Great apes use self-experience to anticipate an agent's action in a false-belief test

Fumihiro Kano[a,b]

Christopher Krupenye[c]

Satoshi Hirata[a]

Masaki Tomonaga[d]

Josep Call[c,e]

[a]Kumamoto Sanctuary, Wildlife Research Center, Kyoto University, 869-3201 Kumamoto, Japan;
[b]Institute for Advanced Study, Kyoto University, 606-8501 Kyoto, Japan;
[c]School of Psychology and Neuroscience, University of St Andrews, KY16 9JP St Andrews, United Kingdom;
[d]Primate Research Institute, Kyoto University, 484-8506 Inuyama, Japan;
[e]Department of Developmental and Comparative Psychology, Max-Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany

Human social life depends on theory of mind, the ability to attribute mental states to oneself and others. A signature of theory of mind, false belief understanding, requires representing others' views of the world, even when they conflict with one's own. After decades of research, it remains controversial whether any nonhuman species possess a theory of mind. One challenge to positive evidence of animal theory of mind, the behavior-rule account, holds that animals solve such tasks by responding to others' behavioral cues rather than their mental states. We distinguish these hypotheses by implementing a version of the "goggles" test, which asks whether, in the absence of any additional behavioral cues, animals can use their own self-experience of a novel barrier being translucent or opaque to determine whether another agent can see through the same barrier. We incorporated this paradigm into an established anticipatory-looking false-belief test for great apes. In a between-subjects design, apes experienced a novel barrier as either translucent or opaque, although both looked identical from afar. While being eye tracked, all apes then watched a video in which an actor saw an object hidden under 1 of 2 identical boxes. The actor then scuttled behind the novel barrier, at which point the object was relocated and then removed. Only apes who experienced the barrier as opaque visually anticipated that the actor would mistakenly search for the object in its previous location. Great apes, therefore, appeared to attribute differential visual access based specifically on their own past perceptual experience to anticipate an agent's actions in a false-belief test.

anticipatory looking; behavior rule; goggles test; nonhuman animals; theory of mind

In a seminal article entitled "Does the chimpanzee have a theory of mind?," Premack and Woodruff (1) defined theory of mind as the ability to impute unobservable mental states to oneself and others. After decades of research, it remains controversial whether theory of mind is uniquely human (2, 3). Particularly central to this debate is the question of whether any nonhuman animals exhibit false-belief understanding, the hallmark of a representational theory of mind. False-belief understanding reflects awareness that others' behavior is driven not by reality but by beliefs about reality—and that those beliefs may differ from one's own.

Call and Tomasello (2) argued that chimpanzees do, indeed, possess a theory of mind—as evidenced by convergent performance on diverse social cognitive experiments—but that they likely lack its richest signature, an understanding of others' false beliefs. More recently, however, great apes (chimpanzees, bonobos, and orangutans) have passed several false-belief tests (4, 5), adapted from minimally demanding paradigms originally developed for human infants (6–8). For example, Krupenye, Kano, and colleagues (4) developed an anticipatory-looking test (based on ref. 6) in which apes watched videos while their gaze was noninvasively eye tracked. In the videos, an actor looked on as an object was hidden in 1 of 2 locations, but the object was, subsequently, moved or removed while the actor was away. As the actor approached centrally toward the 2 locations, in anticipation of his search for the hidden object, apes looked to the location where the actor falsely believed the object to be, even though they knew it was no longer there. The experimental design controlled for several lower-level explanations and a subsequent control experiment ensured that apes were not simply responding to domain-general cues (9–11). These results are consistent with the possibility that apes anticipated the actor's action by tracking his beliefs about the object's location.

However, they are also open to the possibility that apes relied on a behavior rule, innate or learned, that agents tend to search for things where they last saw them (4, 12, 13). Similar behavior- rule or behavior-abstraction accounts have been levied against findings from human infants (7, 13, 14) as well as the entire corpus of theory-of-mind research in nonhuman animals (3, 15, 16). Despite the development of numerous elegant paradigms, in most instances, participants have access to some behavioral cues that may permit them to correctly anticipate an actor's behavior without inferring the content of her mind.

Heyes (15) proposed an experimental design [inspired by Novey (17) and several other researchers], later known as the goggles test, to distinguish mind reading from alternatives, such as behavior reading. The key question is whether a participant can use her own past experience of visual access through novel goggles to determine whether others can see through the same goggles; in other words, whether she understands others not based on observable behavioral cues but rather by projecting her own mental state onto them. Meltzoff and Brooks (18) first successfully implemented this manipulation with young human children in a gaze-following test. More recently, Senju et al. (19) integrated this manipulation into an anticipatory-looking false-belief test (6): Two groups of 18-mo-old infants first experienced either an opaque blindfold or a trick blindfold that looked identical but could be seen through. While being eye tracked, both groups subsequently watched the same video sequence in which an actor saw an object be hidden in one location and then put on a blindfold before the object was removed. The children who had worn the opaque blindfold—but not those who had worn the trick blindfold—treated the actor as though she had a false belief: They looked in anticipation of her searching for the object in its previous location.

Chimpanzees have also been tested on several versions of the goggles task. They failed to preferentially beg for food from, or to gaze-follow, a human experimenter wearing a see-through visor or mask as compared with an experimenter wearing an opaque one (16, 20). However, in a food-competition test, chimpanzees correctly attempted to steal a competitor's food from behind an opaque screen as compared with a screen that appeared identical but that the chimpanzees had previously seen through (20). This last result suggested that apes can pass the goggles test and represent what others can see, at least, in certain contexts where apes show enhanced motivations to solve the task (21, 22).

The present study tested whether apes pass a goggles version of an anticipatory-looking test modeled on previous tests with apes and human infants (4, 6, 19). Apes first experienced seeing or not seeing through 1 of 2 identical novel barriers (a between- subject design). Then, while being eye tracked, they all watched the same video in which a human actor first saw an object hidden in one location. The actor then moved behind the novel barrier before the object was shifted and then removed. If apes can use self-experience to infer the actor's perspective, those who experienced the barrier as translucent should attribute a true belief to the actor, but those who experienced the barrier as opaque should instead attribute a false belief. Accordingly, we predicted that only the apes in the opaque condition should look in anticipation of the actor searching for the object in its original location. Critically, previous anticipatory-looking tasks involved 2 matched conditions, FB1 and FB2 (4, 6). In FB1, the actor watched an object be hidden in one location and moved to a second, and then was absent when the object was ultimately re- moved. In FB2, the actor watched the object be hidden in one location and then was absent when the object was moved to the second and ultimately removed. Although Senju et al. used the FB1 design of Southgate et al., we decided to instead use the FB2 design because recent attempts to replicate Southgate et al. with human populations found greater difficulty replicating the FB2 design (23–26). We, thus, thought that the FB2 design would constitute a more stringent test of action anticipation in great apes.

Methods

Participants. Forty-seven great apes (29 chimpanzees, Pan troglodytes, 14 bonobos, Pan paniscus, 4 orangutans, Pongo abelii from Kumamoto Sanctuary, Kumamoto, Japan, Primate Research Institute, Inuyama, Japan, and Leipzig Zoo, Leipzig, Germany) were randomly assigned to either the opaque (n = 23) or the see-through (n = 24) condition (between subjects). Two additional apes (1 chimpanzee and 1 orangutan) participated in the familiarization trial but refused to approach the test setup on the following day and, thus, could not be included in the test trial. We initially considered testing apes on a within-subject design and, consequently, all Kumamoto apes first experienced a test trial involving black barriers (translucent and opaque properties counterbalanced across individuals) in real life and in a video followed by a second trial with white barriers (on a separate day). However, we noticed a few problems with the black barrier (one of the chimpanzees was afraid of this barrier, and the translucent version was more difficult to see through than its white counterpart). Therefore, in proceeding with the experiment, we instead used white barriers only (in real life and in the video) for all other apes. Because we wanted to ensure that apes were responding to identical actions in the videos, we analyzed the data from the white barrier conditions alone (i.e., trial 2 for Kumamoto apes, trial 1 for all other apes). We confirmed that performance of the Kumamoto apes, who had experienced the additional black barrier, did not differ significantly from that of the other apes (see SI Appendix for this analysis). Most apes had participated in the original study (4) approximately 3 y earlier; however, false-belief trials always ended before the actor searched in either location and, thus, did not provide experience about how individuals with false beliefs behave. See SI Appendix, Table S1 for details about each participant.

Ethics Statement. All participants were tested in the testing rooms prepared for each species, and their daily participation in this study was voluntary. They received regular feedings, daily enrichment, and had ad libitum access to water. Animal husbandry and research protocol complied with international standards (the Weatherall report "The use of non-human primates in research") and institutional guidelines (Kumamoto Sanctuary: Wildlife Re- search Center "Guide for the Animal Research Ethics"; Primate Research Institute: Primate Research Institute 2002 version of "The Guidelines for the Care and Use of Laboratory Primates"; Wolfgang Koehler Primate Research Center: "EAZA [European Association of Zoos and Aquaria] Minimum Standards for the Accommodation and Care of Animals in Zoos and Aquaria," "WAZA [World Association of Zoos and Aquariums] Ethical Guidelines for the Conduct of Research on Animals by Zoos and Aquariums," "Guidelines for the Treatment of Animals in Behavioral Research and Teaching" of the Association for the Study of Animal Behavior).

Apparatus. Eye tracking was performed with apes following an established procedure (27). Apes' eyes were recorded by an infrared head-free eye tracker (60 Hz by X120 in Leipzig and 300 Hz by TX300 in Japan; Tobii Technology AB). The eye tracker and monitor were installed outside of the testing room. Apes were allowed to sip juice via a custom-made dispenser (irrespective of their gaze behavior) and to watch videos presented through a transparent panel. The videos were presented with a resolution of 1,280 × 720 pixels at a viewing distance of 70 cm on a 23-inch LCD monitor (43 × 24°) with Tobii Studio software (v3.2.1). Two-point automated calibration was conducted for each ape by presenting a small object or video clip on each reference point. We then assessed the quality of calibrations by visually checking the calibration results indicated in the software after the calibration and checking whether the ape's gaze point was not substantially deviated from a small reference icon presented before each recording session. We repeated the calibration procedure whenever necessary. Calibration errors in apes are typically within a degree, following these procedures (27).

Stimuli and Procedure. The procedure consisted of the presentation of a familiarization video, self-experience of an opaque or translucent barrier in real life, and the presentation of the test video. Videos largely followed the design of the FB2 condition of experiment 2 of Krupenye, Kano, and colleagues (4). As in that study, we maximized apes' engagement by embedding false-belief manipulations within agonistic conflicts between an ape-like character (Kong) and a human actor (Actor). On the first day, apes watched a pair of familiarization trials, identically designed to those used in Krupenye, Kano, and colleagues (4) (Exp. 2). In the first familiarization trial, Kong hid an object in 1 of 2 identical boxes in front of the Actor. The Actor attempted to obtain the object by reaching ambiguously toward the 2 boxes, finally overturning the correct box and grasping the object. The second familiarization was identical, except the object was hidden in, and retrieved from, the other box. The purpose of the familiarization trial was to demonstrate that the object could be hidden in either box, and that the Actor would search for it after witnessing its hiding. The familiarization trials were presented 1 day before the test trials, following the original study with apes (4), unlike studies involving human infants (6, 19). We made this decision to minimize the duration of video presented each day and thereby ensure apes' attention to the video throughout. A previous study confirmed that apes can remember video events across consecutive days (28).

After watching the familiarization video, we introduced apes to either an opaque or a see-through barrier in real life (on the same day for Leipzig and Inuyama apes but on the next day for Kumamoto apes) (Fig. 1 and SI Appendix, Fig. S2). Both barriers consisted of a wheeled frame (80 cm in width × 160 cm in height) covered with white fabric. For the see-through barrier, several layers of white mesh cloth were used. For the opaque barrier, a similar looking but opaque white meshlike cloth was used. All materials were novel to the apes. Both barriers were decorated with shiny red and green tinsel to be distinguishable from any other normal barrier. The see- through barrier was translucent and could be seen through at a close distance, but both barriers appeared identical from afar (e.g., the distance of the barrier in the test video). During the self-experience phase, an experimenter brought either barrier into the room and positioned it in front of the ape (~1 to 2 m from the ape) (Movie S1). The experimenter then ostensively presented various objects and food beside, and then behind, the barrier, demonstrating its occlusive property. In each facility, the same experimenter performed the same actions across the conditions. During the presentation of an object/food, the experimenter ensured 1) that the ape was attending to the object before moving it behind the barrier (and called the ape's name or changed the object if not) and 2) that the trajectories of the object and the experimenter's hand (and gaze) were identical across conditions. This self-experience phase lasted for ~5 min.

The test trial occurred on the following day (i.e., day 2 for Leipzig and Inuyama apes and day 3 for Kumamoto apes). On the test day, apes again experienced a 1- to 2-min self-experience reminder phase. The barrier was removed from the room and then, while being eye tracked, the apes watched the test video, which did not differ between conditions (Fig. 1; also see Movie S1). The test events did not differ from those in the Krupenye, Kano, and colleagues (4) study except that the actor's perceptual access was manipulated not by leaving through the door but by hiding behind the barrier. In the test video, the Actor first brought the barrier (which appeared opaque at this distance) into the scene. Kong then appeared and hid the object in 1 of the 2 boxes while the Actor watched (Fig. 1A). The Actor then scuttled behind the barrier with legs visible below, to ensure that apes tracked his presence (Fig. 1B). Kong then removed the object from the box (target), relocated it to the other box (distractor), retrieved it once more, and left the scene with it (Fig. 1C). The Actor then returned from behind the barrier and attempted to obtain the object by reaching ambiguously toward the 2 boxes (Fig. 1D). This middle reach lasted for 6 s. We counterbalanced across participants as evenly as possible the locations where the object was hidden in the familiarization (L then R vs. R then L) as well as the locations where the object was last hidden in the test (L or R), (i.e., 4 combinations: LRL, LRR, RLL, and RLR).

To assay apes' anticipations about where the Actor would search, during the 6 s of the Actor's ambiguous middle reach, we coded to which box each ape looked first as well as each ape's differential learning score (DLS; total viewing times to target minus total viewing times to distractor, divided by the sum of these values). First looks and viewing times were automatically calculated in Tobii Studio based on predefined areas of interest (SI Appendix, Fig. S1). The seminal study by Southgate et al. (6) used both first look and DLS measures, although more recent studies have focused primarily on DLS because this measure appears to show more consistent patterns (24, 29, 30). If apes can use their own past experience to determine an agent's mental state, we predicted that their patterns of anticipatory looking should differ between conditions; specifically, their looking should be more biased toward the target than the distractor (accordant with attribution of a false belief) in the opaque condition than in the see-through condition (where they should instead attribute a true belief), and they should show above- chance looking toward the target in the opaque condition only.

Results

Great apes did not differentiate between conditions in their first looks to target vs. distractor (Table 1) ($P = 0.46$, Fisher's exact test). However, as predicted, their DLS was significantly more biased toward the target in the opaque condition than in the see- through condition (Fig. 2) (Wilcoxon rank sum test, $Z = 2.13$, $n = 30$, $P = 0.033$, and $r = 0.39$). DLS was also significantly higher than chance (i.e., biased toward the target; Wilcoxon signed rank test, $Z = 2.35$, $n = 13$, $P = 0.019$, and $r = 0.65$) in the opaque condition but at chance level in the see-through condition ($Z = 0.80$, $n = 17$, $P = 0.42$, and $r = 0.19$).

Moreover, no significant difference was detected between species in either the opaque condition [Kruskal–Wallis rank sum test, $\chi2(2) = 0.81$, and $P = 0.66$] or the see-through condition [$\chi2(2) = 2.06$ and $P = 0.35$]. The opaque condition was akin to a replication of the FB2 condition of the Krupenye, Kano, and colleagues (4) study. To test this with an independent sample of apes, we reanalyzed the data from the opaque condition, excluding the 3 ape participants from this condition who had previously been tested in the FB2 condition of that study (i.e., all apes included in this analysis either had not participated in that study or had been tested in the FB1 condition). DLS remained significantly above chance (0.52 mean ± 0.66, 95% CI, $Z = 2.15$, $n = 11$, $P = 0.032$, and $r = 0.60$), consistent with the DLS data from the FB2 condition in Krupenye, Kano, and colleagues (4) (0.65 mean ± 0.38, 95% CI, $Z = 2.94$, $n = 12$, $P = 0.0033$, and $r = 0.82$).

Discussion

Our results suggest that great apes used their own past perceptual experiences to determine an agent's perceptual access and anticipate how the agent would behave. Apes who had experienced the barrier as opaque treated the actor as not having seen the object's movement and removal and as having a false belief that the object remained in its original location: They showed a bias in looking toward the location the object originally inhabited, consistent with anticipation that the actor would search for the object there. Apes who had experienced the barrier as translucent showed no such bias, consistent with attributing to the actor a true belief that the object had been removed and with having no expectation that the actor would search in any particular place.

Despite viewing identical familiarization and test events, apes' looking behavior differed between conditions. Critically, apes had never witnessed any other agents interacting with the novel barriers, precluding reliance on an abstraction or rule about how others behave in this novel context. Their differences in looking, which accord with attribution of true vs. false beliefs, therefore, derive specifically from their differential perceptual experience of these novel barriers. Our results, thus, support the theory-of-mind account.

Heyes (31) summarized several alternatives in response to Senju and colleague's (19) related findings with human infants. Heyes questioned whether "the opaque group did, and the trick group did not, discriminate between the boxes on the test be- cause the infants in the trick group were less distracted by the blindfold and, therefore, were more likely themselves to see the removal of the toy (object) from the scene during the belief induction trial." However, in this study, by measuring apes' looking times to the barrier and boxes at each event in the video sequence, we confirmed that apes' attention to the object dis- placement events did not differ between conditions—nor did their attention to the barrier at any particular event (SI Appendix, Fig. S3). Heyes also proposed that human infants might have perceived similarities between the cloth materials used for the blindfolds and the common cloth materials used in households. This concern does not apply to our study because we selected cloth materials that were not usually seen in our facilities. Finally, Heyes questioned whether participants could perceive any differences in actions performed by the experimenter during familiarization of the opaque and see-through barriers. To pre- vent this possibility, we ensured that the same experimenters performed the same actions across conditions.

Although the goggles test is seen as a method for distinguishing the theory-of-mind account from nonmentalistic alternatives (15), the paradigm has received some philosophical scrutiny (e.g., refs. 12, 32, and 33). Specifically, the theory-of-mind ac- count proposes that, in this paradigm, participants experience seeing or not seeing through a barrier as an unobservable mental state and attribute that same inner experience to the agent behind the barrier. Nonmentalistic accounts generally conceive of seeing in behavioral terms: as there being an unobstructed line of sight between an agent and an object (32). In principle, participants could solve the goggles task in a nonmentalistic way if, in the self- experience phase, they are simply learning whether the barrier obstructs a line of sight and, in the test, applying existing rules or abstractions about how agents usually behave around barriers that do or do not obstruct the line of sight (12). Thus, the behavior rule previously applied to other false belief tests—"agents tend to search for things where they last saw them" (13) could be modified as "agents tend to search for things in the last location where they established a line of sight unobstructed by a learned property of a barrier." At a minimum, however, as Meltzoff and Brooks (18) argued, participants are "learning a psychological affordance, not a simple physical or motor affordance—and, moreover, are learning how it affects the self and applying it to others." Importantly, the original behavior-rule account—that agents search for things where they last saw them (13)—cannot be applied in any straightforward manner to this or Senju et al.'s (19) study because, across conditions, great apes and infants generate different anticipations in response to an identical set of actions.

If our current results are not explained by behavior rules, do they constitute evidence of full-blown false belief understanding? Apperly and Butterfill (34, 35) have argued for one prominent alternative: That humans are endowed with a minimal mind reading system from infancy, allowing for efficient but inflexible anticipation of behavior through tracking of belief-like states, and that human adults have a second cognitively effortful system for flexibly representing propositional attitudes. The minimal system can represent encountering (i.e., that an object has been within an agent's visual field) and registering (i.e., it can represent the last place the agent encountered the object, even if the object has been moved after the agent departed the scene). Minimal mind readers can, therefore, track belief-like states that are akin to beliefs about an object's location. Accordingly, they can accurately anticipate an agent's action in change-of-location false-belief tasks, the paradigms most commonly used in research with human infants and nonhuman animals. Critically, the signature limit of such a minimal system is its inability to track genuine beliefs about object identity (36, 37). Testing this limitation in the future will be important for precisely specifying the representational mechanisms underlying action anticipation in both human and nonhuman primates.

Our results derive from the DLS rather than apes' first looks (where no clear effects were obtained). This mosaic of results is consistent with several other anticipatory-looking studies with human participants in which DLS also provided more consistent findings than first looks and may emerge from common elements of many anticipatory-looking paradigms (29, 30). DLS results derive from a longer time window (6 s in this study), presumably providing more time for participants to generate their anticipations and potentially making them more robust to noise. Because looking can be sensitive to perceptual and cognitive influences beyond action anticipation, it is likely that the strength of first look vs. DLS measures will vary between individual paradigms. Importantly, our DLS results are consistent with similar work in 18-mo-old humans (19), and they replicate previous findings with apes (4).

Recent replication attempts for the Southgate–Senju paradigm were successful overall in replicating the FB1 design but not the FB2 design with both human adults and infants (for a summary and commentary, see ref. 24). It is, therefore, important to highlight that this study replicated our previous DLS findings in nonhuman apes, specifically from the FB2 design (4). Although it remains unclear which factors impact replicability with human participants, researchers have proposed that several procedural differences between studies may be responsible (24). Of special note, one interesting result from previous replication studies is that, while some researchers found a chance-level effect for FB2 (26), others found a below-chance effect for FB2 (as well as an above-chance effect for FB1) in the DLS results (23, 25). One interpretation for this puzzling pattern is that human participants showed a location bias: they may have simply looked at the last location that the object inhabited (before it was finally re- moved). Interestingly, in previous anticipatory-looking studies that adopted similar designs for great apes and human infants, human infants anticipatorily looked at the location where a target object was before its relocation (in an inanimate control condition ref. 38), while apes did not show the same position bias (39). The reason for this potential species difference is unknown but could be related to a species difference in general eye- movement properties; specifically, that the ape eye is less likely than the human eye to be locked in a certain fixation lo- cation (27). Most relevant here, replication of FB2 with non- human apes supports the general validity of the Southgate–Senju design, at least, for certain populations/species. However, to in- crease the utility of this paradigm, future studies should identify which aspects of the procedure are likely to cause replication failures (or position biases) in human participants. In addition, as Southgate and colleagues (24) pointed out, it is likely critical that researchers ensure that participants are "highly engaged by the agent's actions so that they are entirely focused on predicting what she will do next" in an anticipatory-looking test. As we have shown here, the use of dramatic stories optimized for the target age/ species may be one way to achieve more reliable gaze-based measures of action anticipation (40).

In conclusion, we provide evidence that great apes can use their own past experience of visual access to attribute perception and, potentially, resultant beliefs to others. Our results are consistent with findings from 18-mo-old humans (19). Notably, they are also consistent with findings from corvids showing that they use their own visual and pilfering experiences to guard their caches against pilferers in a food-competition task (41, 42). These results together corroborate the idea that young infants and nonhuman animals have a theory of mind and do not simply rely on behavior rules to interpret and anticipate others' actions. They also highlight the important role that self-experience may play in expanding the states of mind that infants and animals can attribute to others (43). However, additional tests are necessary to further rule out nonmentalistic alternatives. Particularly, in the case of this study, rather than attributing their own inner experience to the actor, it is possible that apes learned about the psychological affordances of the barriers and used this information to determine whether the actor had an unobstructed line of sight (but see ref. 44 for recent evidence against this possibility in a different research paradigm). Moreover, to determine whether apes represent agents' beliefs or belief-like states, it will be necessary to investigate their ability to anticipate an agent's action based on false beliefs about object identity. At a minimum, the present study has shown that, in the absence of any differential behavioral cues during the test, apes make different anticipations of an agent's actions depending on their own past experiences of perceptual access. Our findings, therefore, contest the straightforward conception of the behavior-rule ac- count.

Acknowledgments

We thank Naruki Morimura, Yutaro Sato, Yuri Kawaguchi, Daniel Hanus, Hanna Petschauer, and the staff at Kumamoto Sanctuary, Primate Research Institute, and Wolfgang Koehler Primate Research Center for their assistance in conducting the experiments. Financial support came from the Japan Society for the Promotion of Science KAKENHI Grants 18H05072, 19H01772, 16H06301 (to F.K.), 18H05524 (to S.H.), 16H06283, LDG-U04, GAIN (to M.T.); European Commission Marie Skłodowska-Curie Fellowship MENTALIZINGORIGINS (to C.K.); and European Research Council Synergy Grant 609819 SOMICS (to J.C.).

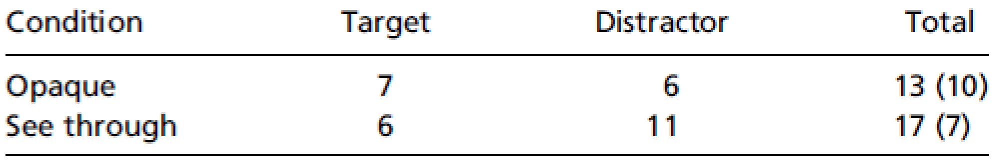
References

1. D. Premack, G. Woodruff, Does the chimpanzee have a theory of mind. *Behav. Brain Sci.* 1, 515–526 (1978).
2. J. Call, M. Tomasello, Does the chimpanzee have a theory of mind? 30 years later.*Trends Cogn. Sci.* 12, 187–192 (2008).
3. C. Krupenye, J. Call, Theory of mind in animals: Current and future directions. *Wiley Interdiscip. Rev. Cogn. Sci.*, 10.1002/wcs.1503 (2019).
4. C. Krupenye, F. Kano, S. Hirata, J. Call, M. Tomasello, Great apes anticipate that other individuals will act according to false beliefs. *Science* 354, 110–114 (2016).
5. D. Buttelmann, F. Buttelmann, M. Carpenter, J. Call, M. Tomasello, Great apes distinguish true from false beliefs in an interactive helping task. *PLoS One* 12, e0173793 (2017).

6. V. Southgate, A. Senju, G. Csibra, Action anticipation through attribution of false belief by 2-year-olds. *Psychol. Sci.* 18, 587–592 (2007).
7. K. H. Onishi, R. Baillargeon, Do 15-month-old infants understand false beliefs? *Science* 308, 255–258 (2005).
8. D. Buttelmann, M. Carpenter, M. Tomasello, Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112, 337–342 (2009).
9. C. Heyes, Apes submentalise. *Trends Cogn. Sci.* 21, 1–2 (2017).
10. F. Kano, C. Krupenye, S. Hirata, J. Call, M. Tomasello, Submentalizing cannot explain belief-based action anticipation in apes. *Trends Cogn. Sci.* 21, 633–634 (2017).
11. C. Krupenye, F. Kano, S. Hirata, J. Call, M. Tomasello, A test of the submentalizing hypothesis: Apes’ performance in a false belief task inanimate control. *Commun. Integr. Biol.* 10, e1343771 (2017).
12. D. Scarf, T. Ruffman, Great apes’ insight into the mind: How great? [Online commentary]. https://science.sciencemag.org/content/354/6308/110/tab-e-letters. Accessed 31 January 2017.
13. J. Perner, T. Ruffman, Psychology. Infants’ insight into the mind: How deep? *Science* 308, 214–216 (2005).
14. R. Baillargeon, R. M. Scott, Z. He, False-belief understanding in infants. *Trends Cogn. Sci.* 14, 110–118 (2010).
15. C. M. Heyes, Theory of mind in nonhuman primates. *Behav. Brain Sci.*21, 101–114, *discussion in Behav. Brain Sci.* 21, 115–148 (1998).
16. D. C. Penn, D. J. Povinelli, On the lack of evidence that non-human animals possess anything remotely resembling a ‘theory of mind’. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 731–744 (2007).
17. M. S. Novey, “The development of knowledge of others’ ability to see,” PhD dissertation, Harvard University, Boston, MA (1975).
18. A. N. Meltzoff, R. Brooks, Self-experience as a mechanism for learning about others: A training study in social cognition. *Dev. Psychol.* 44, 1257–1265 (2008).
19. A. Senju, V. Southgate, C. Snape, M. Leonard, G. Csibra, Do 18-month-olds really attribute mental states to others? A critical test. *Psychol. Sci.* 22, 878–880 (2011).
20. K. Karg, M. Schmelz, J. Call, M. Tomasello, The goggles experiment: Can chimpanzees use self-experience to infer what a competitor can see? *Anim. Behav.* 105, 211–221 (2015).
21. B. Hare, J. Call, B. Agnetta, M. Tomasello, Chimpanzees know what conspecifics do and do not see. *Anim. Behav.* 59, 771–785 (2000).
22. A. P. Melis, J. Call, M. Tomasello, Chimpanzees (*Pan troglodytes*) conceal visual and auditory information from others. *J. Comp. Psychol.* 120, 154–162 (2006).
23. L. Kulke, B. von Duhn, D. Schneider, H. Rakoczy, Is implicit theory of mind a real and robust phenomenon? Results from a systematic replication study. *Psychol. Sci.* 29, 888–900 (2018).
24. R. Baillargeon, D. Buttelmann, V. Southgate, Invited commentary: Interpreting failed replications of early false-belief findings: Methodological and theoretical consider- ations. *Cogn. Dev.* 46, 112–124 (2018).
25. S. Dörrenberg, H. Rakoczy, U. Liszkowski, How (not) to measure infant theory of mind: Testing the replicability and validity of four non-verbal measures. *Cogn. Dev.* 46, 12–30 (2018).
26. C. Grosse Wiesmann, A. D. Friederici, D. Disla, N. Steinbeis, T. Singer, Longitudinal evidence for 4-year-olds’ but not 2- and 3-year-olds’ false belief-related action anticipation. *Cogn. Dev.* 46, 58–68 (2018).
27. F. Kano, S. Hirata, J. Call, M. Tomonaga, The visual strategy specific to humans among hominids: A study using the gap-overlap paradigm. *Vision Res.* 51, 2348– 2355 (2011).
28. F. Kano, S. Hirata, Great apes make anticipatory looks based on long-term memory of single events. *Curr. Biol.* 25, 2513–2517 (2015).
29. A. Senju *et al.*, Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Dev. Psychopathol.* 22, 353–360 (2010).
30. A. Senju, V. Southgate, S. White, U. Frith, Mindblind eyes: An absence of spontaneous theory of mind in asperger syndrome. *Science* 325, 883–885 (2009).

31. C. Heyes, False belief in infancy: A fresh look. *Dev. Sci.* 17, 647–659 (2014).
32. R. Lurz, If chimpanzees are mindreaders, could behavioral science tell? Toward a solution of the logical problem. *Philos. Psychol.* 22, 305–328 (2009).
33. G. Csibra, Seeing is not believing. *Behav. Brain Sci.* 21, 117–118 (1998).
34. I. A. Apperly, S. A. Butterfill, Do humans have two systems to track beliefs and belief- like states? *Psychol. Rev.* 116, 953–970 (2009).
35. S. A. Butterfill, I. A. Apperly, How to construct a minimal theory of mind. *Mind Lang.* 28, 606–637 (2013).
36. J. Low, J. Watts, Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychol. Sci.* 24, 305–311 (2013).
37. E. Fizke, S. Butterfill, L. van de Loo, E. Reindl, H. Rakoczy, Are there signature limits in early theory of mind? *J. Exp. Child Psychol.* 162, 209–224 (2017).
38. E. N. Cannon, A. L. Woodward, Infants generate goal-based action predictions. *Dev. Sci.* 15, 292–298 (2012).
39. F. Kano, J. Call, Great apes generate goal-based action predictions: An eye-tracking study. *Psychol. Sci.* 25, 1691–1698 (2014).
40. F. Kano, C. Krupenye, S. Hirata, J. Call, Eye tracking uncovered great apes' ability to anticipate that other individuals will act according to false beliefs. *Commun. Integr. Biol.* 10, e1299836 (2017).
41. N. J. Emery, N. S. Clayton, Effects of experience and social context on prospective caching strategies by scrub jays. *Nature* 414, 443–446 (2001).
42. T. Bugnyar, S. A. Reber, C. Buckner, Ravens attribute visual access to unseen com- petitors. *Nat. Commun.* 7, 10506 (2016).
43. A. N. Meltzoff, 'Like me': A foundation for social cognition. *Dev. Sci.* 10, 126–134 (2007).
44. R. Lurz, C. Krachun, L. Mahovetz, M. J. G. Wilson, W. Hopkins, Chimpanzees gesture to humans in mirrors: Using reflection to dissociate seeing from line of gaze. *Anim. Behav.* 135, 239–249 (2018).

Table 1. Number of participants who made first looks to either the target or the distractor during the Actor's middle reach



Condition	Target	Distractor	Total
Opaque	7	6	13 (10)
See through	6	11	17 (7)

Values in parentheses indicate the number of participants who did not look at either.

Fig. 1. The test video sequence used in this study (Movie S1). (A) The Actor watches Kong hide an object in the left (target) box. (B) The Actor scuttles behind the barrier. (C) Kong moves object to the right (distractor) box and then removes object. (D) The Actor reaches ambiguously toward

boxes. If apes, who experienced the barrier as opaque, attributed no visual access (and a false belief that the object remained in the target location) to the Actor, their looking in anticipation of the Actor's search should be biased toward the target box. Conversely, if those, who experienced the barrier as see through, attributed a true belief to the Actor (that the object had been removed), their looking should be at chance, consistent with no prediction.

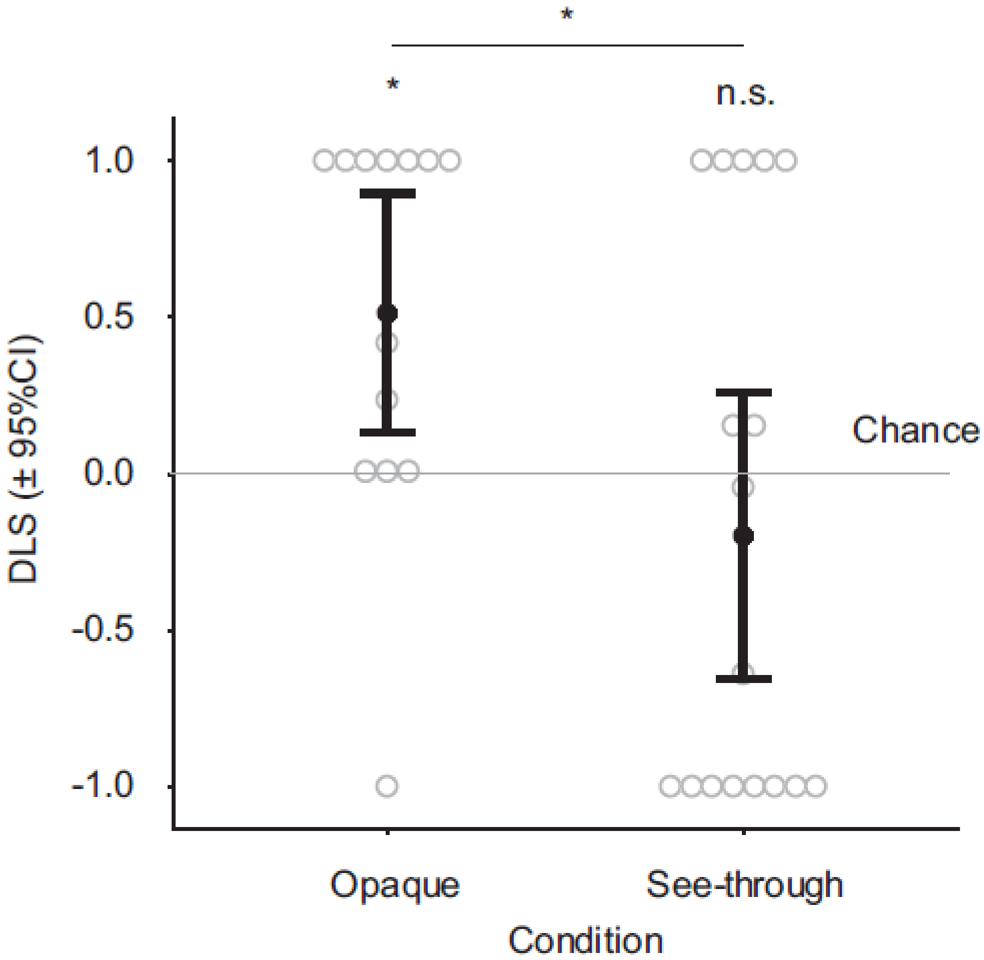


Fig. 2. DLS (±95% CIs; calculated as looking to the target minus looking to the distractor divided by the sum of these values; i.e., positive DLS indicates looking bias toward the target) as the Actor reached ambiguously toward the target and distractor boxes in the opaque and see-through conditions. The target is the location where the Actor last saw the object before moving behind the barrier. During the ambiguous reach in the opaque and see- through conditions, respectively, 13 and 17 apes looked at the boxes, while the rest did not look at either (i.e., no DLS data). The chance level is 0. The dots indicate participants' individual data. *$P < 0.05$ in nonparametric tests.