Jennifer Nagel, University of Toronto
June 5, 2023

Lecture 3: Mental state recognition among animals

1. Agent-focused and world-focused mindreading
When your environment contains another agent (the 'target agent'),

- You face a learning challenge because the target agent's behavior is hard to predict* (a challenge solved by *agent-focused mindreading*).
- You gain a learning bonus because the target agent's intelligent actions serve as cues to reality (a bonus gained through *world-focused mindreading*).

*Agents make exceptions to the rules of folk physics by means of which we anticipate the trajectories of everything else: agents *act* as a function of their epistemic and motivational states. The motions of a floating log are immediately dictated by the mechanical forces on it, while mental states produce action in a way that plays out over variable time scales (for discussion of this point, see Hohwy and Palmer 2014).

Agent-focused mindreading: the unexpected transfer false belief task (Wimmer and Perner 1983).
World-focused mindreading: the container-swap task (Krachun, Carpenter et al. 2009).

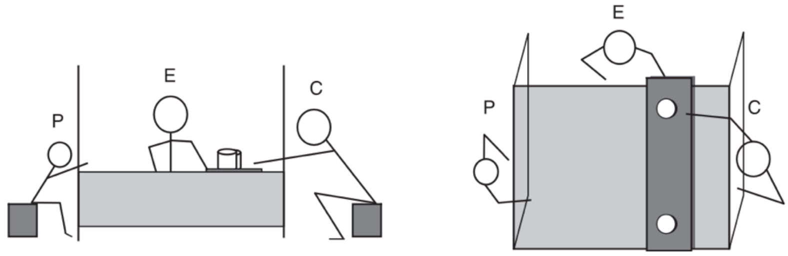


Figure 1 *Experimental set-up: front view on left and top view on right. P = Participant, E = Experimenter, C = Competitor.*

2. Varieties of gaze tracking
Direct gaze (eye contact between agents) is necessarily mutual, and a powerful signal of engagement or threat (for a review, see Shepherd 2010). Simply being in this symmetrical relation is enough to trigger relevant innate and learned responses; the animals involved do not also need to represent the symmetry of the relation.
Deictic gaze (gaze directed toward another object) supplies information both about the target animal and about the object that this target animal is looking at. In deictic gaze tracking, animals look towards the target of another creature's object-directed gaze, first to objects within the shared visual field, and then to objects beyond, for example when a creature facing you evidently sees something happening behind your back (if you turn to look, this is an instance of 'geometrical gaze tracking'). Many species seem to show some capacity for geometrical gaze tracking, including all kinds of primates (Tomasello, Call et al. 1998), bottlenose dolphins (Johnson, Ruiz-Mendoza et al. 2022), goats (Kaminski, Riedel et al. 2005), and even lizards (Simpson and O'Hara 2019).

Curious animals can learn to follow gaze through ordinary reinforcement learning (Deák, Triesch et al. 2013). Physical and social environments make a difference to how well animals track gaze. The open cliff-dwelling Northern Bald Ibis follows gaze into distant space, but not around barriers (Loretto, Schloegl et al. 2010); ravens, who live in dense forests, understand visual barriers by six months (Bugnyar, Stöwe et al. 2004). In general, natural selection and reinforcement learning make animals as smart as they need to be, for their environment.

A) B)

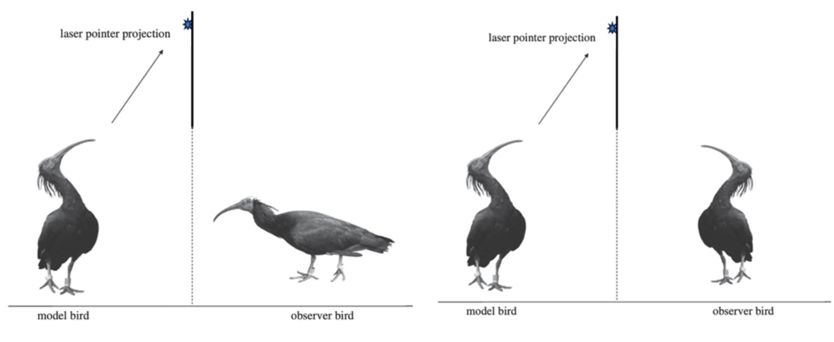


Diagram adapted from: Loretto et al., p.15.

Corvids (ravens, crows, jays) have superior gaze tracking, in part because they switch between roles of caching food for later consumption, as privately as possible, and raiding the caches of other corvids they have spied on. Ravens figure out that they can be seen through tiny peepholes they themselves had spied through earlier, even when they are themselves far enough away from those peepholes that they couldn't see anything through them, and even when they have never seen the face of another bird through such a peephole (Bugnyar, Reber et al. 2016).

In learning about seeing, animals gain some general lessons about what points positively fall within another creature's field of view, and also some general lessons about what points are excluded.

Infant apes follow gaze to objects in the shared visual field (Okamoto, Tanaka et al. 2004), and 2-3-year-olds follow gaze geometrically. By the age of four, most apes understand more complex barriers with windows (Okamoto-Barth, Call et al. 2007), and by age five, apes will look back at someone who is staring at nothing, as if to double-check the accuracy of their geometrical gaze following (Bräuer, Call et al. 2005).

3. Tracking knowledge and ignorance

Animal behavior depends not only what is in its present sensory scope, but what it has sensed over time. Social animals can track what others know about aspects of reality that are not presently within their sensory scopes (for example, which closed box now contains a fruit).

Scrub jays who cache worms while being watched will tend to relocate those worms to new hiding spots later, much more than jays who cached in privacy (Emery and Clayton 2001). Ravens keep track of which particular bird watched them cache (Bugnyar and Heinrich 2005).

Eurasian jays seem unable to exploit the false belief of a would-be pilferer in a competitive caching paradigm (Brecht 2017, chapter 2).

Some researchers believe that corvids are capable of episodic memory (or 'mental time travel') because they choose peanuts over worms if they cached the worms five days ago, but worms over peanuts if they cached the worms 4 hours ago (Clayton and Dickinson 1998). Christoph Hoerl and Teresa McCormack argue that the world models of nonhuman animals only ever represent how things are right now, and never how things were, will be, or might be. Perishables fade out from the map of the world as it now is, thanks to an interval timing mechanism: there is 'a change of representations, but no representation of change', as they put it (Hoerl and McCormack 2019, p.2). These changes of representations can be learned from experience, and can constitute knowledge of the relative permanence of things. In this theory, only humans have a full-blown temporal reasoning system with the potential for mental time travel (and we have that system alongside our more primitive temporal updating system).

Nonhuman animal mentalizing seems to be limited to a model of current reality (which can include hidden and remote contents, but nothing in the 'decoupled' realm of the past, future, or hypothetical).

There is no robust evidence that any nonhuman animal can past tests of false belief attribution. Not dolphins (Hill, Dietrich et al. 2018), not apes in Buttelmann-style (2017) helping paradigms (Priewasser, Rafetseder et al. 2018), not rhesus monkeys in tests of 'implicit false belief' (Marticorena, Ruiz et al. 2011), not chimpanzees, despite (Krupenye, Kano et al. 2016), now disavowed by its last author (Tomasello 2018), and further undermined by larger worries about the anticipatory looking paradigm (Wang and Leslie 2016, Barone, Corradi et al. 2019).

But even if nonhuman animals can't represent false beliefs, could they still be representing true beliefs (as opposed to knowledge)?

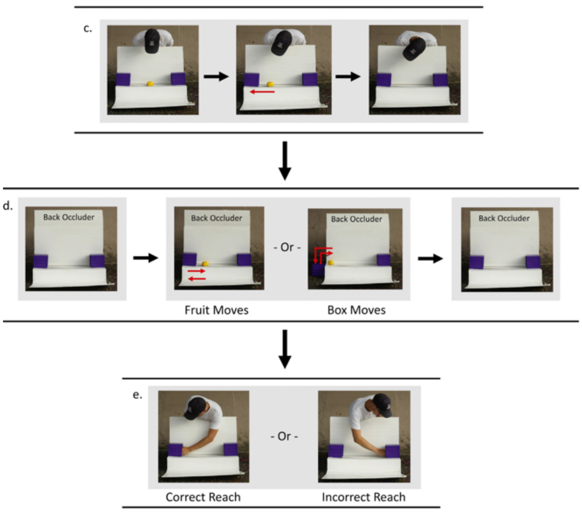
4. Gettier case recognition in nonhuman animals

Background: chimpanzees compete for food subject to a strict dominance hierarchy. If a subordinate sees food being hidden in the presence of a watching dominant, she will not challenge him for it, but if his view is occluded, or if he is taken away and a new dominant is brought in, she will go for it (Hare, Call et al. 2001). Hare and colleagues took this as a sign that chimpanzees distinguish knowledgeable and ignorant competitors, but philosophers might worry that the subordinate just saw the watching dominant as having a true belief, rather than knowledge.

I think the first experimental demonstration of chimpanzees recognizing a Gettiered agent as ignorant is (Kaminski, Call et al. 2008); however, these researchers misdescribe their target (true belief) agent as knowing, and are confused that the focal chimpanzee does not see it this way. This paper also uses a somewhat complex gambling paradigm.

Laurie Santos's lab has produced a clearer demonstration of primate Gettier case recognition, but they see their results as supporting the thesis that primates think in terms of 'awareness relations' as opposed to knowledge and ignorance (Martin and Santos 2016, Horschler, Santos et al. 2019).

HS&M: "Awareness relations are mentalistic representations but are fundamentally different than knowledge- or belief-based representations; awareness relations are hypothesized to have an "on/off" quality such that primates either represent an awareness relation linking a (true) piece of information to the agent or fail to represent any relation between the agent and the information whatsoever. The awareness relations hypothesis predicts that if the information linked to an agent changes while outside that agent's awareness, the awareness relation is 'turned off', regardless of whether this event results in a mismatch between reality and the agent's belief. In this way, primates tend to make no predictions about how an agent will behave when he or she is ignorant or holds a false belief." (Horschler, Santos et al. 2019, 73)



When the lemon moves (while unobserved), monkeys no longer expect an accurate reach.

Brute fact: "awareness relations are disrupted by physical movement of objects being tracked by an agent when this movement occurs outside the agent's awareness" (Horschler, Santos et al. 2019, 79).

HS&M: "A subject who represented these events in terms of the agent's knowledge/ignorance should have expected the agent to reach correctly regardless of whether the fruit or the box moved when the agent could not see, as in both cases the agent maintained knowledge of the object's ultimate location. Instead, monkeys' conceptions of the agent's knowledge of the fruit's location appear to be disrupted, or rendered null, when the target object moves while out of the agent's view." (2019, 78)

Santos elsewhere characterizes ignorance as —"a relation between an agent and a piece of information that is not part of our current reality" (Martin and Santos 2016, 376).

Monkeys are performing poorly on this task if they are compared with human mindreaders who have the concept of belief, as well as the concept of knowledge. However, if we evaluate them strictly on the question of whether they discriminate between states of knowledge and ignorance, their performance looks better: they are right to react differently to these cases. In general, the object-seeking behavior of knowledgeable and ignorant agents does differ, even if this difference is absent in some Gettier cases.

5. Representing and exploiting the epistemic powers of others

Agent-focused mindreading aims simply to predict behavior, where this can be driven by factors like false belief, independent from current reality; world-focused mindreading aims to take the actions of sensitive and smart agents as cues to reality. We need to learn generalizations about when to heed those cues (when target agents are knowledgeable) and when to drop them (when target agents are ignorant). Ignorant agents can be right by chance (for example in Gettier cases), but systematic social learning demands generalization. Generalizations about knowledge are easier to learn because animals are sharply limited in the ways they can come to gain knowledge; by contrast, states of belief can be entered in many ways.

Assuming the motivational state of the agent is grasped, then because knowledge is factive, mindreaders who can spot knowledge in other animals can infer the state of the world from those animals' actions: social cognition allows animals to help ourselves to the knowledge of others in making sense of reality.

References:

Barone, P., G. Corradi and A. Gomila (2019). "Infants' performance in spontaneous-response false belief tasks: A review and meta-analysis." Infant Behavior and Development **57**: 101350.
Bräuer, J., J. Call and M. Tomasello (2005). "All great ape species follow gaze to distant locations and around barriers." Journal of Comparative Psychology **119**(2): 145.
Brecht, K. F. (2017). A multi-facetted approach to investigating theory of mind in corvids, University of Cambridge.

Bugnyar, T. and B. Heinrich (2005). "Ravens, Corvus corax, differentiate between knowledgeable and ignorant competitors." Proceedings of the Royal Society B: Biological Sciences **272**(1573): 1641-1646.
Bugnyar, T., S. A. Reber and C. Buckner (2016). "Ravens attribute visual access to unseen competitors." Nature Communications **7**(1): 1-6.
Bugnyar, T., M. Stöwe and B. Heinrich (2004). "Ravens, Corvus corax, follow gaze direction of humans around obstacles." Proceedings of the Royal Society of London. Series B: Biological Sciences **271**(1546): 1331-1336.
Buttelmann, D., F. Buttelmann, M. Carpenter, J. Call and M. Tomasello (2017). "Great apes distinguish true from false beliefs in an interactive helping task." PLoS One **12**(4): e0173793.
Clayton, N. S. and A. Dickinson (1998). "Episodic-like memory during cache recovery by scrub jays." Nature **395**(6699): 272-274.
Deák, G. O., J. Triesch, A. Krasno, K. de Barbaro and M. Robledo (2013). Learning to share: The emergence of joint attention in human infancy. Cognition and brain development: Converging evidence from various methodologies. B. R. Kar. Washington, D.C., American Psychological Association**:** 173-210.
Emery, N. J. and N. S. Clayton (2001). "Effects of experience and social context on prospective caching strategies by scrub jays." Nature **414**(6862): 443-446.
Hare, B., J. Call and M. Tomasello (2001). "Do chimpanzees know what conspecifics know?" Animal Behaviour **61**(1): 139-151.
Hill, H. M., S. Dietrich, A. Cadena, J. Raymond and K. Cheves (2018). "More than a fluke: Lessons learned from a failure to replicate the false belief task in dolphins." International Journal of Comparative Psychology **31**.
Hoerl, C. and T. McCormack (2019). "Thinking in and about time: A dual systems perspective on temporal cognition." Behavioral and Brain Sciences **42**: e244.
Hohwy, J. and C. Palmer (2014). Social cognition as causal inference: implications for common knowledge and autism. Perspectives on social ontology and social cognition, Springer**:** 167-189.
Horschler, D. J., L. R. Santos and E. L. MacLean (2019). "Do non-human primates really represent others' ignorance? A test of the awareness relations hypothesis." Cognition **190**: 72-80.
Johnson, C. M., C. Ruiz-Mendoza and C. Schoenbeck (2022). "Conspecific" gaze following" in bottlenose dolphins." Animal Cognition **25**(5): 1219-1229.
Kaminski, J., J. Call and M. Tomasello (2008). "Chimpanzees know what others know, but not what they believe." Cognition **109**(2): 224-234.
Kaminski, J., J. Riedel, J. Call and M. Tomasello (2005). "Domestic goats, Capra hircus, follow gaze direction and use social cues in an object choice task." Animal behaviour **69**(1): 11-18.
Krachun, C., M. Carpenter, J. Call and M. Tomasello (2009). "A competitive nonverbal false belief task for children and apes." Developmental Science **12**(4): 521-535.
Krupenye, C., F. Kano, S. Hirata, J. Call and M. Tomasello (2016). "Great apes anticipate that other individuals will act according to false beliefs." Science **354**(6308): 110-114.
Loretto, M.-C., C. Schloegl and T. Bugnyar (2010). "Northern bald ibises follow others' gaze into distant space but not behind barriers." Biology letters **6**(1): 14-17.
Marticorena, D. C. W., A. M. Ruiz, C. Mukerji, A. Goddu and L. R. Santos (2011). "Monkeys represent others' knowledge but not their beliefs." Developmental Science **14**(6): 1406-1416.
Martin, A. and L. R. Santos (2016). "What cognitive representations support primate theory of mind?" Trends in cognitive sciences **20**(5): 375-382.
Okamoto, S., M. Tanaka and M. Tomonaga (2004). Looking back: The "representational mechanism" of joint attention in an infant chimpanzee (Pan troglodytes) 1, Wiley Online Library. **46:** 236-245.
Okamoto-Barth, S., J. Call and M. Tomasello (2007). "Great apes' understanding of other individuals' line of sight." Psychological science **18**(5): 462-468.
Priewasser, B., E. Rafetseder, C. Gargitter and J. Perner (2018). "Helping as an early indicator of a theory of mind: Mentalism or Teleology?" Cognitive Development **46**: 69-78.
Shepherd, S. V. (2010). "Following gaze: gaze-following behavior as a window into social cognition." Frontiers in integrative neuroscience **4**: 5.
Simpson, J. and S. J. O'Hara (2019). "Gaze following in an asocial reptile (Eublepharis macularius)." Animal cognition **22**(2): 145-152.
Tomasello, M. (2018). "How children come to understand false beliefs: A shared intentionality account." Proceedings of the National Academy of Sciences **115**(34): 8491-8498.
Tomasello, M., J. Call and B. Hare (1998). "Five primate species follow the visual gaze of conspecifics." Animal behaviour **55**(4): 1063-1069.
Wang, L. and A. M. Leslie (2016). "Is implicit theory of mind the 'Real Deal'? The own-belief/true-belief default in adults and young preschoolers." Mind & Language **31**(2): 147-176.
Wimmer, H. and J. Perner (1983). "Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception." Cognition **13**(1): 103-128.