Animal Behavior and Cognition
ABC 2020, 7(4):653-657
DOI:https://doi.org/10.26451/abc.07.04.10.2020
©Attribution 3.0 Unported (CC BY 3.0)

Perspective Piece 1

Deconstructing the Raven's Theory of Mind: An Analysis of Bugnyar et al. (2016)

Shannon L. Kuznar[1], Mateja Pavlic[2], Gabrielle C. Glorioso[2], and Daniel J. Povinelli[1, *]

[1]Department of Biology, University of Louisiana
[2]Department of Psychology, University of Louisiana

[*]Corresponding author (Email: povinelli@louisiana.edu)

Citation – Kuznar, S.L., Pavlic, M., Glorioso, G.C., & Povinelli, D.J. (2020). Deconstructing the raven's theory of mind: An Analysis of Bugnyar et al. (2016). *Animal Behavior and Cognition, 7*(4), 653-657. https://doi.org/10.26451/abc.07.04.10.2020

Over four decades ago, Premack and Woodruff (1978) proposed a seemingly simple question, do chimpanzees possess a theory of mind? They argued that organisms possess a theory of mind if they are able to make inferences about the unobservable mental states of others, and to use such representations to generate predictions concerning their behaviors. Some have claimed the genre of available tests possess a logical flaw, so that even in principle they cannot provide an affirmative answer to this question (see Povinelli & Vonk, 2004). In response, others have argued that these tests can, in fact, provide strong experimental evidence for theory of mind in nonhuman animals (Bräuer et al., 2007; Emery & Clayton, 2001; Hare et al., 2000). Still others believe that despite logical flaws in *current* experimental designs, more diagnostic tests can be developed (Lurz & Krachun, 2011; Lurz et al., 2014; but see Lurz et al., 2018).

Bugnyar et al. (2016) reported what they describe as "strong evidence" that ravens create mental representations of the mental states of other ravens, in particular, that they represent the mental state of <*seeing*>. These researchers side with those who believe that previous experiments have suffered from a common problem: the inability to distinguish between animals relying solely on first-order, perceptual-based relational reasoning, versus additionally relying upon higher-order mental states such as <*perceiving*>, <*intending*>, <*knowing*>, <*believing*>, etc. In the case of <*seeing*>, in particular, they argue that previous tests have failed to isolate any cognitive work of higher-order representations above and beyond the causal role played by the (undisputed) ability of animals to represent agents, resources, and opaque/transparent obstructions and their geometric relations. For example, a subordinate chimpanzee may be unwilling to reach for food when they are in an unobstructed geometric relation to (or the "line of sight" of) another dominant individual, but yet be perfectly willing to do so if the dominant is located behind an opaque barrier. Thus, Bugnyar et al. (2016) agree with Povinelli and Vonk (2004) that this general problem invalidates several widely cited previous studies (Bräuer et al., 2007; Emery & Clayton, 2001; Hare et al., 2000). Thus, the value of Bugnyar et al.'s (2016) experimental procedure lies in its ability or inability to escape this limitation. Once Bugnyar et al.'s (2016) procedures and arguments are formally specified, however, it becomes evident that they suffer from the same limitation they sought to overcome.

We begin by emphasizing that we agree with Bugnyar et al.'s (2016) critical analysis of previous experimental designs. Indeed, the problem identified above arises from three related challenges facing researchers who attempt to experimentally distinguish between the causal power of first-order geometric relations among objects and agents versus any additional causal work that might be performed by the representation of mental states *qua* mental states:

First-order perceptual representations persist over time. As Bugnyar et al. (2016) acknowledge, animals form perceptual representations of individuals, objects, events, and the geometric relations among them, and these representations do not "disappear" when they are absent from the organism's sensory fields. Seed-caching birds, for example, can utilize environmental landmarks to locate thousands of seeds, even months after caching (Kamil & Balda, 1990). Thus, a demonstration that an organism can learn and later retrieve an important relation between Agent [X] and Contingent State of Affairs [Y] and incorporate it into their mental representation, does not qualify as reasoning about mental states *qua* mental states (see Penn et al., 2008).

Representations of mental states are causally dependent on first-order perceptual representations, but not vice versa. Although science frequently relies upon underspecified models, given the standard tools of cognitive science (see Penn et al., 2008), the necessity and sufficiency of lower-order systems to explain the behaviors in question, are not, in this case, *ad hoc* alternatives to higher-order explanations of animal behavior. Rather, the higher-order systems (e.g., reasoning about <*seeing*>) *are causally dependent upon the lower order systems.* To wit, any organism that can reason about <*seeing*> as a mental state must also be able to keep track of the relevant geometrical first-order relations described above. An organism cannot represent that another organism can <*see*> (or has <*seen*>) something without keeping track of the geometrical relations in question.[1] Importantly, the reverse does not follow: organisms could track relevant geometrical relations and act upon them without having any higher-order understanding of a mental state like <*seeing*>. Thus, the causal dependency of the purported connection between first- and higher-order mental states is *asymmetrical*.

First-order representations of agents, objects, and events are modality redundant. By way of example, imagine a gazelle walking past a thicket wherein a tiger is visually obscured. The gazelle's first-order representation of *tiger-in-the-thicket* can be independently activated by a variety of modality-independent sensory inputs including, but not limited to, growling vocalizations, the characteristic sound of tiger movements, the olfactory profile of tigers, etc. In short, it is not requisite to see a tiger to represent that one is present. Any of the stimuli listed above are likely to prove sufficient to allow the gazelle to execute a survival-promoting behavioral response (running away, mobbing, etc.).

We now outline how each of these uncontroversial claims undermine the theory of mind inference Bugnyar et al. (2016) wish to draw from their experimental procedure.

The study was conducted with ravens in two adjoining rooms separated by an opaque partition. This partition could be manipulated so that it either contained a large transparent window or a very small transparent window (a "peephole"). The "windows" and "peepholes" could be opened or closed. In a "baseline" phase, a "subject" (the raven whose behavior was observed by experimenters) was placed in one of the rooms and a second raven was placed in the adjoining room. The subject was given food to cache. In an "observed condition" one window was open (hereafter, "window open" trials). In the "non-observed

[1] The failure to attend to this distinction has led to considerable confusions in the empirical literature. For example, one could argue that there is nothing more to understanding "seeing" than tracking and productively using such geometric relations (see discussion in Hare et al., 2000). But such first-order geometric representations do not qualify as representations of mental states qua <*mental states*> (Butterfill & Apperly, 2013; Povinelli & Vonk, 2004). Indeed, Bugnyar et al. (2016) explicitly reject previous demonstrations that animals can and do keep track of such agent-based geometric relations as evidence of reasoning about the mental state of <*seeing*>.

condition" the window was closed (hereafter "window closed" trials).[2] In both "window open" and "window closed" trials, auditory communication between the subject and the other bird occurred *ad hoc*.

Importantly, the baseline phase was used to define the subject's baseline rates of caching when the other raven was either (a) in principle *visually observable* to the subject ("window open" trials) or (b) *not visually observable* to the subject ("window closed" trials). This phase was not used to determine whether the subjects were reasoning about whether the other raven could <*see*> them. Such a higher-order account of the "window open" trials can be constructed as a series of premises (P1-P3) and a conclusion (C) from the subject's point of view:

(1) P1 Visual signal that a raven is located at *A*,
P2 Unobstructed geometric relation between *A* and here,
P3 Other raven can <*see*> me, will <*know*> where I have hidden food and may therefore steal it later,
C *Execute less re-caching.*

Where '*A*' denotes the entire space in the adjoining room (assuming the set of all points in that space are co-extensive with the sets of all points for which is an unobstructed geometric relation to the subject's position).

Note that Bugnyar et al. (2016) implicitly acknowledge that the following reduced account (devoid of any mental state ascriptions) is sufficient:

(1.1) P1 Visual signal that a raven is located at *A*.
P2 Unobstructed geometric relation between *A* and here.
C *Execute less re-caching.*

Admittedly, our human folk psychology imposes a mentalistic reinterpretation of the geometric relation in question. But as noted above, the facts about the world are such that the subjects need not know anything about higher-order mental states to respond intelligently.

In an attempt to isolate a unique causal role for at least some of the mentalistic representational content in (1), Bugnyar et al. (2016) implemented a final (in their view, diagnostic) "experimental phase."[3] This phase presented the subjects with three types of trials: (a) the original two baseline conditions ("window open" and "window closed") and (b) a new condition labeled a "playback/peephole" trial. On "playback/peephole" trials, the subject was present in its normal room with food and the peephole to the

[2] We have rephrased the language here for two important reasons. First, even during the "window closed" trials the subject could hear the other raven, and thus there is every reason to believe the subject represented another raven as being present in the adjoining room. Indeed, Bugnyar et al.'s (2016) key experimental manipulations (see below) depend on this fact. Second, what the experimenters actually manipulated was whether the window was open or shut, not whether the subject *was observed* or not. Further, although the subject could never see the other raven during "window closed" trials, during the "window open" trials it remains unclear how often and when the subjects visually perceived the other raven. We do not question the effectiveness of the manipulation, but rather we attempt to elucidate the underlying claims of the experimenters. Is the claim that the elevated levels of caching in the "window open" condition were solely triggered by the visual input of the raven?

[3] Between the baseline and experimental phases, the subjects were led through the following exercise. From the adjoining room, an experimenter opened the peephole and obtained the subject's "attention." As the subject looked through the peephole, the experimenter cached a piece of food. The experimenter then let the subject into the room. Subjects were required to successfully uncover the food within 30 seconds on two consecutive trials in order to advance to the next "experimental phase." Although our human folk psychology seems to warrant this procedure, only subjects that "passed" were allowed to participate in the experimental phase, so it cannot be known whether this step was necessary. Thus, we do not consider it further.

adjoining room was open. However, no raven was present in the adjoining room. Instead, the experimenters played back raven noises (recorded during earlier phases) from a hidden speaker in the adjoining room.

The researchers report that the subjects finished their caches quickly and engaged in few re-caches during both "playback/peephole" trials and "window open" trials. In contrast, subjects spent more time completing caches and re-cached more during the "window closed" trials. Bugnyar et al. (2016) concluded that because the profile of caching behavior in the "playback/peephole" and "window open" trials were comparable—despite the fact that the subject never saw another raven during "playback/peephole" trials—this should be taken as strong evidence that subjects represented the other raven's mental state of <*seeing*>. Specifically, Bugnyar et al. (2016) argue:

> Because conspecifics *were never present* [emphasis added] in the peephole context in either the pretraining or the test conditions, these data demonstrate a flexible ability to individually modulate distinct behaviours to specific and novel circumstances, which cannot be reduced to a tracking of gaze cues (p. 4).

We strongly disagree.

To begin, recall that Bugnyar et al. (2016) acknowledge that reduced levels of re-caching behavior during the "window open" as compared to "window closed" trials can be completely explained by the subject's use of the geometric relations in question (see [1], above). However, the subject's representational state described in (1) covers only the case wherein the subject has ongoing access to the visual stimuli of the other raven. But as Bungyar et al. (2016) and many others note (see above), animals can and do remember previous instances of such geometric relations *and these representations can influence present behavior*. Thus:

(2)	P1	The raven located at *A* was present when food was cached here.
	P2	There was an unobstructed geometric relation between *A* and here.
	C	*Execute less re-caching.*

We stress that (2) is a precise formulation of the general criticism Bugnyar et al. (2016) level against the previous research of Clayton and Emery (2001).

The previous analysis applies with equal force to Bugnyar et al.'s (2016) "playback condition." Recall our earlier example: gazelles do not need to visually perceive the tiger to execute an appropriate behavioral response. In the playback trials, when the subject hears (but does not see) another raven in the adjoining room, *and there is the possibility of an unobstructed geometric relation between the spatial source of that other bird and themselves*, the subject executes the appropriate behavior as follows:

(3)	P1	Auditory signal that a raven is located at *A*.
	P2	There exists a set of locations at *A* (A_{peep}) that yield an unobstructed geometric relation between *A* and here.
	P3	Raven at *A* could be located at $A_{peep.}$
	C	*Execute less re-caching.*

Thus, the only relevant difference between "playback trials" and "window open" trials is that during the former, the proximate trigger of the representation of the other raven is an auditory (as opposed to visual) stimulus (the size of the unobstructed geometric relation is irrelevant to this analysis). The same is true of "window closed" trials except for one critical difference—the geometric relation relevant to elevated re-caching is different:

(4)	P1	Auditory ~~Visual~~ signal that a bird is located at *A*.
	P2	There is *no* possibility of an unobstructed geometric relation between *A* and here.

C *Execute higher levels of re-caching*

Crucially, during all of the trials used in the experimental phase (including the purportedly diagnostic "playback/peephole"), the subject is afforded direct knowledge of a raven's presence in the opposite room (NB: But by the authors' own logic, it is no "less direct" simply because it is triggered by auditory stimuli). The only relevant difference is whether the relevant geometric relationship (the unobstructed path) holds. Indeed, the only variable that was shown to influence the caching behavior was whether there was an unobstructed geometric relation between the two rooms. Thus, the variable Bugnyar et al. (2016) sought to eliminate as a confound is, in fact, the only variable to which the ravens appeared sensitive.

This confound is not explicitly stated or eliminated by the authors because of the same assumptions:

(5)
- P1 Auditory signal of raven located at *A*.
- C1 Raven present at A.
- P2 There exists a set of locations at *A* (namely, A_{peep}) that yield an unobstructed geometric relation between *A* and here.
- C2 Raven present at *A* could be in an unobstructed geometric relation to here (spatial positions $A_{i\text{-}n}$).
- C3 If raven at *A* were at A_{peep}, that raven might be able to <*see*> me and therefore <*know*> where food is hidden and possibly steal it.
- C4 *Execute less re-caching.*

Thus, on both Bugnyar et al.'s (2016) account, as well as the representationally reduced account provided in (3), subjects have the ability to predict the possible spatial positions of other ravens and use such feed-forward models to exploit salient geometric relations. Once this is granted, the causally superfluous role of the mental state representations in (C3) above becomes evident.

We conclude that Bugnyar et al. (2016) do not employ a method that can determine whether or not ravens represent <*seeing*>. Could such a study be designed to do so? Barring a drastic, fundamental shift in the testing methodologies previously and currently employed by comparative psychologists, we remain skeptical.

References

Bräuer, J., Call, J., & Tomasello, M. (2007). Chimpanzees really know what others can see in a competitive situation. *Animal Cognition, 10*(4), 439–448.

Bugnyar, T., Reber, S. A., & Buckner, C. (2016). Ravens attribute visual access to unseen competitors. *Nature Communications, 7*, 1-6.

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language, 28*(5), 606-637.

Emery, N. J., & Clayton, N. S. (2001). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature, 414*(6862), 443–446.

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour, 59*(4), 771–785.

Kamil, A. C., & Balda, R. P. (1990). Spatial memory in seed-caching Corvids. In G. H. Bower (Ed.). *The psychology of learning and motivation: Advanccs in research and theory, Vol. 26.* (pp. 1-25). Academic Press.

Lurz, R. W., & Krachun, C. (2011). How could we know whether nonhuman primates understand others' internal goals and intentions? Solving Povinelli's problem. *Review of Philosophy and Psychology, 2*, 449-481.

Lurz, R. W., Kanet, S., & Krachun, C. (2014). Animal mindreading: A defense of optimistic agnosticism. *Mind & Language 29*(4), 428-454.

Lurz, R., Krachun, C., Mahovetz, L., Wilson, M. J., & Hopkins, W. (2018). Chimpanzees gesture to humans in mirrors: Using reflection to dissociate seeing from line of gaze. *Animal Behaviour, 135*, 239-249.

Penn, D. C, Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *The Behavioral and Brain Sciences, 31(2)*, 109-130.

Povinelli, D. J., & Vonk, J. (2004). We don't need a microscope to explore the chimpanzee's mind. *Mind & Language, 19*(1), 1-28.

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