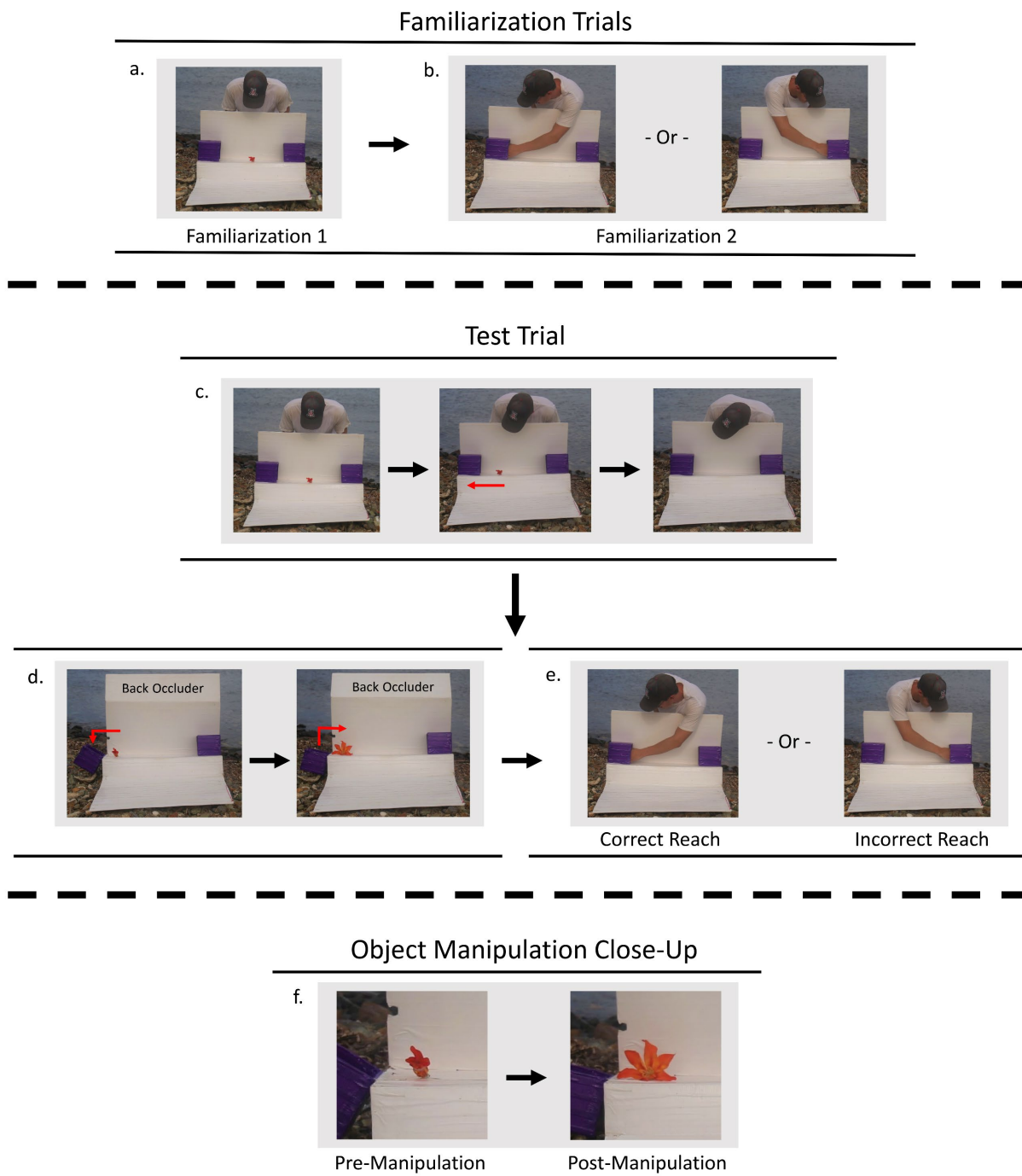
137 To better understand the specific factors governing when NHPs’ representations of others’
138 awareness break down, we assessed the impacts of a novel geometric and non-geometric change
139 to the object’s properties while it was outside the agent’s perceptual access. The geometric
140 manipulation was designed to alter spatial properties of the object (i.e., size and shape) without
141 impacting the object’s location, while the non-geometric manipulation was designed to alter a non-
142 spatial property (i.e., color). Our method was based on a commonly-used violation of expectation
143 task (Martcorena et al., 2011), in which monkeys have robustly demonstrated an understanding

144 of agents' awareness of a hidden object (see also Arre, Clark, & Santos, 2019; Arre et al., n.d.;
145 Drayton & Santos, 2018; Horschler et al., 2019; Martin & Santos, 2014). In this task, subjects
146 watch as an agent observes an object hidden in one of two boxes. Typically, monkeys look
147 significantly longer when the agent reaches into the incorrect as opposed to correct box to search
148 for it, suggesting that they expect the agent to search correctly based on his awareness of the
149 object's location. In our experiments, we similarly created situations in which an agent always
150 maintained awareness about the correct location of an object and tested whether monkeys had
151 differing expectations about where the agent would search after the object's size/shape or color
152 were manipulated while the agent was not looking.

153 **Table 1**

Experiment	Test Event	n	Age	Male	Female
1	Correct Reach	50	5.81 ± 3.85	28	22
	Incorrect Reach	49	5.52 ± 3.69	28	21
2	Correct Reach	41	4.73 ± 3.66	25	16
	Incorrect Reach	41	4.69 ± 3.51	23	18
3	Control	29	4.45 ± 2.74	16	13
	Color Change	30	4.43 ± 2.75	16	14

154 **Table 1.** Descriptive statistics of the sample in each condition of all three experiments,
155 including sample size (n), age in years (mean ± standard deviation), number of males, and
156 number of females.



158

159 **Figure 1.** A depiction of the procedure for Experiment 1. **(a)** Familiarization 1: The agent stared at the un-blossomed
 160 flower positioned in the middle of the stage for 10 seconds. **(b)** Familiarization 2: The agent reached into one of the
 161 two boxes (consistent with the final reach direction in the test trial) for 10 seconds. **(c)** Test trials: All monkeys watched
 162 as the agent observed the flower moving into one of the two boxes. **(d)** While the agent’s view of the stage was
 163 occluded, the box hiding the flower opened over the end of the stage, the flower bloomed, and the box closed. **(e)**
 164 Finally, the agent reached into either the correct or incorrect box while the monkey’s looking behavior was recorded
 165 for 10 seconds. Panel **(f)** provides a close-up view of the flower’s manipulation in panel **(d)**.

166 2. Experiment 1

167 In Experiment 1, monkeys watched as an agent saw a target object—an unblossomed
168 flower—move into one of two boxes. While the agent’s view was blocked, monkeys watched as
169 the box containing the flower flipped open and the flower bloomed (i.e., a geometric manipulation
170 that changed the object’s shape and made it appear larger). The box then flipped closed, and the
171 agent reappeared and reached into either the correct or incorrect box to search for the object. If
172 monkeys represented that the agent was aware of this object’s location after its change, we
173 expected them to look significantly longer when the agent reached into the incorrect box compared
174 to the correct box (indicating violation of expectation). However, if monkeys were no longer able
175 to represent the agent’s awareness after this change, we expected that monkeys would look equally
176 long in both conditions.

177 2.1. Methods

178 **2.1.1 Subjects:** We tested 99 free-ranging rhesus macaques at the Cayo Santiago Biological
179 Field Station (Rawlins & Kessler, 1986) based on sample sizes reported in previous looking time
180 studies in this population (Drayton & Santos, 2018; Horschler et al., 2019). This population is well
181 habituated to participation in cognitive experiments, including those using violation of expectation
182 paradigms (Arre et al., 2019; Drayton & Santos, 2018; Horschler et al., 2019; Marticorena et al.,
183 2011; Martin & Santos, 2014). Individual monkeys were identified after testing by a unique three-
184 digit tattoo. Our sample included 56 males and 43 females (mean age 5.66 ± 3.77 years, Table 1).
185 Other monkeys were approached for testing but did not contribute to our final sample because they
186 failed to watch critical components of the presentation ($n = 17$), left the presentation area ($n = 35$),
187 became distracted due to interference from other monkeys ($n = 4$), had been tested previously in
188 the same experiment ($n = 10$), or due to experimental error ($n = 2$). Decisions to abort were made

189 by the cameraperson, who was blind to all conditions. The exclusion rate was similar to or lower
190 than that of previous studies in this population (Arre et al., 2019; Drayton & Santos, 2018;
191 Horschler et al., 2019; Marticorena et al., 2011; Martin & Santos, 2014).

192 **2.1.2. Methods and Apparatus:** To assess monkeys' expectations about an agent's
193 awareness of the object's location, we used a violation of expectation looking time method. We
194 chose a flower as the target object because flowers are commonly consumed by these monkeys
195 and because it allowed for a size/shape manipulation by opening and closing the petals. Subjects
196 were approached opportunistically when relatively isolated from other group members. In each
197 trial, the experimenter knelt behind the apparatus approximately 2m in front of the subject, with a
198 cameraperson filming the subject's face while standing approximately 1m behind the
199 experimenter. As in previous studies we used a foamcore stage for our presentation (Arre et al.,
200 2019; Drayton & Santos, 2018; Marticorena et al., 2011; Martin & Santos, 2014), identical to that
201 used in Horschler et al. (2019). The stage was 76cm long \times 26cm wide, sitting 14cm off the ground
202 with a 56cm tall back panel (Figure 1). Two boxes, situated on opposite ends of the stage, measured
203 15cm \times 15cm \times 15cm. A front occluder (50cm tall) could be raised to block the subject's view of
204 the stage, and a back occluder (20cm tall) could be raised to block the experimenter's view. A
205 71cm long track was cut into the stage between the boxes along which the flower could travel. The
206 side of each box facing the center of the stage was cut out so that the flower could enter and exit
207 both boxes. Both boxes were fixed to the outside of the stage, allowing them to flip open over the
208 side of either end of the stage. Neither the experimenter nor the subject had visual access to the
209 contents of either box while the boxes were sitting on the stage. As in previous studies, the
210 experimenter controlled the movement of the target object and the boxes behind the stage

211 surreptitiously via wooden dowels attached to each object (Arre et al., 2019; Horschler et al., 2019;
212 Marticorena et al., 2011; Martin & Santos, 2014).

213 **2.1.3. Procedure:** Testing sessions consisted of two familiarization trials and one test trial
214 (Figure 1). Familiarization trials served to familiarize the subject to the apparatus and the potential
215 for the experimenter to reach into a box. In the first familiarization trial, the experimenter dropped
216 the front occluder to reveal a small, un-blossomed flower situated on the stage between the two
217 boxes. When the occluder dropped, the experimenter stared downward at the flower and said
218 “now” to begin the 10 second trial. After the first familiarization trial, the cameraperson announced
219 the condition using an alphanumeric code. The cameraperson was blind to which codes
220 corresponded to which conditions. Subjects were assigned to conditions pseudo-randomly to
221 balance mean ages and sex ratios across conditions. The condition code was announced after the
222 first familiarization trial because a subject’s condition determined which box the experimenter
223 reached into in the second familiarization trial.

224 In the second familiarization trial, the experimenter dropped the front occluder and reached
225 into one of the two boxes while the flower was no longer visible. The box reached into (left vs.
226 right) was balanced between subjects within each condition (correct or incorrect reach; described
227 below) but was always consistent with which box the experimenter ultimately reached into in the
228 test trial. The experimenter held the reaching motion and said “now” to begin the 10 second trial,
229 remaining still throughout.

230 After the second familiarization trial, the test trial began (Video S1). In the test trial, the
231 experimenter dropped the front occluder to reveal the unblossomed flower situated on the stage
232 between the two boxes, and subsequently watched the flower as it moved into one of the boxes.
233 The experimenter then raised the back occluder to block his view of the stage. While the

234 experimenter's view was occluded, the box containing the flower flipped open and the flower
235 "bloomed" (i.e., growing larger while exposing the interior of the petals) while in view of the
236 subject but not the agent. The experimenter controlled the flower's bloom surreptitiously by
237 pulling back on a 2cm transparent ring from behind the stage via attached fishing line. The box
238 then flipped closed, the experimenter dropped the back occluder such that he could again see the
239 stage, and then reached into one of the boxes (between-subjects). After reaching, the experimenter
240 said "now" and subjects' looking was filmed for 10 seconds. If monkeys are able to represent that
241 the agent was still aware of the flower's location after its change in size/shape, we expected them
242 to look significantly longer when the agent reached into the incorrect rather than the correct box.
243 However, if subjects were no longer able to represent the agent's awareness after the flower
244 changed size/shape, then monkeys should look equally long in both conditions.

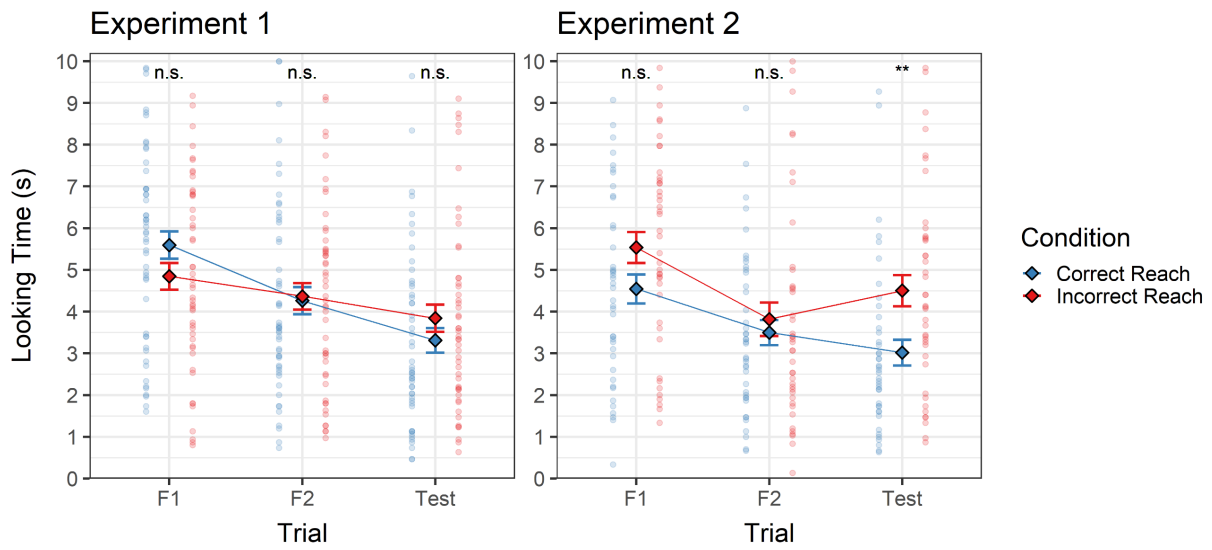
245 **2.1.4. Video Coding:** As in previous studies (Arre et al., 2019; Drayton & Santos, 2018;
246 Horschler et al., 2019; Martin & Santos, 2014), all trials were coded independently by two coders
247 using the programs MPEG Streamclip or BORIS (Friard & Gamba, 2016). Videos were clipped
248 such that coders were blind to the condition and trial. Only the subject's face (and occasionally
249 body depending on the level of zoom in the original recording) was visible in the 10-second-long
250 clips, and all subjects were centered in front of the apparatus to begin each trial. Coders assessed
251 each 10 second trial at 30 frames/second beginning immediately after the experimenter said "now"
252 by recording each frame where the subject was looking at the apparatus. Interrater reliability
253 between coders was excellent (Pearson's $R = .92$).

254 **Table 2**

Experiment	Trial	Looking Time: Incorrect Reach	Looking Time: Correct Reach	Contrast Estimate	<i>t</i>	<i>df</i>	<i>p</i>
1	Familiarization 1	4.85 ± .32	5.60 ± .38	.18	1.49	243	.36
	Familiarization 2	4.37 ± .32	4.26 ± .33	-.02	-.20	243	.99
	Test	3.84 ± .32	3.31 ± .30	-.19	-1.50	243	.36
2	Familiarization 1	5.54 ± .37	4.55 ± .35	-.24	-1.64	181	.28
	Familiarization 2	3.82 ± .40	3.50 ± .30	.004	.03	181	1.00
	Test	4.50 ± .37	3.02 ± .31	-.43	-2.99	181	.01**
Experiment	Trial	Looking Time: Color Change	Looking Time: Control	Contrast Estimate	<i>t</i>	<i>df</i>	<i>p</i>
3	Familiarization 1	4.48 ± .45	4.11 ± .52	-.20	-.83	123	.93
	Familiarization 2	2.87 ± .40	2.83 ± .38	-.02	-.01	123	1.00
	Test	4.82 ± .43	2.24 ± .35	-1.14	-4.70	123	<.001***

255 **Table 2.** Mean looking time ± standard error for each trial in all three experiments broken down by condition (i.e., monkeys
256 who ultimately saw an incorrect reach versus a correct reach in the test trial for Experiment 1 and 2; monkeys who ultimately
257 saw a green [color change] versus a yellow lemon [control] in the test trial for Experiment 3), along with test statistics for
258 all comparisons, including estimates from planned contrasts examining log-transformed looking times between conditions
259 within each trial type, *t* statistics, degrees of freedom, and *p*-values. Asterisks indicate significant differences at an alpha
260 level of .05 (*), .01 (**), or .001 (***).

261 **Figure 2**



262
263 **Figure 2.** Diamond points show mean looking time in seconds \pm standard error grouped by condition in the first
264 familiarization trial (F1), the second familiarization trial (F2), and the test trial (Test) in Experiments 1 and 2. Circular
265 points show individual-level data. Asterisks indicate significant differences at an alpha level of .05 (*), .01 (**), or
266 .001 (***), while 'n.s.' denotes no significant difference.

267 2.2. Results

268 Analyses were conducted in the R environment (v4.0.0; R Core Team, 2020), and looking
269 times were log-transformed in all analyses to improve model fit. Linear mixed models were fit
270 using the “lmer” function with a Gaussian error structure and identity link function from the
271 “lme4” package (Bates, Mächler, Bolker, & Walker, 2015). We assessed model assumptions by
272 visually examining diagnostic plots to check for normality of residuals and linear fit. Additionally,
273 we used the “var.test” function to conduct F-tests for homogeneity of variances between
274 conditions. In all cases, model assumptions were met. We assessed the effects of the predictors
275 using the “Anova” function from the “car” package (Fox & Weisberg, 2019) to produce an analysis
276 of deviance table using Type II Wald chi-squared tests. To assess significant interactions and
277 planned contrasts, we used the “emmeans” function from the “emmeans” package (Lenth, 2020)
278 to conduct t-tests between estimated marginal means of interest.

279 To assess differences in looking time between monkeys who saw the agent reach correctly
280 versus incorrectly in the test trial, we fit a linear mixed model with a random intercept for subject
281 predicting looking time as a function of trial type (Familiarization 1, Familiarization 2, or Test),
282 condition (correct or incorrect reach), and the interaction between trial type and condition. Results
283 showed a significant interaction between trial type and condition ($\chi^2(2) = 6.52, p = .04$). We next
284 examined planned contrasts comparing looking time on each of the three trial types across
285 conditions. After Dunn-Sidak correction for multiple comparisons, there was no significant
286 difference in looking time between monkeys who saw a correct reach and those who saw an
287 incorrect reach on the test trial in either Familiarization 1 (Correct: $M = 5.60s$; Incorrect: $M =$
288 $4.85s$; $t(243) = 1.49, p = .36$; Table 2; Figure 2) or Familiarization 2 (Correct: $M = 4.26s$; Incorrect:
289 $M = 4.37s$; $t(243) = -.20; p = .99$; Table 2; Figure 2). Thus, there were no general differences in

290 how interested each group of monkeys was in viewing events on the stage. In the test trial, there
291 was also no significant difference in looking time between the monkeys who saw a correct reach
292 ($n = 50$; $M = 3.31s$) versus an incorrect reach ($n = 49$; $M = 3.84s$; $t(243) = -1.50$, $p = .36$, Dunn-
293 Sidak corrected; Table 2; Figure 2). This suggests that monkeys had no expectation about the
294 agent's behavior despite his awareness of the flower's location after its size/shape changed outside
295 of the agent's perpetual access.

296 **2.3. Discussion**

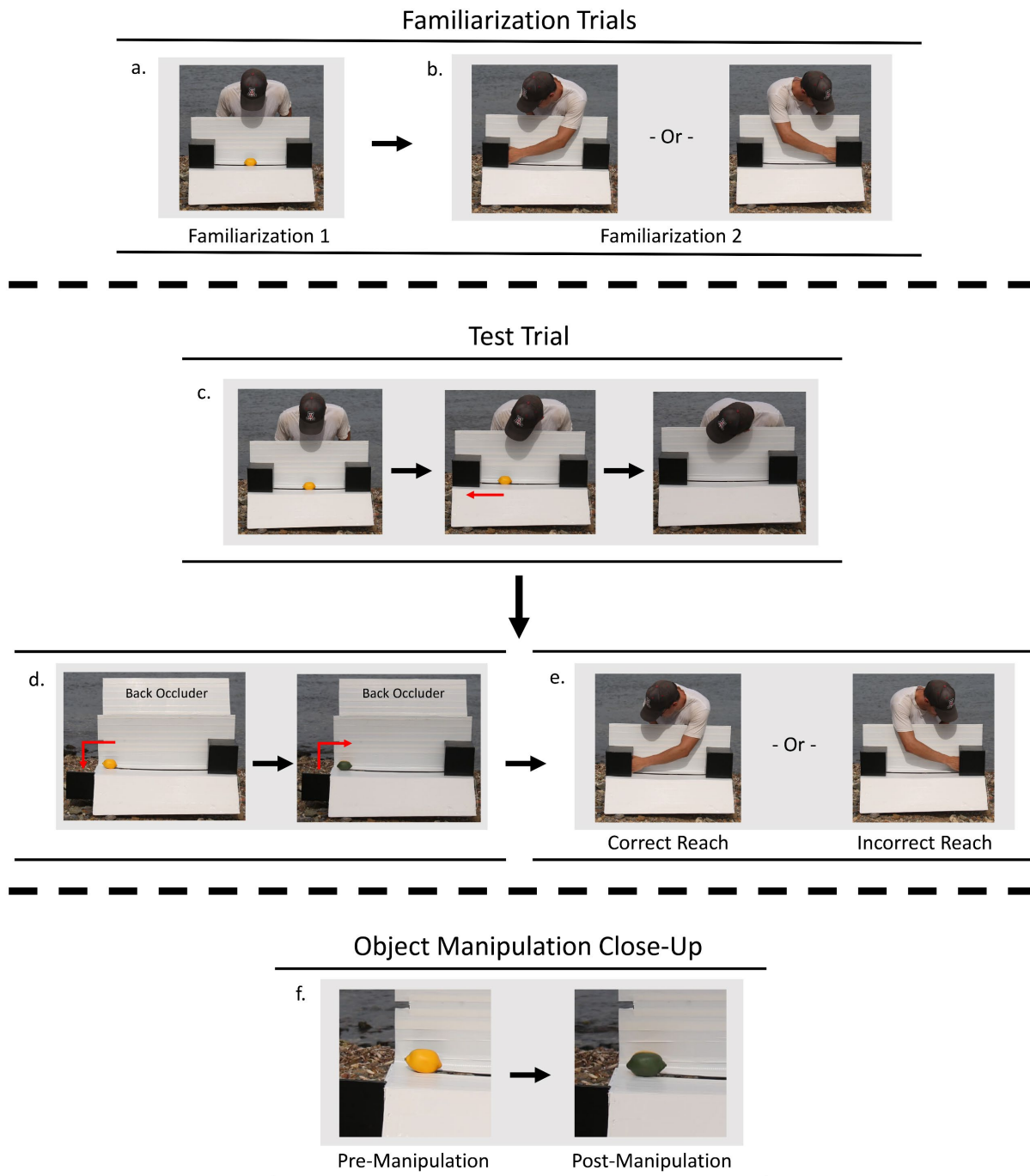
297 Even though the agent saw where the flower was hidden, we found that monkeys looked
298 equally when the agent reached into the incorrect versus the correct box after the flower changed
299 size/shape outside of the agent's perceptual access. As reviewed above, our method was based on
300 a commonly-used task in which subjects have robustly demonstrated an understanding of agents'
301 awareness of the location of objects they previously witnessed being hidden (Arre et al., 2021;
302 Drayton & Santos, 2018; Horschler et al., 2019; Marticorena et al., 2011). Therefore, in otherwise
303 identical situations where the object is not manipulated in any way, monkeys who see an agent
304 observe a target object enter a hiding location expect the agent to reach correctly for it (Arre et al.,
305 2021; Horschler et al., 2019; Marticorena et al., 2011). More specifically, in the box moves
306 condition of Horschler et al. (2019), monkeys who saw an agent observe an object enter one of
307 two boxes still anticipated that the agent would reach correctly for it after the box containing the
308 object flipped open and closed while the agent could not see it. In that condition, monkeys looked
309 significantly longer at incorrect as compared to correct reaches, suggesting that movement of the
310 box did not disrupt monkeys' expectations about the agent's awareness of the object it contained,
311 absent direct manipulation of the object. Therefore, our current results suggest that the flower's
312 size/shape change was responsible for disrupting monkeys' predictions about the agent's action.

313 We designed the current procedure such that in reality the agent always maintained
314 awareness of the flower's location—the flower never changed locations after the agent observed
315 where it was hidden. A monkey who maintained a representation of the agent's awareness should
316 have expected him to reach correctly rather than incorrectly to obtain the flower. However, we
317 found no evidence for expectations of where the agent would reach after the flower's size/shape
318 changed. While monkeys may have perceived this change as an alteration to an object's identity,
319 this possibility seems unlikely given that monkeys always had direct perceptual access to the
320 flower's transformation (see 'General Discussion'). These findings are consistent with previous
321 studies showing that spatial transformations of an object appear to disrupt representations of an
322 agent's awareness of this object's location (e.g., Horschler et al., 2019; Kaminski et al., 2008). Our
323 findings extend this work by showing that geometric manipulations which never impact an object's
324 location/position also disrupt monkeys' representations, suggesting that even more subtle spatial
325 manipulations interfere with representations about whether the agent is aware of the object's
326 location.

327 **3. Experiment 2**

328 To test the prediction that spatial manipulations *specifically* break representations of an
329 agent's awareness of an object's location, Experiment 2 introduced a non-geometric manipulation
330 of a different property of a target object (color). We hypothesized that non-geometric
331 manipulations should not break these representations if our findings from Experiment 1 are the
332 result of geometric object manipulations interfering with the spatial information contained in
333 awareness relations about the object's location. In Experiment 2, monkeys watched as an agent
334 saw a target object (a lemon) move into one of two boxes. While the agent's view was blocked,
335 the box containing the lemon flipped open and the lemon appeared to change color. The box then

336 flipped closed, and the agent reappeared and reached into either the correct or incorrect box. If any
337 manipulation (geometric or non-geometric) of a target object breaks monkeys' representations of
338 the agent's awareness, we expected that monkeys would look equally long regardless of where the
339 agent reached. However, if non-geometric manipulations (i.e., a change to the object's color in this
340 case) do not break monkeys' representations of the agent's awareness, we predicted that monkeys
341 would expect the agent to reach correctly for the object, and thus look significantly longer at
342 incorrect as compared to correct reaches.



344

345 **Figure 3.** A depiction of the procedure for Experiment 2. **(a)** Familiarization 1: The agent stared at the lemon
 346 positioned in the middle of the stage for 10 seconds. **(b)** Familiarization 2: The agent reached into one of the two
 347 boxes (consistent with the final reach direction in the test trial) for 10 seconds. **(c)** Test trials: All monkeys watched
 348 as the agent observed the lemon moving into one of the two boxes. **(d)** While the agent’s view of the stage was
 349 occluded, the box hiding the lemon opened over the end of the stage, the lemon turned green, and the box closed. **(e)**
 350 Finally, the agent reached into either the correct or incorrect box while the monkey’s looking behavior was recorded
 351 for 10 seconds. Panel **(f)** provides a close-up view of the lemon’s manipulation in panel **(d)**.

352 **3.1. Methods**

353 **3.1.1 Subjects:** We tested 82 rhesus macaques, including 48 males and 34 females with a
354 mean age of 4.71 ± 3.59 years (Table 1). Other monkeys were approached for testing but did not
355 contribute to our final subject group because they failed to watch the presentation ($n = 27$), left the
356 presentation area ($n = 22$), became distracted due to interference from other monkeys ($n = 2$), had
357 been tested previously ($n = 9$), or due to experimental error ($n = 1$). We intended to test an equal
358 number of subjects as in Experiment 1, but data collection was limited by time constraints at the
359 field site. We analyzed these data rather than collecting more data for this experiment on future
360 field trips to minimize potential differences in looking time due to seasonality, as social behavior
361 in this population has been shown to vary across the mating and birth seasons (Brent, MacLarnon,
362 Platt, & Semple, 2013).

363 **3.1.2. Methods and Apparatus:** All testing methods were identical to Experiment 1 except
364 for the apparatus: we used a foamcore stage with slightly different dimensions (Figure 3). The
365 front occluder measured 43cm and the track along which the object could travel measured 64cm.

366 **3.1.3. Procedure:** Testing sessions consisted of two familiarization trials and one test trial,
367 and the procedure was identical to Experiment 1 except for following (Figure 3; Video S2). Rather
368 than a flower, we used a lemon as the target object. We chose a lemon based on previous studies
369 using similar paradigms (Horschler et al., 2019; Marticorena et al., 2011), and because it allowed
370 for color manipulation by painting one half of it green (such that it looked like a lime). During the
371 familiarization trials and at the beginning of the test trial, only the yellow side of the lemon was
372 visible to the subject. In the test trial, the experimenter dropped the front occluder to reveal the
373 yellow lemon situated on the stage between the two boxes, and subsequently watched the lemon
374 as it moved into one of the boxes. The experimenter then raised the back occluder to block his

375 view of the stage. While the experimenter's view was occluded, the box containing the lemon
376 opened and the lemon rapidly rotated on its axis, appearing to turn green without moving. The
377 experimenter surreptitiously controlled the rotation by quickly twisting a wooden dowel attached
378 to the lemon from underneath the stage. The box then closed, the experimenter dropped the back
379 occluder such that he could again see the stage, and then reached into either the correct or incorrect
380 box (between-subjects).

381 **3.1.4. Video Coding:** All coding procedures were the same as in Experiment 1. Interrater
382 reliability between coders was excellent (Pearson's $R = .94$).

383 **3.2. Results**

384 Our analytical approach was identical to Experiment 1. Results showed a significant
385 interaction between trial type and condition ($\chi^2(2) = 7.64, p = .02$). After Dunn-Sidak correction,
386 there were no significant differences in looking time between monkeys who saw a correct reach
387 and those who saw an incorrect reach on the test trial in either Familiarization 1 (Correct: $M =$
388 $4.55s$; Incorrect: $M = 5.54s$; $t(181) = -1.64, p = .28$; Table 2; Figure 2) or Familiarization 2 (Correct:
389 $M = 3.50s$; Incorrect: $M = 3.82s$; $t(181) = .03, p = 1.00$; Table 2; Figure 2). However, in the test
390 trial, monkeys who saw an incorrect reach ($n = 41$; $M = 4.50s$) looked significantly longer than
391 monkeys who saw a correct reach ($n = 41$; $M = 3.02s$; $t(181) = -2.99, p = .01$, Dunn-Sidak
392 corrected; Table 2; Figure 2) with a large effect size ($d = -.85$). This suggests that monkeys
393 expected the agent to still be aware of the lemon's location and were thus surprised when he
394 reached incorrectly, even after it changed color while the agent was not looking.

395 **3.3. Discussion**

396 After monkeys saw an agent observe where the lemon was hidden and watched as the
397 lemon changed color outside of the agent’s perceptual access, monkeys still expected the agent to
398 reach correctly for it. This result contrasts with Experiment 1, in which monkeys did not expect
399 the agent to reach correctly for an object after it changed size/shape while the agent could not see,
400 but aligns with previous work in which monkeys expected a knowledgeable agent to search
401 correctly (Arre et al., 2021; Drayton & Santos, 2018; Horschler et al., 2019; Marticorena et al.,
402 2011). Together with Experiment 1 and results from previous work (Drayton & Santos, 2018;
403 Horschler et al., 2019; Kaminski et al., 2008), these results suggest that spatial but not non-spatial
404 changes to a target object while an agent cannot see it disrupt monkeys’ representations about the
405 agent’s awareness. We hypothesize that spatial (but not non-spatial) object manipulations have
406 this effect because they interfere with the spatial information contained in these awareness
407 relations (i.e., information about the object’s location) in a way that decouples this information
408 from representation of the agent’s awareness.

409 However, one alternative explanation for the results of Experiment 2 is that the object’s
410 color change may have been less salient than the flower’s size/shape change used in Experiment
411 1, raising the question of whether monkeys even detected the color change. Under this explanation,
412 monkeys may have maintained their representations of the agent’s awareness in Experiment 2
413 simply because they did not perceive a change to the object at all. To explore this alternative,
414 Experiment 3 directly tested whether monkeys are able to detect this change in the lemon’s color.

415 **4. Experiment 3**

416 In Experiment 3, we used a habituation-dishabituation paradigm—a method well-suited to
417 examining whether subjects can discriminate between two different stimuli—to ensure that
418 monkeys perceived the lemon’s change from yellow to green in a situation not requiring

419 representation of the agent's awareness. Monkeys were first habituated to the yellow lemon across
420 two familiarization trials, and then saw either the yellow lemon (control condition) or a green
421 lemon (color change condition) in a test trial. We predicted that if monkeys detected the color
422 change, they should exhibit dishabituation in the color change, but not the control condition.

423 **4.1. Methods**

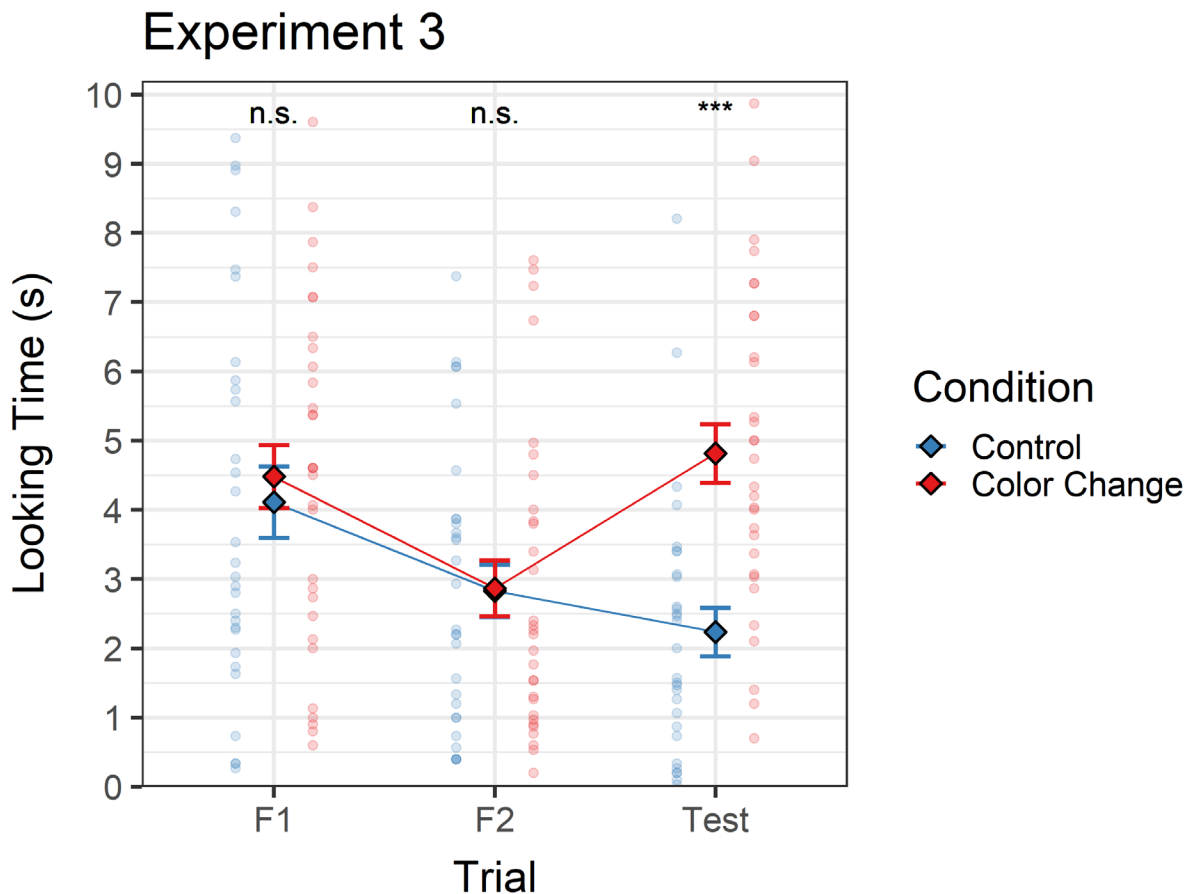
424 **4.1.1 Subjects:** We tested 59 rhesus macaques, including 32 males and 27 females (mean
425 age of 4.44 ± 2.74 years; Table 1). Other monkeys were approached for testing but did not
426 contribute to our final sample because they failed to watch the presentation ($n = 8$), left the
427 presentation area ($n = 14$), became distracted due to interference from other monkeys ($n = 2$), had
428 been tested previously ($n = 3$), or due to experimental error ($n = 2$). We intended to test a similar
429 number of subjects as in Experiments 1 and 2, but data collection ended prematurely due to the
430 imposition of travel restrictions by the University of Arizona related to the COVID-19 pandemic
431 in March 2020.

432 **4.1.2. Methods and Apparatus:** The apparatus and target object were identical to
433 Experiment 2.

434 **4.1.3. Procedure:** Sessions consisted of two familiarization trials and one test trial. The
435 familiarization trials were identical across conditions (control or color change; described below)
436 and served to familiarize subjects with the apparatus and habituate them to the yellow lemon. In
437 both familiarization trials, the experimenter dropped the front occluder to reveal an empty stage.
438 The experimenter then surreptitiously flipped open one of the boxes (balanced between subjects
439 but consistent within subjects) to reveal a yellow lemon and stared downward at the lemon for 10
440 seconds, during which the experimenter and lemon remained motionless. After the second

441 familiarization trial, the cameraperson announced the condition using an alphanumeric code. In
442 the control condition, the test trial was identical to both familiarization trials (i.e., the lemon
443 appeared as yellow) and in the ‘color change’ condition, the lemon instead appeared green. As in
444 Experiments 1 and 2, test trial condition varied between-subjects.

445 **4.1.4. Video Coding:** All coding procedures were the same as in Experiment 2. Interrater
446 reliability between coders was excellent (Pearson’s $R = .92$).



448
 449 **Figure 4.** Diamond points show mean looking time in seconds \pm standard error grouped by condition in the first
 450 familiarization trial (F1), the second familiarization trial (F2), and the test trial (Test) in Experiment 3. Circular points
 451 show individual-level data. Asterisks indicate significant differences at an alpha level of .05 (*), .01 (**), or .001
 452 (***) , while 'n.s.' denotes no significant difference.

453 4.2. Results and Discussion

454 Our analytical approach was the same as in Experiments 1 and 2, but with the addition of
455 three planned contrasts (described below). To assess differences in looking time between monkeys
456 who saw a color change versus no color change, we ran a linear mixed model with a random
457 intercept for subject predicting looking time as a function of trial type (Familiarization 1,
458 Familiarization 2, or Test), condition (control or color change), and the interaction between trial
459 type and condition. Results showed a significant interaction between trial type and condition ($\chi^2(2)$
460 = 21.81, $p < .001$). After Dunn-Sidak correction, there was no significant difference in looking
461 time between the monkeys who saw a color change and those who saw no color change on the test
462 trial in either Familiarization 1 (Control: $M = 4.11$ s; Color Change: $M = 4.48$ s; $t(123) = -.83$, $p =$
463 $.93$; Table 2; Figure 4) or Familiarization 2 (Control: $M = 2.83$ s; Color Change: $M = 2.87$ s; $t(123)$
464 = $-.01$, $p = 1.00$; Table 2; Figure 4). However, in the test trial, monkeys who saw the lemon as
465 green ($n = 30$; $M = 4.82$ s) looked significantly longer than monkeys who saw the lemon as yellow
466 ($n = 29$; $M = 2.24$ s; $t(123) = -4.70$, $p < .001$, Dunn-Sidak corrected; Table 2; Figure 4) with a large
467 effect size ($d = -1.63$). This suggests that monkeys expected to again see a yellow lemon in the test
468 trial (as in the familiarization trials) and thus looked longer (i.e., were surprised) when the lemon
469 was instead green.

470 Of the three additional planned contrasts (subject to Dunn-Sidak correction where
471 appropriate), two tested for significant differences in looking time between the second
472 familiarization trial and the test trial within each condition (dishabituation), and the third was an
473 interaction contrast to test for a significant difference in the extent of looking time change from
474 the second familiarization trial to the test trial across conditions. Monkeys who saw the lemon as
475 yellow in the test trial (control condition) showed no significant difference in looking time between

476 the second familiarization trial ($M = 2.83s$) and the test trial ($M = 2.24s$; $t(114) = 2.29$, $p = .11$).
477 However, monkeys who saw the green lemon in the test trial looked significantly longer in the test
478 trial ($M = 4.82s$) as compared to the second familiarization trial ($M = 2.87s$; $t(114) = -3.87$, $p <$
479 $.001$) with a large effect size ($d = -1.00$). Additionally, an interaction contrast showed that the
480 degree of looking time change from the second familiarization trial to the test trial was significantly
481 greater in the color change (1.95s mean increase) compared to the control condition (0.59s mean
482 decrease; $t(114) = 4.35$, $p < .001$), with a large effect size ($d = 1.60$). Taken together, these results
483 suggest that monkeys attended to and perceived the color change, showing dishabituation in the
484 test trial of the color change but not the control condition. Furthermore, this supports the
485 conclusion that in Experiment 2, monkeys maintained representations of the agent's awareness of
486 the hidden object's location despite likely perceiving its color change.

487 **5. General Discussion**

488 Our experiments aimed to assess the content of the representations that underlie NHPs'
489 understanding of how knowledgeable agents use knowledge to guide their behavior. We adapted
490 a commonly-used task in which monkeys have robustly demonstrated an understanding of agents'
491 awareness of the location of a hidden object and manipulated the object's size/shape (Experiment
492 1) or color (Experiment 2) while the agent was not looking to assess whether these manipulations
493 interfered with subjects' attributions of awareness to the agent. Importantly, the agent maintained
494 awareness of the object's location in both experiments because he witnessed the object move into
495 its hiding location and it never changed locations while he could not see it. Nevertheless, monkeys
496 did not expect the agent to maintain awareness when the object changed size/shape (Experiment
497 1) but did expect the agent to maintain awareness after the object changed colors (Experiment 2).
498 Experiment 3 suggested that monkeys are able to notice the object's color change, making it

499 unlikely that their performance in Experiment 2 resulted from a failure to perceive the object
500 manipulation. Taken together, these results suggest that, at least in cases of awareness of an
501 object's location, monkeys' representations of other agents' awareness are disrupted specifically
502 by spatial manipulations of the object.

503 Our results fit with previous work showing that NHPs do not *always* make correct
504 predictions about agents who are aware of an object's location; the failure we observed in
505 Experiment 1 is similar to the results that Kaminski and colleagues (2008, unknown lift condition)
506 observed in apes and that Horschler and colleagues (2019, fruit moves condition) observed on a
507 related task in this same macaque population. In both studies, NHPs no longer expected an agent
508 to be aware of an object that they saw hidden after it was moved out of and back into the same
509 location while the agent could not see. Relatedly, Drayton and Santos (2018) found that monkeys
510 no longer expected an agent to be aware of the location of an object she saw hidden in one of two
511 differently colored boxes after the boxes switched positions while the agent could not see. In that
512 study, monkeys did expect the agent to maintain awareness of the object's location when she
513 witnessed the boxes' positional movement, further suggesting that positional movements have this
514 effect only when they happen outside the agent's perceptual access.

515 In the current study, we implemented a much more subtle spatial manipulation (a change
516 in size/shape) which did not involve a change of location. However, our results suggest that even
517 simple geometric transformations of a stationary object appear to disrupt representations of an
518 agent's awareness, similarly to larger scale positional movements—ones more likely to disrupt
519 what an agent knew. Importantly, the results of Experiment 2 show that not all manipulations to
520 an object's properties are sufficient to interfere with monkeys' representations of awareness; we
521 found that non-geometric manipulations (i.e., a color change) do not have this same effect. We

522 hypothesize that spatial but not non-spatial object manipulations have this effect because we
523 specifically tested monkeys' representations of an agent's awareness of an object's location, which
524 must contain spatial information about the object. Thus, spatial manipulations may result in
525 monkeys updating the spatial content of these representations in a way that decouples
526 representation of the agent's awareness from representation of the object.

527 Importantly, this possibility likely also depends on the specific ways that monkeys
528 represent objects, which remain largely unexplored. In the domain of intuitive physics, analogies
529 to machine physics engines are producing compelling parallels to human mental representation of
530 spatial object manipulations (Ullman, Spelke, Battaglia, & Tenenbaum, 2017) which align well
531 with our results in monkeys. For example, physics engines often use approximate body
532 representations such as “bounding boxes”—which encompass objects but contain no specific
533 information about the object's other properties—to represent rough shape approximations
534 separately from detailed object graphics. These approximate body representations are typically
535 called on when assessing how an object may behave in the future, allowing for rapid simulations.
536 If monkeys use something similar to a bounding box as a placeholder when representing a target
537 object in awareness relations, it is reasonable that object changes altering the state of the bounding
538 box (i.e., positional movement or size/shape change like we used in Experiment 1) may cause the
539 bounding box to be updated whereas object changes irrelevant to the bounding box (i.e., color
540 change) would not. Thus, future work probing these social cognitive mechanisms will benefit from
541 a deeper understanding of how monkeys represent the relevant physical stimuli, including the
542 agent and objects being monitored.

543 Our findings extend previous research demonstrating that NHPs often understand how
544 knowledgeable agents will behave by outlining features of the cognitive mechanisms that enable

545 these behavioral predictions. Our results cannot be explained by the knowledge-ignorance account
546 of NHP ToM because this account proposes that NHPs represent others' knowledge and ignorance
547 states as such. In both Experiments 1 and 2, the agent always maintained knowledge of the object's
548 location. After the object's geometric properties were manipulated while the agent could not see,
549 monkeys had no expectations about where the agent would search. In contrast, the knowledge-
550 ignorance account predicts that monkeys should have expected the agent to search correctly even
551 after this geometric manipulation, because the agent's information about the object's location
552 never changed. By showing that NHPs do not always make predictions consistent with the agent's
553 knowledge, our results support the idea that simpler, heuristic-based awareness relations may offer
554 a better explanation for how NHPs form and maintain expectations about the actions of
555 knowledgeable agents. In the awareness relations account, NHPs are hypothesized to represent an
556 agent as either being aware of true information, or to represent nothing at all about the relationship
557 between this information and the agent. This account explains why monkeys did not expect the
558 agent to search correctly in Experiment 1, because the spatial manipulation to the object while it
559 was outside of the agent's perceptual access is hypothesized to disrupt the representation of the
560 agent's awareness of the object's location (see Horschler et al., 2019; Kaminski et al., 2008). While
561 it would be possible to amend the knowledge-ignorance account such that geometric manipulations
562 disrupt knowledge attribution, this phenomenon is more consistent with heuristic-based models,
563 which are built around simple rules governing when links between an agent and information are
564 formed or eliminated (see Horschler et al., 2019; Martin & Santos, 2016).

565 In our study, monkeys expected the agent to still be aware of an object's location after it
566 changed colors outside of his perceptual access, but did not expect the agent to still be aware of an
567 object's location after it changed sizes/shapes. One alternative interpretation of these results is that

568 monkeys perceived the flower's size/shape but not the lemon's color change as an alteration to an
569 object's *identity* (i.e., they may have represented the altered flower as a different object altogether,
570 while representing the altered lemon as having the same identity despite a change to one of its
571 properties) or simply as a greater or more consequential transformation. Previous work with
572 humans suggests that infants as young as 14 months of age reason about the identity of an object
573 as it relates to an agent's beliefs (F. Buttelmann, Suhrke, & Buttelmann, 2015; Scott & Baillargeon,
574 2009; Scott, Richman, & Baillargeon, 2015; Song & Baillargeon, 2008). Less work has explored
575 representations of object identity in NHPs, but one study (Krachun, Carpenter, Call, & Tomasello,
576 2010) found hints that chimpanzees may be able to reason about other agents' representations of
577 object identity in some contexts. In this study, chimpanzees first learned that an agent would
578 reliably hide a box containing one of two different types of food (i.e., a banana slice or a grape) in
579 one of two different locations based on which type of food was inside. When one type of food was
580 swapped for the other, five chimpanzees showed a trend toward understanding an agent's
581 knowledge of the identity change when the agent observed the food being swapped (i.e., they
582 correctly responded on 80% of trials, but this result was not statistically significant) but not their
583 false beliefs about object identity when the food was swapped while the agent was absent (i.e.,
584 they responded correctly on less than 50% of trials). In Experiments 1 and 2 of the present study,
585 monkeys witnessed each object's transformation such that they directly observed that the object
586 was not substituted for another. Therefore, our study differs from previous work using object
587 identity manipulations in that it involved the transformation of a single object rather than the
588 substitution of one object for another. It therefore seems unlikely that monkeys represented the
589 altered versions of the objects as different objects altogether given that monkeys always witnessed
590 each object's transformation. However, given how little is known about monkeys' representations

591 of objects, we cannot rule out the possibility that monkeys perceived the transformation in
592 Experiment 1 as being more consequential than that of Experiment 2.

593 A second limitation of our study is that the manipulation in Experiment 2 was
594 accomplished by an extremely rapid spatial change to the lemon (quickly rotating on its axis). To
595 a human observer, this manipulation appears as a non-spatial change to the color of the object, but
596 whether monkeys were able to detect the rapid rotation inducing this effect remains unknown. If
597 this rotation was perceived as a spatial manipulation, it remains possible that some minimal spatial
598 changes can occur (e.g., rotation in place, but not displacement or expansion into new spatial
599 coordinates) while leaving monkeys' representations of an agent's awareness of the object's
600 location intact.

601 Although more work is needed to assess monkeys' perceptions of object identity and how
602 other manipulations may influence representations of awareness, our results suggest that these
603 representations share some similarities with theoretical accounts of implicit ToM in humans. Some
604 scholars have proposed that adult humans employ two separate and dissociable ToM systems: a
605 slower, effortful, and flexible explicit system alongside a more rapid, automatic, and efficient
606 implicit system (Apperly & Butterfill, 2009; Butterfill & Apperly, 2013; Carruthers, 2017;
607 Schneider, Slaughter, & Dux, 2017). The explicit system is proposed to allow for conscious
608 representation of an agent's mental state, whereas the implicit system is used for rapid,
609 unconscious predictions about how an agent's mental state may influence their behavior
610 (Schneider et al., 2017). Apperly and Butterfill proposed that an implicit "minimal ToM" system
611 could operate using representations of simple relational states in a similar (but not identical) way
612 to awareness relations (Butterfill & Apperly, 2013). Minimal ToM predicts that subjects "register"
613 objects at the last location they were "encountered," and that these registrations are maintained

614 regardless of any spatial manipulation of the object while the agent cannot see it. Therefore, under
615 this account, subjects should make correct behavioral predictions about agents with false beliefs,
616 at least in change-of-location paradigms. Importantly, the awareness relations account predicts that
617 subjects will have no expectations about an agent's behavior in these change-of-location false
618 belief tests: the spatial manipulations requisite of inducing a false belief in an agent are predicted
619 to also break the subject's representation of the agent's awareness.

620 Our study was predicated on the well-established finding that NHPs often (but not always)
621 correctly anticipate how knowledgeable agents will behave (Arre et al., 2021; Drayton & Santos,
622 2018; Hare et al., 2001; Horschler et al., 2019; Kaminski et al., 2008; Krachun et al., 2009, 2010;
623 MacLean & Hare, 2012; Marticorena et al., 2011). A separate but related question concerns
624 whether NHPs can also correctly anticipate how agents with false beliefs will behave. Although
625 the extent to which NHPs represent false beliefs remains controversial (see Horschler et al., 2020b;
626 Horschler, MacLean, & Santos, 2020a; Kano, Call, & Krupenye, 2020), some recent studies have
627 argued for belief representation in NHPs (D. Buttelmann, Buttelmann, Carpenter, Call, &
628 Tomasello, 2017; Hayashi et al., 2020; Kano, Krupenye, Hirata, Tomonaga, & Call, 2019;
629 Krupenye, Kano, Hirata, Call, & Tomasello, 2016). We suggest that future work on belief
630 representation in NHPs may benefit from using the present study's approach to outlining the
631 specific circumstances in which positive performance on ToM tasks begins to break down.
632 Exploring novel manipulations that disrupt performance indicative of false belief representation is
633 likely to provide a better understanding of the representational content underlying this
634 performance.

635 Although we find support for the idea that geometric but not non-geometric object
636 manipulations disrupt monkeys' representations linking locational content about an object to an

637 agent, there remains much to learn about how these representations are formed and maintained.
638 For example, would a change to the identity of an object (e.g., F. Buttelmann et al., 2015; Krachun
639 et al., 2010; Scott & Baillargeon, 2009; Scott et al., 2015; Song & Baillargeon, 2008) disrupt
640 NHPs' representations of an agent's awareness about the initial object? Would awareness relations
641 that do not contain locational content about an object similarly be disrupted by geometric
642 manipulations? Or do geometric manipulations disrupt these representations simply because
643 current tasks have assessed subjects' expectations about the spatial location of a target object? By
644 addressing these questions, we will not only learn more about how NHPs reason about others'
645 intentional actions, but we will also be able to formulate testable hypotheses about a potentially
646 evolutionarily conserved implicit ToM system in humans.

647 **Acknowledgments**

648 We would like to thank Alyssa Arre, Yiyun Huang, Amelia Linett, Miriam Ross, Rebecca
649 Sellati, Ellen Stumph, and Katherine Ziska for their help in apparatus construction, data collection,
650 and coding. We are also grateful to Nahiri Rivera Barreto, Giselle Caraballo, Angelina Ruiz
651 Lambides, and Bianca Giura Negru for their help in securing the Cayo Santiago Biological Field
652 Station.

653 **Data Accessibility**

654 Data are available as electronic supplementary material.

655 **Funding**

656 This research was supported in part by grants from the NIMH (R01MH096875) and the
657 NCRR (CM-5-P40RR003640-13) to the Caribbean Primate Research Center. DJH was supported
658 by a Haury Dissertation Fellowship from the School of Anthropology at the University of Arizona.
659 LRS was supported by Yale University.

660 **Conflict of Interest**

661 The authors declare no conflict of interest.

662 **Ethics Approval**

663 This work was approved by the Cayo Santiago IACUC committee and conforms to
664 guidelines for the use of animals in research.

665 **References**

- 666 Apperly, I. A., & Butterfill, S. A. (2009). Do Humans Have Two Systems to Track Beliefs and
667 Belief-Like States? *Psychological Review*, *116*(4), 953–970.
668 <https://doi.org/10.1037/a0016923>
- 669 Arre, A. M., Clark, C. S., & Santos, L. R. (2019). Do young rhesus macaques know what others
670 see?: A comparative developmental perspective. *American Journal of Primatology*,
671 (March), 1–9. <https://doi.org/10.1002/ajp.23054>
- 672 Arre, A. M., & Santos, L. R. (n.d.). Mentalizing in Non-Human Primates. In *Neural Basis of*
673 *Mentalizing*.
- 674 Arre, A. M., Stumph, E., & Santos, L. R. (2021). Macaque species with varying social tolerance
675 show no differences in understanding what other agents perceive. *Animal Cognition*.
- 676 Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in*
677 *Cognitive Sciences*, *14*(3), 110–118. <https://doi.org/10.1016/j.tics.2009.12.006>
- 678 Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects
679 models using lme4. *Journal of Statistical Software*, *67*(1).
680 <https://doi.org/10.18637/jss.v067.i01>
- 681 Brent, L. J. N., MacLarnon, A., Platt, M. L., & Semple, S. (2013). Seasonal changes in the
682 structure of rhesus macaque social networks. *Behavioral Ecology and Sociobiology*, *67*(3),
683 349–359. <https://doi.org/10.1007/s00265-012-1455-8>.Seasonal
- 684 Buttelmann, D., Buttelmann, F., Carpenter, M., Call, J., & Tomasello, M. (2017). Great apes
685 distinguish true from false beliefs in an interactive helping task. *PLoS ONE*, *12*(4), 1–13.
686 <https://doi.org/10.1371/journal.pone.0173793>
- 687 Buttelmann, F., Suhrke, J., & Buttelmann, D. (2015). What you get is what you believe:
688 Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task.
689 *Journal of Experimental Child Psychology*, *131*, 94–103.
690 <https://doi.org/10.1016/j.jecp.2014.11.009>
- 691 Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind and*

692 *Language*, 28(5), 606–637. <https://doi.org/10.1111/mila.12036>

693 Byrne, R. W., & Whiten, A. (1991). Computation and mindreading in primate tactical deception.
694 In A Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of*
695 *everyday mindreading* (pp. 127–141). Cambridge, MA.

696 Call, J., & Santos, L. R. (2012). Understanding other minds. In *The Evolution of Primate*
697 *Societies* (pp. 664–681).

698 Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later.
699 *Trends in Cognitive Sciences*, 12(5), 187–192. <https://doi.org/10.1016/j.tics.2008.02.010>

700 Carruthers, P. (2017). Mindreading in adults: evaluating two-systems views. *Synthese*, 194(3),
701 673–688. <https://doi.org/10.1007/s11229-015-0792-3>

702 Drayton, L. A., & Santos, L. R. (2018). What do monkeys know about others' knowledge?
703 *Cognition*, 170(October 2017), 201–208. <https://doi.org/10.1016/j.cognition.2017.10.004>

704 Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of
705 more: Object files versus analog magnitudes. *Psychological Science*, 13(2), 150–156.
706 <https://doi.org/10.1111/1467-9280.00427>

707 Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third edit). Thousands
708 Oaks, CA: Sage.

709 Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software
710 for video/audio coding and live observations. *Methods in Ecology and Evolution*, 7(11),
711 1325–1330. <https://doi.org/10.1111/2041-210X.12584>

712 Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational
713 action. *Trends in Cognitive Sciences*, 7(7), 287–292. [https://doi.org/10.1016/S1364-](https://doi.org/10.1016/S1364-6613(03)00128-1)
714 6613(03)00128-1

715 Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know?
716 *Animal Behaviour*, 61(1), 139–151. <https://doi.org/10.1006/anbe.2000.1518>

717 Hare, B., Call, J., & Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding.
718 *Cognition*, 101(3), 495–514. <https://doi.org/10.1016/j.cognition.2005.01.011>

- 719 Hauser, M. D., & Carey, S. (2003). *Spontaneous representations of small numbers of objects by*
720 *rhesus macaques: Examinations of content and format. Cognitive Psychology* (Vol. 47).
721 [https://doi.org/10.1016/S0010-0285\(03\)00050-1](https://doi.org/10.1016/S0010-0285(03)00050-1)
- 722 Hayashi, T., Akikawa, R., Kawasaki, K., Egawa, J., Minamimoto, T., Kobayashi, K., ...
723 Hasegawa, I. (2020). Macaques Exhibit Implicit Gaze Bias Anticipating Others' False-
724 Belief-Driven Actions via Medial Prefrontal Cortex. *Cell Reports*, 30(13), 4433-4444.e5.
725 <https://doi.org/10.1016/j.celrep.2020.03.013>
- 726 Heyes, C. (2014). Submentalizing: I Am Not Really Reading Your Mind. *Perspectives on*
727 *Psychological Science*, 9(2), 131–143. <https://doi.org/10.1177/1745691613518076>
- 728 Heyes, C. (2017). Apes Submentalise. *Trends in Cognitive Sciences*, 21(1), 1–2.
729 <https://doi.org/10.1016/j.tics.2016.11.006>
- 730 Horschler, D. J., MacLean, E. L., & Santos, L. R. (2020a). Advancing gaze-based research on
731 primate theory of mind. *Trends in Cognitive Sciences*, 24(10), 778–779.
- 732 Horschler, D. J., MacLean, E. L., & Santos, L. R. (2020b). Do non-human primates really
733 represent others' beliefs? *Trends in Cognitive Sciences*, 24(8), 594–605.
- 734 Horschler, D. J., Santos, L. R., & MacLean, E. L. (2019). Do non-human primates really
735 represent others' ignorance? A test of the awareness relations hypothesis. *Cognition*,
736 190(March), 72–80. <https://doi.org/10.1016/j.cognition.2019.04.012>
- 737 Huntley-Fenner, G., Carey, S., & Solimando, A. (2002). Objects are individuals but stuff doesn't
738 count: Perceived rigidity and cohesiveness influence infants' representations of small
739 groups of discrete entities. *Cognition*, 85(3), 203–221. [https://doi.org/10.1016/S0010-](https://doi.org/10.1016/S0010-0277(02)00088-4)
740 [0277\(02\)00088-4](https://doi.org/10.1016/S0010-0277(02)00088-4)
- 741 Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not
742 what they believe. *Cognition*, 109(2), 224–234.
743 <https://doi.org/10.1016/j.cognition.2008.08.010>
- 744 Kano, F., Call, J., & Krupenye, C. (2020). Primates pass dynamically social false-belief tests.
745 *Trends in Cognitive Sciences*, 1–2. <https://doi.org/10.1016/j.tics.2020.07.003>

746 Kano, F., Krupenye, C., Hirata, S., Tomonaga, M., & Call, J. (2019). Great apes use self-
747 experience to anticipate an agent's action in a false-belief test. *Proceedings of the National*
748 *Academy of Sciences*, *116*(42), 20904–20909. <https://doi.org/10.1073/pnas.1910095116>

749 Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2015). Chimpanzees strategically manipulate
750 what others can see. *Animal Cognition*, *18*(5), 1069–1076. [https://doi.org/10.1007/s10071-](https://doi.org/10.1007/s10071-015-0875-z)
751 [015-0875-z](https://doi.org/10.1007/s10071-015-0875-z)

752 Kovacs, A. M., Teglas, E., & Endress, A. D. (2010). The Social Sense: Susceptibility to Others'
753 Beliefs in Human Infants and Adults. *Science*, *330*, 1830–1834.
754 <https://doi.org/10.1126/science.1190792>

755 Krachun, C., Carpenter, M., Call, J., & Tomasello, M. (2009). A competitive nonverbal false
756 belief task for children and apes. *Developmental Science*, *12*(4), 521–535.
757 <https://doi.org/10.1111/j.1467-7687.2008.00793.x>

758 Krachun, C., Carpenter, M., Call, J., & Tomasello, M. (2010). A New Change-of-Contents False
759 Belief Test: Children and Chimpanzees Compared. *International Journal of Comparative*
760 *Psychology*, *23*, 145–165.

761 Krupenye, C., & Call, J. (2019). Theory of mind in animals: Current and future directions. *Wiley*
762 *Interdisciplinary Reviews: Cognitive Science*, (June 2018), e1503.
763 <https://doi.org/10.1002/wcs.1503>

764 Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that
765 other individuals will act according to false beliefs. *Science*, *354*(6308), 110–113.

766 Lenth, R. (2020). emmeans: Estimated Marginal Means, aka Least-Squares Means.

767 MacLean, E. L., & Hare, B. (2012). Bonobos and chimpanzees infer the target of another's
768 attention. *Animal Behaviour*, *83*(2), 345–353.
769 <https://doi.org/10.1016/j.anbehav.2011.10.026>

770 Marticorena, D. C. W., Ruiz, A. M., Mukerji, C., Goddu, A., & Santos, L. R. (2011). Monkeys
771 represent others' knowledge but not their beliefs. *Developmental Science*, *4*(164), 1406–
772 1416. <https://doi.org/10.1126/scisignal.2001449.Engineering>

- 773 Martin, A., & Santos, L. R. (2014). The origins of belief representation: Monkeys fail to
774 automatically represent others' beliefs. *Cognition*, *130*(3), 300–308.
775 <https://doi.org/10.1016/j.cognition.2013.11.016>
- 776 Martin, A., & Santos, L. R. (2016). What Cognitive Representations Support Primate Theory of
777 Mind? *Trends in Cognitive Sciences*. Elsevier Ltd.
778 <https://doi.org/10.1016/j.tics.2016.03.005>
- 779 Mitroff, S. R., Scholl, B. J., & Wynn, K. (2004). Divide and Conquer: How Object Files Adapt
780 When a Persisting Object Splits Into Two. *Psychological Science*, *15*(6), 420–425.
781 <https://doi.org/10.1038/s41579-018-0146-5>
- 782 Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs?
783 *Science*, *308*(5719), 255–258. <https://doi.org/10.1126/science.1107621>
- 784 Penn, D. C., & Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess
785 anything remotely resembling a “theory of mind.” *Philosophical Transactions of the Royal*
786 *Society of London. Series B, Biological Sciences*, *362*(1480), 731–744.
787 <https://doi.org/10.1098/rstb.2006.2023>
- 788 Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in*
789 *Cognitive Sciences*, *16*(10), 519–525. <https://doi.org/10.1016/j.tics.2012.08.004>
- 790 Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in*
791 *Cognitive Sciences*, *7*(4), 157–160. [https://doi.org/10.1016/S1364-6613\(03\)00053-6](https://doi.org/10.1016/S1364-6613(03)00053-6)
- 792 Povinelli, D. J., & Vonk, J. (2012). We don't need a microscope to explore the chimpanzee's
793 mind. In *Mind & Language* (Vol. 19, pp. 1–28).
794 <https://doi.org/10.1093/acprof:oso/9780198528272.003.0018>
- 795 Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral*
796 *and Brain Sciences*, *4*(1978), 515–526.
- 797 R Core Team. (2020). A Language and Environment for Statistical Computing. *Vienna, Austria:*
798 *R Foundation for Statistical Computing*. <https://doi.org/citeulike-article-id:2400517>
- 799 Rawlins, R. G., & Kessler, M. J. (1986). *The Cayo Santiago macaques: History, behavior, and*

800 *biology*. SUNY Press.

801 Rosati, A. G., Santos, L. R., & Hare, B. (2010). Primate social cognition: Thirty years after
802 Premack and Woodruff. In *Primate Neuroethology* (pp. 117–144).

803 Saxe, R. (2013). The New Puzzle of Theory of Mind Development. *Navigating the Social World*,
804 107–112. <https://doi.org/10.1093/acprof:oso/9780199890712.003.0020>

805 Schneider, D., Slaughter, V. P., & Dux, P. E. (2017). Current evidence for automatic Theory of
806 Mind processing in adults. *Cognition*, *162*, 27–31.
807 <https://doi.org/10.1016/j.cognition.2017.01.018>

808 Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1–2), 1–46.
809 [https://doi.org/10.1016/S0010-0277\(00\)00152-9](https://doi.org/10.1016/S0010-0277(00)00152-9)

810 Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking Multiple Items Through Occlusion: Clues to
811 Visual Objecthood. *Cognitive Psychology*, *38*(2), 259–290.
812 <https://doi.org/10.1006/cogp.1998.0698>

813 Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? attributing false beliefs about
814 object identity at 18 months. *Child Development*, *80*(4), 1172–1196.
815 <https://doi.org/10.1111/j.1467-8624.2009.01324.x>

816 Scott, R. M., & Baillargeon, R. (2017). Early False-Belief Understanding. *Trends in Cognitive*
817 *Sciences*, *21*(4), 237–249. <https://doi.org/10.1016/j.tics.2017.01.012>

818 Scott, R. M., Richman, J. C., & Baillargeon, R. (2015). Infants understand deceptive intentions
819 to implant false beliefs about identity: New evidence for early mentalistic reasoning.
820 *Cognitive Psychology*, *82*, 32–56. <https://doi.org/10.1016/j.cogpsych.2015.08.003>

821 Slaughter, V. (2015). Theory of Mind in Infants and Young Children: A Review. *Australian*
822 *Psychologist*, *50*(3), 169–172. <https://doi.org/10.1111/ap.12080>

823 Song, H. joo, & Baillargeon, R. (2008). Infants' Reasoning About Others' False Perceptions.
824 *Developmental Psychology*, *44*(6), 1789–1795. <https://doi.org/10.1037/a0013774>

825 Southgate, V., Senju, A., & Csibra, G. (2007). Action Anticipation Through Attribution of False
826 Belief by 2-Year-Olds. *Psychological Science*, *18*(7), 587–592.

- 827 Ullman, T. D., Spelke, E., Battaglia, P., & Tenenbaum, J. B. (2017). Mind Games: Game
828 Engines as an Architecture for Intuitive Physics. *Trends in Cognitive Sciences*, 21(9), 649–
829 665. <https://doi.org/10.1016/j.tics.2017.05.012>
- 830 VanMarle, K., Aw, J., McCrink, K., & Santos, L. R. (2006). How capuchin monkeys (*Cebus*
831 *apella*) quantify objects and substances. *Journal of Comparative Psychology*, 120(4), 416–
832 426. <https://doi.org/10.1037/0735-7036.120.4.416>
- 833 Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of theory-of mind development:
834 The truth about false belief. *Child Development*, 72(3), 655–684.
835 [https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5)
- 836 Whiten, A., & Byrne, R. W. (1988). Tactical deception in primates. *Behavioral and Brain*
837 *Sciences*, 11, 233–273.
- 838 Whiten, Andrew. (2013). Humans are not alone in computing how others see the world. *Animal*
839 *Behaviour*, 86(2), 213–221. <https://doi.org/10.1016/j.anbehav.2013.04.021>
- 840 Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach.
841 *Cognition*, 69(1), 1–34. [https://doi.org/10.1016/S0010-0277\(98\)00058-4](https://doi.org/10.1016/S0010-0277(98)00058-4)
- 842