Cognition 212 (2021) 104658

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

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

Daniel J. Horschler [a,b,*], Laurie R. Santos [c], Evan L. MacLean [a,b,d,e]

[a] School of Anthropology, University of Arizona, Tucson, AZ 85719, USA
[b] Cognitive Science Program, University of Arizona, Tucson, AZ 85719, USA
[c] Department of Psychology, Yale University, New Haven, CT 06520, USA
[d] Department of Psychology, University of Arizona, Tucson, AZ 85719, USA
[e] College of Veterinary Medicine, University of Arizona, Tucson, AZ 85719, USA

ARTICLE INFO

Keywords:
Comparative cognition
Non-human primates
Social cognition
Theory of mind
Knowledge representation
Violation of expectation

ABSTRACT

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

1. Introduction

Theory of mind—the ability to represent the knowledge, ignorance, beliefs, and desires of others (Premack & Woodruff, 1978)—is a critical component of human social cognition. As adult humans, our theory of mind (ToM) enables us to explicitly understand that others have mental states that often differ from our own. For example, by representing a person's *desire* to obtain an object and her *belief* about its location, we readily predict that she will search for the object in the location where she believes it to be. This situation involves meta-representation of others' mental states—explicitly reasoning about how concepts like desires and beliefs interact to guide behavior. The developmental and evolutionary origins of such capacities have been studied extensively throughout the last four decades (for reviews, see Arre & Santos, 2021; Baillargeon, Scott, & He, 2010; Call & Tomasello, 2008; Horschler, MacLean, & Santos, 2020b; Krupenye & Call, 2019; Rosati, Santos, & Hare, 2010; Saxe, 2013; Scott & Baillargeon, 2017; Slaughter, 2015; Wellman, Cross, & Watson, 2001). Many comparative studies show that non-human animals behave in ways that are consistent with an understanding of at least some mental states. However, whether any non-human animals possess a human-like representational ToM is still debated.

Most comparative ToM studies have focused on non-human primates (NHPs). This work has shown that NHPs are skilled at understanding whether other agents are aware of objects they have seen hidden in the past, and specifically that they have correct expectations about how *knowledgeable* agents will behave. For example, both monkeys (Arre, Stumph, & Santos, 2021; Drayton & Santos, 2018; Horschler, Santos, & MacLean, 2019; Marticorena, Ruiz, Mukerji, Goddu, & Santos, 2011) and apes (Hare, Call, & Tomasello, 2001; Kaminski, Call, & Tomasello,

---

* Corresponding author at: School of Anthropology, University of Arizona, Tucson, AZ 85719, USA.
E-mail address: horschler@email.arizona.edu (D.J. Horschler).

https://doi.org/10.1016/j.cognition.2021.104658
Received 24 September 2020; Received in revised form 2 March 2021; Accepted 3 March 2021

2008; Krachun, Carpenter, Call, & Tomasello, 2009) typically expect agents who have previously seen where a desired object was hidden to search for it in this location. In these cases, NHPs must represent some connection between the location of an object and an agent to infer that the agent will search for the object correctly, but the precise contents of these representations remain poorly understood.

Scholars have proposed a wide range of mechanisms to account for these findings. Low-level behavior-reading accounts propose that NHPs use rigid behavioral rules associating cues (e.g., an agent's eyes being visible when food was hidden) with observed outcomes (e.g., the agent correctly reaches for the food) to make predictions about how agents will behave (Penn & Povinelli, 2007; Povinelli & Vonk, 2003, 2012; see also the sub-mentalizing account: Heyes, 2014, 2017). Others have explained NHPs' performance using the same non-mentalistic explanations for successful performance in human infants (e.g., Kovacs, Teglas, & Endress, 2010; Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007; Woodward, 1998), such as minimal ToM (Apperly & Butterfill, 2009; Butterfill & Apperly, 2013), experiential record-keeping (Perner & Roessler, 2012), and teleological reasoning (Gergely & Csibra, 2003).

In contrast, some scholars have proposed more mentalistic hypotheses for NHPs' performance. One such account—the knowledge-ignorance account—hypothesizes that NHPs represent others' knowledge and ignorance states as such[1] (Call & Santos, 2012; Call & Tomasello, 2008; Kaminski et al., 2008; Rosati et al., 2010; Whiten, 2013). Although many previous findings are consistent with the knowledge-ignorance account, some have argued that this account cannot fully explain NHPs' performance (Martin & Santos, 2016). For example, the knowledge-ignorance account predicts that NHPs should understand what it means for an agent to be ignorant—that the agent does not know where an object is. However, there is currently no experimental evidence that NHPs make positive predictions about how ignorant agents will behave, such as demonstrating an expectation that an ignorant agent is likely to search for information that they lack. Additionally, there is no evidence that NHPs intentionally induce ignorance in others (e.g., by taking an object that an agent is aware of and hiding it in a new location), despite clear evidence for acting in ways that avoid making others aware (e.g., leaving information that an agent is unaware of hidden from them) (Byrne & Whiten, 1991; Hare, Call, & Tomasello, 2006; Karg, Schmelz, Call, & Tomasello, 2015; Whiten & Byrne, 1988).

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

Numerous hypotheses about the representations that guide NHPs' ToM have thus been proposed, but how should researchers distinguish between these nuanced accounts? One strategy is to explore the circumstances under which these representations break down. This approach has been used effectively to elucidate the contents of different representations in other domains. Consider, for example, the domain of numerical cognition. Initially, nonverbal number discrimination was believed to be wholly dependent on the ratio between numbers being compared (i.e., discrimination obeying Weber's law). However, subsequent work with both humans and NHPs demonstrated surprising scenarios in which this prediction did not hold (Feigenson, Carey, & Hauser, 2002; Hauser & Carey, 2003). Additionally, related studies showed that representations of small sets of objects can break down in both humans and NHPs based on how the objects move (e.g., in cases of discontinuous appearance and disappearance or dispersion and coalescence) or are presented (e.g., in cases where objects are embedded within one another or are connected by lines) (Huntley-Fenner, Carey, & Solimando, 2002; Mitroff, Scholl, & Wynn, 2004; Scholl, 2001; Scholl & Pylyshyn, 1999; VanMarle, Aw, McCrink, & Santos, 2006). Thus, by discovering these limitations, cognitive scientists were able to identify properties of the specific representations underlying performance on diverse number tasks. Here, we apply a similar approach to the study of mental state representation in NHPs. More specifically, we assess the conditions that cause NHPs' representations of an agent's awareness of a hidden object to break down.

Previous work has shown that NHPs no longer expect an agent to be aware of an object they saw hidden after the object was arbitrarily moved out of and back into a hiding location while the agent could not see (Horschler et al., 2019; Kaminski et al., 2008). Additionally, one study found that monkeys no longer expected an agent to be aware of the location of an object she saw hidden in one of two differently colored boxes after the boxes rotated into different positions while the agent's view was occluded (Drayton & Santos, 2018). However, monkeys did attribute awareness to the agent when she saw the rotation, suggesting that monkeys maintained representations of the agent's awareness only when the agent witnessed all positional manipulations. Thus, current evidence strongly suggests that positional movement of an object while outside of an agent's perceptual access is sufficient to disrupt attribution of awareness to the agent. However, the specific factors governing when and why these types of manipulations interfere with knowledge-like attribution remain poorly understood.

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

To better understand the specific factors governing when NHPs' representations of others' awareness break down, we assessed the impacts of a novel geometric and non-geometric change to an object's properties while it was outside the agent's perceptual access. The geometric manipulation was designed to alter spatial properties of the object (i.e., size and shape) without impacting the object's location, while the non-geometric manipulation was designed to alter a non-spatial property (i.e., color). Our method was based on a commonly-used violation of expectation task (Marticorena et al., 2011), in which monkeys have robustly demonstrated an understanding of agents' awareness of a hidden object (see also Arre, Clark, & Santos, 2019; Arre et al., 2020; Drayton & Santos, 2018; Horschler et al., 2019; Martin & Santos, 2014).

---

[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).

In this task, subjects watch as an agent observes an object hidden in one of two boxes. Typically, monkeys look significantly longer when the agent reaches into the incorrect as opposed to correct box to search for it, suggesting that they expect the agent to search correctly based on his awareness of the object's location. In our experiments, we similarly created situations in which an agent always maintained awareness about the correct location of an object and tested whether monkeys had differing expectations about where the agent would search after the object's size/shape or color were manipulated while the agent was not looking.

2. Experiment 1

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

2.1. Methods

2.1.1. Subjects

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

2.1.2. Methods and apparatus

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

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

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

2.1.3. Procedure

Testing sessions consisted of two familiarization trials and one test trial (Fig. 1). Familiarization trials served to familiarize the subject to the apparatus and the potential for the experimenter to reach into a box. In the first familiarization trial, the experimenter dropped the front occluder to reveal a small, un-blossomed flower situated on the stage between the two boxes. When the occluder dropped, the experimenter stared downward at the flower and said "now" to begin the 10 s trial. After the first familiarization trial, the cameraperson announced the condition using an alphanumeric code. The cameraperson was blind to which codes corresponded to which conditions. Subjects were assigned to conditions pseudo-randomly to balance mean ages and sex ratios across conditions. The condition code was announced after the first familiarization trial because a subject's condition determined which box the experimenter reached into in the second familiarization trial.

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

After the second familiarization trial, the test trial began (Video S1). In the test trial, the experimenter dropped the front occluder to reveal the unblossomed flower situated on the stage between the two boxes, and subsequently watched the flower as it moved into one of the boxes. The experimenter then raised the back occluder to block his view of the stage. While the experimenter's view was occluded, the box containing the flower flipped open and the flower "bloomed" (i.e., growing larger while exposing the interior of the petals) while in view of the subject but not the agent. The experimenter controlled the flower's bloom surreptitiously by pulling back on a 2 cm transparent ring from behind the stage via attached fishing line. The box then flipped closed, the experimenter dropped the back occluder such that he could again see the stage, and then reached into one of the boxes (between-subjects). After reaching, the experimenter said "now" and subjects' looking was filmed for 10 s. If monkeys were able to represent that the agent was still aware of the flower's location after its change in size/shape, we expected them to look significantly longer when the agent reached into the incorrect

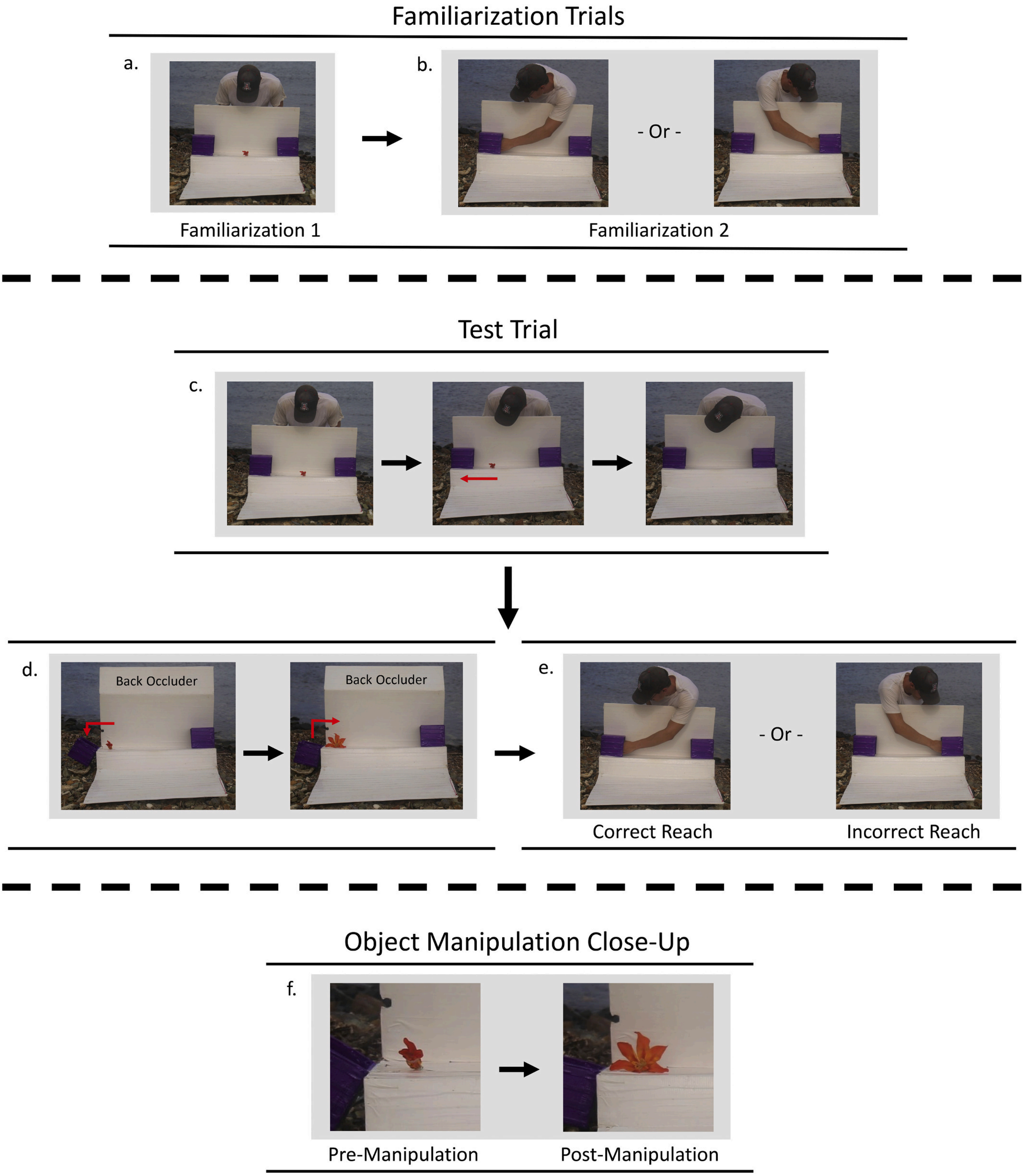

Fig. 1. A depiction of the procedure for Experiment 1. **(a)** Familiarization 1: The agent stared at the un-blossomed flower positioned in the middle of the stage for 10 s. **(b)** Familiarization 2: The agent reached into one of the two boxes (consistent with the final reach direction in the test trial)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

rather than the correct box. However, if subjects were no longer able to represent the agent's awareness after the flower changed size/shape, we expected them to look equally long in both conditions.

2.1.4. Video coding

As in previous studies (Arre et al., 2019; Drayton & Santos, 2018; Horschler et al., 2019; Martin & Santos, 2014), all trials were coded independently by two coders using the programs MPEG Streamclip or BORIS (Friard & Gamba, 2016). Videos were clipped such that coders were blind to the condition and trial. Only the subject's face (and occasionally body depending on the level of zoom in the original recording) was visible in the 10-s-long clips, and all subjects were centered in front of the apparatus to begin each trial. Coders assessed each 10 s trial at 30 frames/s beginning immediately after the experimenter said "now" by recording each frame where the subject was looking at the apparatus. Interrater reliability between coders was excellent (Pearson's $R = 0.92$).

2.2. Results

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

To assess differences in looking time between monkeys who saw the agent reach correctly versus incorrectly in the test trial, we fit a linear mixed model with a random intercept for subject predicting looking time as a function of trial type (Familiarization 1, Familiarization 2, or Test), condition (correct or incorrect reach), and the interaction between trial type and condition. Results showed a significant interaction between trial type and condition ($\chi^2(2) = 6.52$, $p = .04$). We next examined planned contrasts comparing looking time on each of the three trial types across conditions. After Dunn-Sidak correction for multiple comparisons, there was no significant difference in looking time between monkeys who saw a correct reach and those who saw an incorrect reach on the test trial in either Familiarization 1 (Correct: $M = 5.60$ s; Incorrect: $M = 4.85$ s; $t(243) = 1.49$, $p = .36$; Table 2; Fig. 2) or Familiarization 2 (Correct: $M = 4.26$ s; Incorrect: $M = 4.37$ s; $t(243) = -0.20$; $p = .99$; Table 2; Fig. 2). Thus, there were no general differences in how interested each group of monkeys was in viewing events on the stage. In the test trial, there was also no significant difference in looking time between the monkeys who saw a correct reach ($n = 50$; $M = 3.31$ s) versus an incorrect reach ($n = 49$; $M = 3.84$ s; $t(243) = -1.50$, $p = .36$, Dunn-Sidak corrected; Table 2; Fig. 2). This suggests that monkeys had no expectation about the agent's behavior despite his awareness of the flower's location after its size/shape changed outside of the agent's perpetual access.

2.3. Discussion

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

We designed the current procedure such that in reality the agent always maintained awareness of the flower's location—the flower never changed locations after the agent observed where it was hidden. A monkey who maintained a representation of the agent's awareness should have expected him to reach correctly rather than incorrectly to obtain the flower. However, we found no evidence for expectations of where the agent would reach after the flower's size/shape changed. While monkeys may have perceived this change as an alteration to an object's identity, this possibility seems unlikely given that monkeys always had direct perceptual access to the flower's transformation (see 'General discussion'). These findings are consistent with previous studies showing that spatial transformations of an object appear to

Table 2

Mean looking time ± standard error for each trial in all three experiments broken down by condition (i.e., monkeys who ultimately saw an incorrect reach versus a correct reach in the test trial for Experiment 1 and 2; monkeys who ultimately saw a green [color change] versus a yellow lemon [control] in the test trial for Experiment 3), along with test statistics for all comparisons, including estimates from planned contrasts examining log-transformed looking times between conditions within each trial type, *t* statistics, degrees of freedom, and *p*-values. Asterisks indicate significant differences at an alpha level of 0.05 (*), 0.01 (**), or 0.001 (***).

Experiment	Trial	Looking Time: Incorrect Reach	Looking Time: Correct Reach	Contrast Estimate	*t*	*df*	*p*
1	Familiarization 1	4.85 ± 0.32	5.60 ± 0.38	0.18	1.49	243	0.36
	Familiarization 2	4.37 ± 0.32	4.26 ± 0.33	−0.02	−0.20	243	0.99
	Test	3.84 ± 0.32	3.31 ± 0.30	−0.19	−1.50	243	0.36
2	Familiarization 1	5.54 ± 0.37	4.55 ± 0.35	−0.24	−1.64	181	0.28
	Familiarization 2	3.82 ± 0.40	3.50 ± 0.30	0.004	0.03	181	1.00
	Test	4.50 ± 0.37	3.02 ± 0.31	−0.43	−2.99	181	0.01**

Experiment	Trial	Looking Time: Color Change	Looking Time: Control	Contrast Estimate	*t*	*df*	*p*
3	Familiarization 1	4.48 ± 0.45	4.11 ± 0.52	−0.20	−0.83	123	0.93
	Familiarization 2	2.87 ± 0.40	2.83 ± 0.38	−0.02	−0.01	123	1.00
	Test	4.82 ± 0.43	2.24 ± 0.35	−1.14	−4.70	123	<0.001***

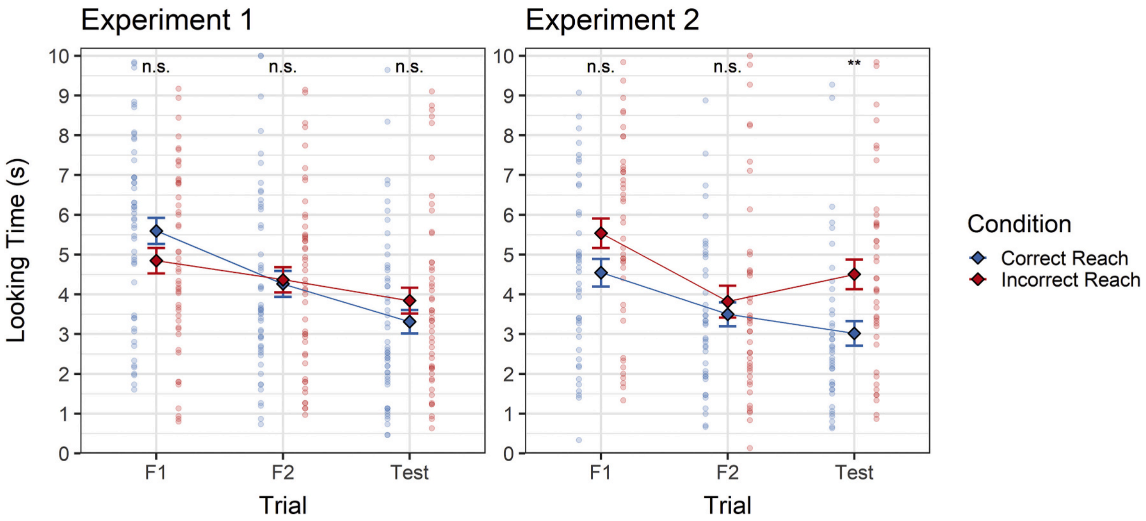

Fig. 2. Diamond points show mean looking time in seconds ± standard error grouped by condition in the first familiarization trial (F1), the second familiarization trial (F2), and the test trial (Test) in Experiments 1 and 2. Circular points show individual-level data. Asterisks indicate significant differences at an alpha level of 0.05 (*), 0.01 (**), or 0.001 (***), while 'n.s.' denotes no significant difference.

disrupt representations of an agent's awareness of this object's location (e.g., Horschler et al., 2019; Kaminski et al., 2008). Our findings extend this work by showing that geometric manipulations which never impact an object's location/position also disrupt monkeys' representations, suggesting that even more subtle spatial manipulations interfere with representations about whether the agent is aware of the object's location.

3. Experiment 2

To test the prediction that spatial manipulations *specifically* break representations of an agent's awareness of an object's location, Experiment 2 introduced a non-geometric manipulation of a different property of a target object (color). We hypothesized that non-geometric manipulations should not break these representations if our findings from Experiment 1 are the result of geometric object manipulations interfering with the spatial information contained in awareness relations about the object's location. In Experiment 2, monkeys watched as an agent saw a target object (a lemon) move into one of two boxes. While the agent's view was blocked, the box containing the lemon flipped open and the lemon appeared to change color. The box then flipped closed, and the agent reappeared and reached into either the correct or incorrect box. If any manipulation (geometric or non-geometric) of a target object breaks monkeys' representations of the agent's awareness, we expected that monkeys would look equally long regardless of where the agent reached. However, if non-geometric manipulations (i.e., a change to the object's color in this case) do not break monkeys' representations of the agent's awareness, we predicted that monkeys would expect the agent to reach correctly for the object, and thus look significantly longer at incorrect as compared to correct reaches.

3.1. Methods

3.1.1. Subjects

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

3.1.2. Methods and Apparatus

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

3.1.3. Procedure

Testing sessions consisted of two familiarization trials and one test trial, and the procedure was identical to Experiment 1 except for following (Fig. 3; Video S2). Rather than a flower, we used a lemon as the target object. We chose a lemon based on previous studies using similar paradigms (Horschler et al., 2019; Marticorena et al., 2011), and because it allowed for color manipulation by painting one half of it green (such that it looked like a lime). During the familiarization trials and at the beginning of the test trial, only the yellow side of the lemon was visible to the subject. In the test trial, the experimenter dropped the front occluder to reveal the yellow lemon situated on the stage between the two boxes, and subsequently watched the lemon as it moved into one of the boxes. The experimenter then raised the back occluder to block his view of the stage. While the experimenter's view was occluded, the box containing the lemon opened and the lemon rapidly rotated on its axis, appearing to turn green without moving. The experimenter surreptitiously controlled the rotation by quickly twisting a wooden dowel attached to the lemon from underneath the stage. The box then closed, the experimenter dropped the back occluder such that he could again see the stage, and then reached into either the correct or incorrect box (between-subjects).

3.1.4. Video coding

All coding procedures were the same as in Experiment 1. Interrater reliability between coders was excellent (Pearson's $R = 0.94$).

3.2. Results

Our analytical approach was identical to Experiment 1. Results showed a significant interaction between trial type and condition ($\chi^2(2)$

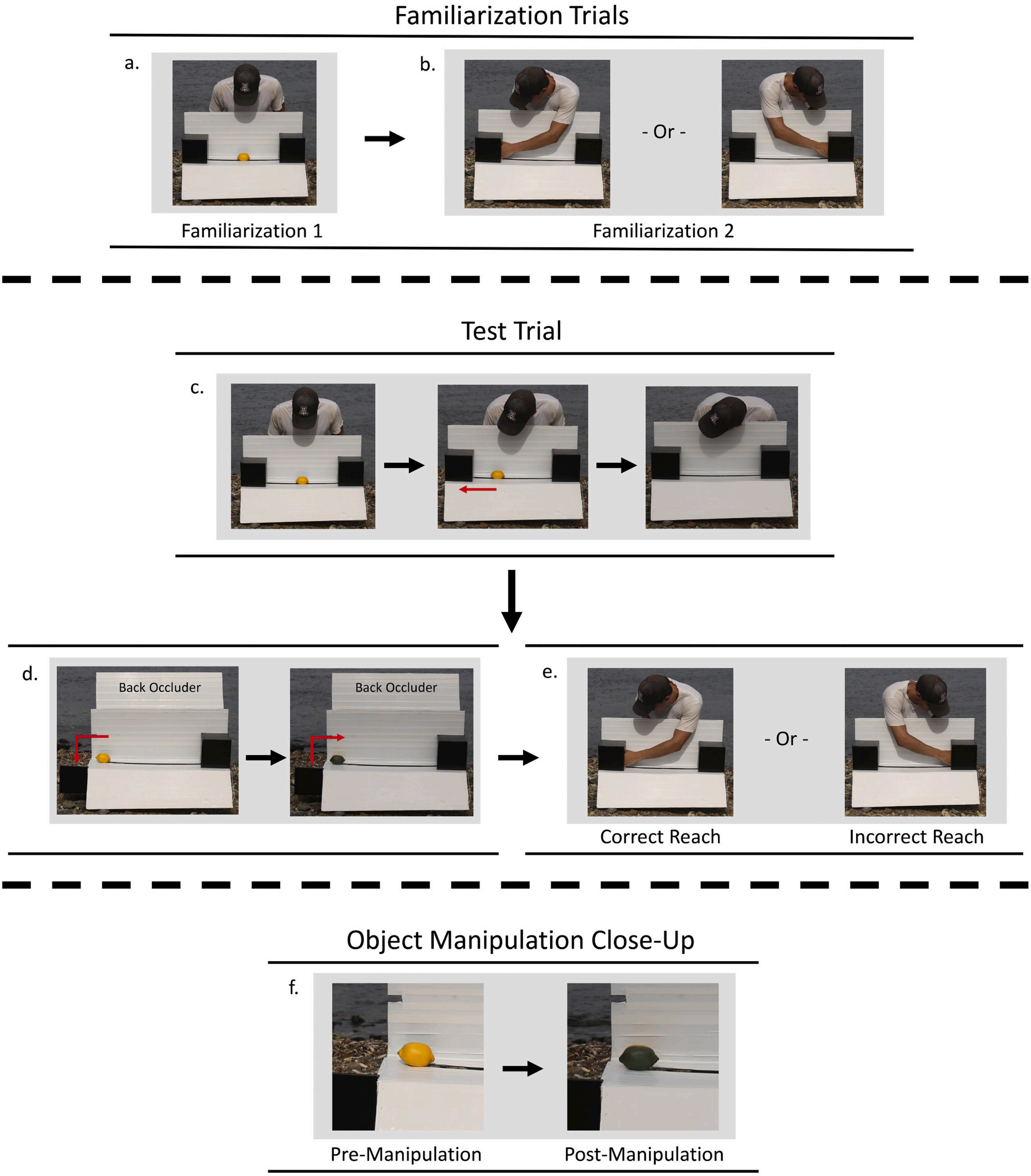

Fig. 3. A depiction of the procedure for Experiment 2. **(a)** Familiarization 1: The agent stared at the lemon positioned in the middle of the stage for 10 s. **(b)** Familiarization 2: The agent reached into one of the two boxes (consistent with the final reach direction in the test trial) for 10 s. **(c)** Test trials: All monkeys watched as the agent observed the lemon moving into one of the two boxes. **(d)** While the agent's view of the stage was occluded, the box hiding the lemon opened over the end of the stage, the lemon turned green, and the box closed. **(e)** Finally, the agent reached into either the correct or incorrect box while the monkey's looking behavior was recorded for 10 s. Panel **(f)** provides a close-up view of the lemon's manipulation in panel **(d)**. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$= 7.64$, $p = .02$). After Dunn-Sidak correction, there were no significant differences in looking time between monkeys who saw a correct reach and those who saw an incorrect reach on the test trial in either Familiarization 1 (Correct: $M = 4.55$ s; Incorrect: $M = 5.54$ s; $t(181) = -1.64$, $p = .28$; Table 2; Fig. 2) or Familiarization 2 (Correct: $M = 3.50$ s; Incorrect: $M = 3.82$ s; $t(181) = 0.03$, $p = 1.00$; Table 2; Fig. 2). However, in the test trial, monkeys who saw an incorrect reach ($n = 41$; $M = 4.50$ s) looked significantly longer than monkeys who saw a correct reach ($n = 41$; $M = 3.02$ s; $t(181) = -2.99$, $p = .01$, Dunn-Sidak corrected; Table 2; Fig. 2) with a large effect size ($d = -0.85$). This suggests that monkeys expected the agent to still be aware of the lemon's location and were thus surprised when he reached incorrectly, even after it changed color while the agent was not looking.

3.3. Discussion

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

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

4. Experiment 3

In Experiment 3, we used a habituation-dishabituation paradigm—a method well-suited to examining whether subjects can discriminate between two different stimuli—to ensure that monkeys perceived the lemon's change from yellow to green in a situation not requiring representation of the agent's awareness. Monkeys were first habituated to the yellow lemon across two familiarization trials, and then saw either the yellow lemon (control condition) or a green lemon (color change condition) in a test trial. We predicted that if monkeys detected the color change, they should exhibit dishabituation in the color change, but not the control condition.

4.1. Methods

4.1.1. Subjects

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

4.1.2. Methods and apparatus

The apparatus and target object were identical to Experiment 2.

4.1.3. Procedure

Sessions consisted of two familiarization trials and one test trial. The familiarization trials were identical across conditions (control or color change; described below) and served to familiarize subjects with the apparatus and habituate them to the yellow lemon. In both familiarization trials, the experimenter dropped the front occluder to reveal an empty stage. The experimenter then surreptitiously flipped open one of the boxes (balanced between subjects but consistent within subjects) to reveal a yellow lemon and stared downward at the lemon for 10 s, during which the experimenter and lemon remained motionless. After the second familiarization trial, the cameraperson announced the condition using an alphanumeric code. In the control condition, the test trial was identical to both familiarization trials (i.e., the lemon appeared as yellow) and in the 'color change' condition, the lemon instead appeared green. As in Experiments 1 and 2, test trial condition varied between-subjects.

4.1.4. Video coding

All coding procedures were the same as in Experiment 2. Interrater reliability between coders was excellent (Pearson's $R = 0.92$).

4.2. Results and discussion

Our analytical approach was the same as in Experiments 1 and 2, but with the addition of three planned contrasts (described below). To assess differences in looking time between monkeys who saw a color change versus no color change, we ran a linear mixed model with a random intercept for subject predicting looking time as a function of trial type (Familiarization 1, Familiarization 2, or Test), condition (control or color change), and the interaction between trial type and condition. Results showed a significant interaction between trial type and condition ($\chi^2(2) = 21.81$, $p < .001$). After Dunn-Sidak correction, there was no significant difference in looking time between the monkeys who saw a color change and those who saw no color change on the test trial in either Familiarization 1 (Control: $M = 4.11$ s; Color Change: $M = 4.48$ s; $t(123) = -0.83$, $p = .93$; Table 2; Fig. 4) or Familiarization 2 (Control: $M = 2.83$ s; Color Change: $M = 2.87$ s; $t(123) = -0.01$, $p = 1.00$; Table 2; Fig. 4). However, in the test trial, monkeys who saw the lemon as green ($n = 30$; $M = 4.82$ s) looked significantly longer than monkeys who saw the lemon as yellow ($n = 29$; $M = 2.24$ s; $t(123) = -4.70$, $p < .001$, Dunn-Sidak corrected; Table 2; Fig. 4) with a large effect size ($d = -1.63$). This suggests that monkeys expected to again see a yellow lemon in the test trial (as in the familiarization trials) and thus looked longer (i.e., were surprised) when the lemon was instead green.

Of the three additional planned contrasts (subject to Dunn-Sidak correction where appropriate), two tested for significant differences in looking time between the second familiarization trial and the test trial within each condition (dishabituation), and the third was an interaction contrast to test for a significant difference in the extent of looking time change from the second familiarization trial to the test trial across conditions. Monkeys who saw the lemon as yellow in the test trial (control condition) showed no significant difference in looking time between the second familiarization trial ($M = 2.83$ s) and the test trial ($M = 2.24$ s; $t(114) = 2.29$, $p = .11$). However, monkeys who saw the green lemon in the test trial looked significantly longer in the test trial ($M = 4.82$ s) as compared to the second familiarization trial ($M = 2.87$ s; $t(114) = -3.87$, $p < .001$) with a large effect size ($d = -1.00$). Additionally, an interaction contrast showed that the degree of looking time

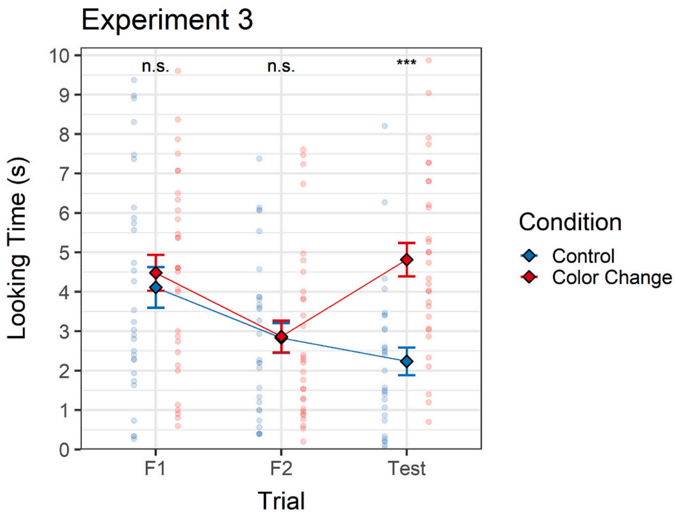

Fig. 4. Diamond points show mean looking time in seconds ± standard error grouped by condition in the first familiarization trial (F1), the second familiarization trial (F2), and the test trial (Test) in Experiment 3. Circular points show individual-level data. Asterisks indicate significant differences at an alpha level of 0.05 (*), 0.01 (**), or 0.001 (***), while 'n.s.' denotes no significant difference.

change from the second familiarization trial to the test trial was significantly greater in the color change (1.95 s mean increase) compared to the control condition (0.59 s mean decrease; $t(114) = 4.35$, $p < .001$), with a large effect size ($d = 1.60$). Taken together, these results suggest that monkeys attended to and perceived the color change, showing dishabituation in the test trial of the color change but not the control condition. Furthermore, this supports the conclusion that in Experiment 2, monkeys maintained representations of the agent's awareness of the hidden object's location despite likely perceiving its color change.

5. General discussion

Our experiments aimed to assess the content of the representations that underlie NHPs' understanding of how knowledgeable agents use knowledge to guide their behavior. We adapted a commonly-used task in which monkeys have robustly demonstrated an understanding of agents' awareness of the location of a hidden object and manipulated the object's size/shape (Experiment 1) or color (Experiment 2) while the agent was not looking to assess whether these manipulations interfered with subjects' attributions of awareness to the agent. Importantly, the agent maintained awareness of the object's location in both experiments because he witnessed the object move into its hiding location and it never changed locations while he could not see it. Nevertheless, monkeys did not expect the agent to maintain awareness when the object changed size/shape (Experiment 1) but did expect the agent to maintain awareness after the object changed colors (Experiment 2). Experiment 3 suggested that monkeys are able to notice the object's color change, making it unlikely that their performance in Experiment 2 resulted from a failure to perceive the object manipulation. Taken together, these results suggest that, at least in cases of awareness of an object's location, monkeys' representations of other agents' awareness are disrupted specifically by spatial manipulations of the object.

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

In the current study, we implemented a much more subtle spatial manipulation (a change in size/shape) which did not involve a change of location. However, our results suggest that even simple geometric transformations of a stationary object appear to disrupt representations of an agent's awareness, similarly to larger scale positional movements—ones more likely to disrupt what an agent knew. Importantly, the results of Experiment 2 show that not all manipulations to an object's properties are sufficient to interfere with monkeys' representations of awareness; we found that non-geometric manipulations (i.e., a color change) do not have this same effect. We hypothesize that spatial but not non-spatial object manipulations have this effect because we specifically tested monkeys' representations of an agent's awareness of an object's location, which must contain spatial information about the object. Thus, spatial manipulations may result in monkeys updating the spatial content of these representations in a way that decouples representation of the agent's awareness from representation of the object.

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

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

In our study, monkeys expected the agent to still be aware of an object's location after it changed colors outside of his perceptual access, but did not expect the agent to still be aware of an object's location after it changed sizes/shapes. One alternative interpretation of these results is that monkeys perceived the flower's size/shape change but not the lemon's color change as an alteration to an object's *identity* (i.e., they may have represented the altered flower as a different object altogether, while representing the altered lemon as having the same identity despite a change to one of its properties) or simply as a greater or more consequential transformation. Previous work with humans suggests that infants as young as 14 months of age reason about the identity of an object as it relates to an agent's beliefs (Buttelmann, Suhrke, & Buttelmann, 2015; Scott & Baillargeon, 2009; Scott, Richman, & Baillargeon, 2015; Song & Baillargeon, 2008). Less work has explored representations of object identity in NHPs, but one study (Krachun, Carpenter, Call, & Tomasello, 2010) found hints that chimpanzees may be able to reason about other agents' representations of object identity in some contexts. In this study, chimpanzees first learned that an agent would reliably hide a box containing one of two different types of food (i.e., a banana slice or a grape) in one of two different locations based on which type of food was inside. When one type of food was swapped for the other, five chimpanzees showed a trend toward understanding an agent's knowledge of the identity change when the agent observed the food being swapped (i.e., they correctly responded on 80% of trials, but this result was not statistically significant) but not their false beliefs about object identity when the food was swapped while the agent was absent (i.e., they responded correctly on less than 50% of trials). In Experiments 1 and 2 of the present study, monkeys witnessed each object's transformation such that they directly observed that the object was not substituted for another. Therefore, our study differs from previous work using object identity manipulations in that it involved the transformation of a single object rather than the substitution of one object for another. It therefore seems unlikely that monkeys represented the altered versions of the objects as different objects altogether given that monkeys always witnessed each object's transformation. However, given how little is known about monkeys' representations of objects, we cannot rule out the possibility that monkeys perceived the transformation in Experiment 1 as being more consequential than that of Experiment 2.

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

Although more work is needed to assess monkeys' perceptions of object identity and how other manipulations may influence representations of awareness, our results suggest that these representations share some similarities with theoretical accounts of implicit ToM in humans. Some scholars have proposed that adult humans employ two separate and dissociable ToM systems: a slower, effortful, and flexible explicit system alongside a more rapid, automatic, and efficient implicit system (Apperly & Butterfill, 2009; Butterfill & Apperly, 2013; Carruthers, 2017; Schneider, Slaughter, & Dux, 2017). The explicit system is proposed to allow for conscious representation of an agent's mental state, whereas the implicit system is used for rapid, unconscious predictions about how an agent's mental state may influence their behavior (Schneider et al., 2017). Apperly and Butterfill proposed that an implicit "minimal ToM" system could operate using representations of simple relational states in a similar (but not identical) way to awareness relations (Butterfill & Apperly, 2013). Minimal ToM predicts that subjects "register" objects at the last location they were "encountered," and that these registrations are maintained regardless of any spatial manipulation of the object while the agent cannot see it. Therefore, under this account, subjects should make correct behavioral predictions about agents with false beliefs, at least in change-of-location paradigms. Importantly, the awareness relations account predicts that subjects will have no expectations about an agent's behavior in these change-of-location false belief tests: the spatial manipulations requisite of inducing a false belief in an agent are predicted to also break the subject's representation of the agent's awareness.

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

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

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2021.104658.

Ethics approval

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

Author statement

All authors contributed to Conceptualization, Methodology, and Writing – Review & Editing. DJH: Data Curation, Formal Analysis, Investigation, Visualization, and Writing – Original Draft. LRS: Funding Acquisition, Project Administration, and Supervision. ELM: Data Curation, Formal Analysis, Project Administration, Supervision, and Visualization.

Data accessibility

Data are available as electronic supplementary material.

Funding

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

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgments

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

References

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*(4), 953–970. https://doi.org/10.1037/a0016923.

Arre, A. M., Clark, C. S., & Santos, L. R. (2019). Do young rhesus macaques know what others see?: A comparative developmental perspective. *American Journal of Primatology, March*, 1–9. https://doi.org/10.1002/ajp.23054.

Arre, A. M., & Santos, L. R. (2021). Mentalizing in Non-Human Primates. In K. Ochsner, & M. Gilead (Eds.), *The Neural Bases of Mentalizing (pp)*. New York, NY: Springer.

Arre, A. M., Stumph, E., & Santos, L. R. (2021). Macaque species with varying social tolerance show no differences in understanding what other agents perceive. *Animal Cognition*.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*(3), 110–118. https://doi.org/10.1016/j.tics.2009.12.006.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1). Doi:10.18637/jss.v067.i01.

Brent, L. J. N., MacLarnon, A., Platt, M. L., & Semple, S. (2013). Seasonal changes in the structure of rhesus macaque social networks. *Behavioral Ecology and Sociobiology, 67* (3), 349–359. https://doi.org/10.1007/s00265-012-1455-8.Seasonal.

Buttelmann, D., Buttelmann, F., Carpenter, M., Call, J., & Tomasello, M. (2017). Great apes distinguish true from false beliefs in an interactive helping task. *PLoS One, 12* (4), 1–13. https://doi.org/10.1371/journal.pone.0173793.

Buttelmann, F., Suhrke, J., & Buttelmann, D. (2015). What you get is what you believe: Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. *Journal of Experimental Child Psychology, 131*, 94–103. https://doi.org/10.1016/j.jecp.2014.11.009.

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language, 28*(5), 606–637. https://doi.org/10.1111/mila.12036.

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

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

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

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

Core Team, R. (2020). *A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. https://doi.org/citeulike-article-id:2400517.

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

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

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression (third ed.)*. Thousans Oaks, CA: Sage.

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

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

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

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

Hauser, M. D., & Carey, S. (2003). Spontaneous representations of small numbers of objects by rhesus macaques: Examinations of content and format. *Cognitive Psychology, 47*. https://doi.org/10.1016/S0010-0285(03)00050-1.

Hayashi, T., Akikawa, R., Kawasaki, K., Egawa, J., Minamimoto, T., Kobayashi, K., … Hasegawa, I. (2020). Macaques exhibit implicit gaze bias anticipating others' false-belief-driven actions via medial prefrontal cortex. *Cell Reports, 30*(13), 4433–4444. e5. https://doi.org/10.1016/j.celrep.2020.03.013.

Heyes, C. (2014). Submentalizing: I am not really Reading your mind. *Perspectives on Psychological Science, 9*(2), 131–143. https://doi.org/10.1177/1745691613518076.

Heyes, C. (2017). Apes Submentalise. *Trends in Cognitive Sciences, 21*(1), 1–2. https://doi.org/10.1016/j.tics.2016.11.006.

Horschler, D. J., MacLean, E. L., & Santos, L. R. (2020a). Advancing gaze-based research on primate theory of mind. *Trends in Cognitive Sciences, 24*(10), 778–779.

Horschler, D. J., MacLean, E. L., & Santos, L. R. (2020b). Do non-human primates really represent others' beliefs? *Trends in Cognitive Sciences, 24*(8), 594–605.

Horschler, D. J., Santos, L. R., & MacLean, E. L. (2019). Do non-human primates really represent others' ignorance? A test of the awareness relations hypothesis. *Cognition, 190*(March), 72–80. https://doi.org/10.1016/j.cognition.2019.04.012.

Huntley-Fenner, G., Carey, S., & Solimando, A. (2002). Objects are individuals but stuff doesn't count: Perceived rigidity and cohesiveness influence infants' representations of small groups of discrete entities. *Cognition, 85*(3), 203–221. https://doi.org/10.1016/S0010-0277(02)00088-4.

Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition, 109*(2), 224–234. https://doi.org/10.1016/j.cognition.2008.08.010.

Kano, F., Call, J., & Krupenye, C. (2020). Primates pass dynamically social false-belief tests. *Trends in Cognitive Sciences*, 1–2. https://doi.org/10.1016/j.tics.2020.07.003.

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

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

Kovacs, A. M., Teglas, E., & Endress, A. D. (2010). The social sense: Susceptibility to Others' beliefs in human infants and adults. *Science, 330*, 1830–1834. https://doi.org/10.1126/science.1190792.

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

Krachun, C., Carpenter, M., Call, J., & Tomasello, M. (2010). A new change-of-contents false belief test: Children and chimpanzees compared. *International Journal of Comparative Psychology, 23*, 145–165.

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

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

Lenth, R. (2020). *Emmeans: Estimated marginal means, aka least-squares means*.

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

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

Martin, A., & Santos, L. R. (2014). The origins of belief representation: Monkeys fail to automatically represent others' beliefs. *Cognition, 130*(3), 300–308. https://doi.org/10.1016/j.cognition.2013.11.016.

Martin, A., & Santos, L. R. (2016). What cognitive representations support primate theory of mind? *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2016.03.005. Elsevier Ltd.

Mitroff, S. R., Scholl, B. J., & Wynn, K. (2004). Divide and conquer: How object files adapt when a persisting object splits into two. *Psychological Science, 15*(6), 420–425. https://doi.org/10.1038/s41579-018-0146-5.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science, 308*(5719), 255–258. https://doi.org/10.1126/science.1107621.

Penn, D. C., & Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a "theory of mind.". *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362*(1480), 731–744. https://doi.org/10.1098/rstb.2006.2023.

Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences, 16*(10), 519–525. https://doi.org/10.1016/j.tics.2012.08.004.

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences, 7*(4), 157–160. https://doi.org/10.1016/S1364-6613(03)00053-6.

Povinelli, D. J., & Vonk, J. (2012). We don't need a microscope to explore the chimpanzee's mind. In *, vol. 19. Mind & Language* (pp. 1–28). https://doi.org/10.1093/acprof:oso/9780198528272.003.0018.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 4*(1978), 515–526.

Rawlins, R. G., & Kessler, M. J. (1986). *The Cayo Santiago macaques: History, behavior, and biology*. SUNY Press.

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

Saxe, R. (2013). The new puzzle of theory of mind development. *Navigating the Social World*, 107–112. https://doi.org/10.1093/acprof:oso/9780199890712.003.0020.

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

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

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual Objecthood. *Cognitive Psychology, 38*(2), 259–290. https://doi.org/10.1006/cogp.1998.0698.

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

Scott, R. M., & Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences, 21*(4), 237–249. https://doi.org/10.1016/j.tics.2017.01.012.

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

Slaughter, V. (2015). Theory of mind in infants and young children: A review. *Australian Psychologist, 50*(3), 169–172. https://doi.org/10.1111/ap.12080.

Song, H. J., & Baillargeon, R. (2008). Infants' reasoning about Others' false perceptions. *Developmental Psychology, 44*(6), 1789–1795. https://doi.org/10.1037/a0013774.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18*(7), 587–592.

Ullman, T. D., Spelke, E., Battaglia, P., & Tenenbaum, J. B. (2017). Mind games: Game engines as an architecture for intuitive physics. *Trends in Cognitive Sciences, 21*(9), 649–665. https://doi.org/10.1016/j.tics.2017.05.012.

VanMarle, K., Aw, J., McCrink, K., & Santos, L. R. (2006). How capuchin monkeys (Cebus apella) quantify objects and substances. *Journal of Comparative Psychology, 120*(4), 416–426. https://doi.org/10.1037/0735-7036.120.4.416.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of mind development: The truth about false belief. *Child Development, 72*(3), 655–684. https://doi.org/10.1016/0010-0277(83)90004-5.

Whiten, A., & Byrne, R. W. (1988). Tactical deception in primates. *Behavioral and Brain Sciences, 11*, 233–273.

Whiten, A. (2013). Humans are not alone in computing how others see the world. *Animal Behaviour, 86*(2), 213–221. https://doi.org/10.1016/j.anbehav.2013.04.021.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4.