Evolutionary psychology, learning, and belief signaling: design for natural and artificial systems

Eric Funkhouser

Received: 24 February 2021 / Accepted: 9 September 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract
Recent work in the cognitive sciences has argued that beliefs sometimes acquire signaling functions in virtue of their ability to reveal information that manipulates “mindreaders.” This paper sketches some of the evolutionary and design considerations that could take agents from solipsistic goal pursuit to beliefs that serve as social signals. Such beliefs will be governed by norms besides just the traditional norms of epistemology (e.g., truth and rational support). As agents become better at detecting the agency of others, either through evolutionary history or individual learning, the candidate pool for signaling expands. This logic holds for natural and artificial agents that find themselves in recurring social situations that reward the sharing of one’s thoughts.

Keywords Belief · Signaling · Evolutionary psychology · Mindreading

Recent work in the cognitive sciences has argued for the possibility of cognitive “displays,” a “signaling mind,” “socially adaptive belief,” and “falsehoods as tools.” On these views, the mind is said to acquire functions in virtue of the fact that others detect and differentially respond to our mental states. It is not just that someone’s words or behaviors are signals. No, the idea is that the underlying mental states are themselves signals that solicit detection and response.

Since signaling is driven by detection and differential response, the candidate pool for signals widens as a species becomes more adept at detecting the nuances of trait and behavioral variation. As we become better at knowing others—not just their physical attributes, but their behavior and minds as well—a wider range of attributes can come

1Sterelny (2015), Funkhouser (2017), Levy (2018), Williams (2021), and Petersen et al. (2021).
to function as signals. We can extract information about them in new ways for our benefit, and they can also communicate information to us in new ways that benefit them. This allows beliefs to have functions beyond their familiar role in epistemology and practical reasoning, where truth and rationality tend to prevail (at least as norms). When beliefs function as signals, social utility or self-esteem tend to dominate.

The varying detection abilities of “mindreaders” give rise to three levels of social signaling sophistication. These levels correspond to different degrees of sensitivity to the mental states of others. First, physical attributes are recruited for signaling purposes, as is common for animal signaling. At this level, there need not be any receiver-side detection of agency. Second, for animals that are capable of detecting behavior “as such,” purposive behavior can function as signals in virtue of the detected underlying teleology and perspective. Such communicative behavior has not, to my knowledge, been explicitly conceptualized in terms of signaling theory. Third, mental states themselves, if detected, can eventually come to function as signals. In Sects. 2–4, we examine each of these levels.

This article explores belief signaling from a particular angle: What does an evolutionary (e.g., natural selection or individual learning) or design approach tell us about how minds—beliefs, specifically—will or should be shaped in virtue of the fact that they are part of the social world? These approaches are complementary, attending to the dynamics and strategy of mind design. The answer to this question—reached through analytical rather than experimental means—leads us to mental states that function as signals. Our investigation begins by considering the evolution of animals in natural environments, but the logic extends to artificial systems with representational states analogous to those found in higher-level animal minds. Our primary questions will be:

1. Under what conditions should we expect beliefs and other mental states that function as social signals? When is it a good and available strategy?
2. What is the evolutionary trajectory for human beliefs that eventually function as signals? What are the receiver-side requirements for different kinds of signaling, and what signaler-side changes do they tend to induce?
3. In what artificial environments would it be advantageous to consider the benefits of something like belief signaling (e.g., if–then-else statements, data structures, etc. that convey information to other systems for the purpose of manipulation)? And what is the design trajectory, with respect to signaling, for artificial systems with belief-like states?

1 Assumptions and background

Our central hypothesis is that once psychological detection mechanisms (“mindreaders”) are in place, signaling minds will often result so long as the informational content conveyed by a mental state induces (mutually or unilaterally) favorable differential responses in some audience and the essential ingredients of evolutionary psychology or learning are in place. That is, signaling minds will often emerge if psychological systems in an ecosystem have:
Beliefs, desires, values, or emotions,
Variation, inheritance, and differential replication (via biology, culture, or individual learning) among these detected psychological states (or in the tendencies to form such states),
Detection mechanisms (e.g., mindreading, especially among conspecifics) for these beliefs, desires, values, or emotions, and
The detected mental states convey information to others that matters to them.

For example, we have beliefs about ourselves, our kin, and others in our tribe that tend to be self-enhancing or favorable to the in-group (Taylor & Brown, 1988). Other people care about these beliefs and can detect them. The beliefs (or belief-forming tendencies) vary, are inheritable, and they can differentially replicate based on how they manipulate the behavior of others. If the induced manipulation is favorable, the self-enhancing and group-favorable beliefs can come to function as signals.

1.1 Belief

While the evolutionary account of social signaling developed here is supposed to apply to a range of psychological states, belief serves as our primary application. So, something should be said about the nature of belief, or at least about the kinds of beliefs that we should expect to acquire signaling functions. Minimally, the nature of belief will have to be compatible with the existence of belief signaling. The approach taken here is as non-committal as possible while acknowledging the conceptual and empirical limits of belief. The philosophical literature on belief largely agrees that belief is the most general attitude of seriously accepting a proposition as true (Schwitzgebel, 2019). Things get complicated when we push further as to what determines whether a proposition is generally (i.e., not merely in limited contexts) and seriously accepted as true. Broadly speaking, there are two prominent answers to the question of what determines whether a proposition is believed: (1) dispositional or functional accounts and (2) interpretationist accounts. We see the influence of behaviorism within each tradition, with belief either functioning to generate behavior (among other input–output relations) or as a posit that rationalizes behavior. As such, we should expect that belief signals tend to generate or explain actions, if only in the form of assertions. Beliefs that function as signals are fully beliefs, after all, and they often lead to consequential outcomes such as voting decisions and moral judgments. Beliefs also shape our social and epistemic networks. Further, there are some rationality restrictions on belief, meaning that certain combinations of beliefs are unsustainable. It might be that individual beliefs cannot be held without at least some connection to evidence, even if that connection is tenuous. Therefore, if a signaling belief deviates from the objective standards of non-motivated reasoning, we should expect some rationalization to sustain that belief as well as to better display it to others (Levy, 2020).

Which beliefs are particularly good candidates for signaling? Section 1.4 makes a more nuanced positive case, but we can say a bit now about which kinds of beliefs are not good candidates for signaling. Beliefs come in many different forms—perceptual beliefs, episodic memories, mundane beliefs about local matters of fact (e.g., the best route from work to home), reflective beliefs, moral beliefs, ideological beliefs,
religious beliefs, and so on. Only some of these categories are going to offer plausible candidates for belief signaling. Basically, only beliefs that are part of the “social economy”—beliefs that are shared by others and with others—are candidates for belief signaling. Belief signals must be widespread enough and non-fleeting, such that informational content can be derived from them. Episodic memories, perceptual beliefs, and mundane beliefs about local matters of fact fail to meet this standard. The beliefs that function as signals will tend to be discussed and displayed before others, which these are not. Given the importance of discussion (social relevance) and display (detection), beliefs that function as signals will often be asserted in at least some contexts. These are often, though not always, what some philosophers call our “opinions” or “commitments” (Dennett, 1978, Chapter 16; Frankish, 2004, pp. 71–80). These count as beliefs by ordinary standards, as well as by the (admittedly imperfect) practices of social science. The view of belief taken here is rather liberal, and beliefs that function as signals can be either conscious or nonconscious, actively or passively acquired.

1.2 Signaling theory

In Funkhouser (2017), I offered an abstract account of signals that is supposed to cover all manner of signaling, from quorum sensing in bacteria, to the peacock’s tail feather (Zahavi, 1975), to signaling by human applicants in job markets (Spence, 1973).

“Signal:
1. Any object that is successfully designed or selected to communicate information,
2. so as to be detected by some receiver,
3. in order to modify its behavior.” (Funkhouser, 2017, pp. 811–812)

This teleological account of signaling will also be accepted here. Biological evolution, enculturation, or individual learning can generate the functionality captured by clause (1). The chief contribution of this article is to illustrate the evolutionary path to belief signaling, via heightened agential detection, rather than making a conceptual (static) case for it.

Belief signaling requires that beliefs can acquire such functions in the first place. The general viability of evolutionary psychology—the application of evolutionary logic, especially natural selection, to psychological systems—will be assumed (Tooby & Cosmides, 2015). This is to say that psychological states or tendencies relating to beliefs, desires, emotions, and the like can be shaped by a process of natural selection. The basic ingredients for natural selection are variation, inheritance, and differential replication. So, for the purposes of this argument, we are assuming that there are psychological states or tendencies that have allowed for variation, inheritance, and differential replication sufficient to acquire functionality. A biological understanding of this process is supported by the presence of substantial cross-cultural uniformity in beliefs, values, psychological drives (desires), emotions, and the like. If pressed, we could make a decent case for some biological or psychological adaptations within

---

2 Social scientists are aware of the problems with survey responses. A respondent’s apparent endorsement of a proposition might be mere “expressive responding” rather than genuine belief, to give just one possibility (Bullock et al., 2015).
each type—e.g., self-enhancing and otherwise self-serving belief-forming tendencies, incest taboos, the desire to belong and have one’s group flourish, or Ekman’s universal emotions. There certainly is no assumption that all psychological phenomena are adaptations or even that all psychological states or tendencies are capable of such evolution. But some mental states are likely biologically adaptive and functional.

Evolutionary psychology in the biological sense is not the only way for beliefs to acquire signaling (or other) functions, however. Skeptics of (biological) evolutionary psychology might view learning mechanisms—both cultural and individual—as alternative sources of teleology. In fact, most belief signals, especially if we get fine-grained enough, are due to cultural evolution or individual learning. Belief signaling functions acquired through phylogenetic evolution are less common, and when they occur, they tend to operate at a coarser grain (e.g., producing psychological mechanisms that favor self-enhancing belief in general rather than skill at managing a hedge fund in particular, say). Learning mechanisms offer other means for mental states to acquire functions, and some psychological states acquire purposefulness this way. I view such adaptation as supplementing the teleology that is due to biological selection.

1.3 Mindreading

The term ‘mindreading’ is used in slightly different senses by psychologists, ethologists, and philosophers. I will limit its use to refer to any ability to understand another by attributing internal mental states to them. Psychological understanding comes in degrees, and we will draw out the signaling consequences of two meaningfully different grades: behavioral and mentalistic (Csibra & Gergely, 1998; Gergely & Csibra, 2003). So-called behavioral mindreading will not count as mindreading for our purposes. Developmentally normal human beings acquire mentalistic mindreading abilities at around 15-months, though this ability is not fully integrated with other systems (e.g., executive systems) until a few years later (Carruthers, 2013; Onishi & Baillargeon, 2005). We can remain neutral regarding the particular mechanisms by which people mindread. The point is that, somehow or other, people nearly universally have this ability, it manifests at an early age, and they employ it rather easily. It may be that belief signaling depends not just on basic mindreading but also on some integration with other cognitive systems. This is because signals must communicate some other information (besides just the detected belief itself) which in turn modifies the behavior of the receiver. This additional information (e.g., climate change denial communicates Republican affiliation) need not be the result of any conscious or explicit inference, however. Nevertheless, mindreaders must possess the relevant background knowledge in order to properly receive a belief signal.

1.4 Social relevance

Mindreading, by itself, does not exert signaling pressure. The psychological state that is being read must communicate information, reliably enough, that matters to the receiver, such that she is incentivized to look for it and then react. The last piece of the puzzle is to find recurring social situations in which mindreading leads to adaptive
solutions for the mindreader and/or the mindread subject. In many cases, the benefits of mindreading will have nothing to do with signaling, but the following are common situations in which signaling benefits (to receiver and/or sender) can emerge:

- **Behavioral Prediction**: predicting the behavior of another.
- **Behavioral Stabilization**: regulating other people’s behavior so that it becomes stable and predictable.
- **Trust Problems**: discerning whether to approach a person, trust them with your property, expect that they will reciprocate your altruism, etc.
- **Commitment Problems**: getting others to commit to a future action or having a reason to commit yourself to a future action.
- **Coordination Problems**: figuring out how to work with others to jointly accomplish a goal or to work in parallel.

The first two categories are quite broad and include the latter three as special instances: predicting and shaping the behavior of others is a way to achieve trust, commitment, and coordination. But what does any of this have to do with belief signaling? Our behaviors are underpinned by beliefs, desires, values and emotions that contribute to solving these problems. *Mindreading allows for the possibility of detecting and molding these mental states into signals. In turn, beliefs that signal tribal loyalties, self-assessments, values, worldviews, and the like allow us to even better anticipate the behavior of others and rely on them. This generates more robust networks of trust, commitment, and coordination.*

These five categories constitute important problem areas for an ultra-social species like ours to solve. They bear on the major biological tasks from our ancestral past, and they manifest in different cultural forms in contemporary life: avoiding predators, finding food and other resources, fighting, and mating. Non-psychological animal signals have evolved to negotiate these same tasks. Since biological natural selection has worked on psychological traits for the sake of sociality, it could also provide the shape and teleology for beliefs and other mental states to function as social signals for such tasks. In the cultural world, things are even more complex, nuanced, and strategic. In these quickly changing environments, we need new signals that biological evolution does not have time to provide, and individual learning and cultural transmission allow for this functionality. New signals are especially needed as modern social interactions involve strangers, complex coalitions (corporate, ideological, religious, etc.), long-term commitments, and great spatial distance.

The informational content that beliefs signal will often have significance not only to the receiver but to the sender as well. The sender will often, though not always, have an incentive to broadcast it, along with the audience having an incentive to pay attention. But the incentives do not always align. The familiar distinction between competitive and cooperative contexts, which crosscuts the above five categories, is relevant here.

### 1.4.1 Competition

Suppose that people are competing for some good (e.g., food, property, mates, etc.). Mindreaders in these situations have an incentive to detect the beliefs, desires, and
emotions relevant to their competitor’s pursuit of that good. We might want to know where the other animals believe that berries can be found, what property they find most desirable, or how jealous they would be if they lost out on a mating opportunity. This information has strategic value to the receiver. It can tell one where to go, what to offer, what to avoid, and so on.

The mindread subject often will have an incentive to hide these very thoughts. So, if they are signals, then they often are such despite the interests of the signaler. In these cases, there is no effort to advertise or display one’s genuine thoughts, and there might even be an effort to present a misleading picture. At least sometimes, though, there is an incentive to reveal one’s genuine thoughts even in competitive settings. There can be value in honestly broadcasting to an adversary one’s willingness to retaliate or follow through on a threat. Of course, other times there are incentives to broadcast dishonest signals.

1.4.2 Cooperation

We have incentives to both mindread and to be mindread for the sake of cooperative action. We will often want others to see that we have the following kinds of beliefs, each of which furthers cooperation:

1. Beliefs that reveal our tribal affiliations.
2. Beliefs that signal our values and good intentions.
3. Beliefs about the courses of action we will take in both actual and counterfactual situations so as to achieve some end.

These beliefs have anticipatory effects (e.g., people anticipate that we can be trusted or that we will act in a particular manner). They also might not have a simple or straightforward connection to overt behavior, so only mentalistic mindreaders (as described later) can detect them.

This is all well and good, but how do we actually get to belief signaling, if at all? When and where should we expect it to emerge? The following sections will show how the candidate pool for signals grows as animals become better at detecting the agency of others. We should expect signaling minds to emerge alongside mentalistic mindreading. Similarly, we should expect the signaling of “psychological” states in intelligently designed or evolving artificial devices (e.g., the intentions of an autonomous vehicle) as these devices better communicate with and understand each other.

2 Solipsistic goal pursuit

2.1 Cognitive abilities

Let’s start by considering simple agents whose behavior is explained in terms of psychological states like beliefs, desires, and emotions. These agents have goals or

---

3 Maybe some of these will not be full-fledged intentional states but simply proto-beliefs or the like. This need not concern us.
motives provided by their desires and emotions, and they manifest enough rationality for us to attribute them with beliefs that are instrumental to their goal achievement. Say, the turtle desires to eat a worm, and it follows a certain path because it believes that doing so is a means to getting the worm. Significantly, though, these agents do not recognize the behavior of others (conspecifics, but also any other kind of agent) as rational goal pursuit. Not only do they fail to attribute internal mental states to others (mentalistic stance), they do not even recognize the goal pursuit behind agential behavior (teleological stance) (Csibra & Gergely, 1998; Gergely & Csibra, 2003). This is solipsistic navigation in the sense that such agents act (though not self-consciously) as if they are the only purposive agents in the world. Solipsistic agents are not literally alone in the world, of course. They interact with other agents, but they do not engage with others as agents acting for purposes.

This is not to say, however, that such agents cannot identify behavioral patterns in any sense. They are still capable of treating other agents as objects that fall under the same folk theories that apply to all physical objects. That is, they can take the physical stance—as well as other non-teleological stances (e.g., learning simple associations)—toward them (Dennett, 1987). These animals might expect certain regular movements from others, but they do not have a general capacity to see the behavior of others as the rational pursuit of a goal. Their ability to predict movements comes from detecting cues, such as a courtship display to solicit mating or the baring of teeth to indicate a forthcoming attack.

2.2 Effects

Solipsistic agents have mental states that help them better navigate their environment and acquire the valuable goods that further replication or reinforce a strategy. The underlying mental states can be selected and shaped by their ability to produce these outcomes, so there is feedback from replicative success (biological, cultural, or learning) to the next generation of mental states. (See Fig. 1.) True beliefs will tend to be rewarded in such systems, as true belief tends to further the acquisition of goods for solipsistic agents (Millikan, 1984). The effectiveness of these mental states can still be somewhat sensitive to the behavior of others, since their behavior, in combination with that of the self, often determines one’s success at acquiring goods and eventual replicative success. It does no good for an ant to devote its energies to building an ant hill unless others do so as well. While ants might not have beliefs and desires, the point generalizes to animals that do. Some mental states de facto better complement behavioral tendencies in others, but in solipsistic agents those states are not selected or shaped by a recognition of teleological or intentional behavior by others—there simply is no such recognition. So, there is not yet a signaling mind.

Nevertheless, solipsistic agents can achieve many social tasks. They can predict, shape, trust, commit, and coordinate with the behavior of others. But they can do this only by detecting physical cues, regularities, or associations absent any teleological or mentalistic interpretation. For them, signaling occurs only in those physical realms.
2.3 Examples in natural and artificial systems

We are searching for examples of animals that have intentional psychological states but which lack the ability to see others as teleological agents. As with other mental phenomena (e.g., consciousness), it is very hard to classify animals in these regards. Perhaps our turtle, which has belief-like and desire-like representations of its environment, can learn a lot from conditioning without appreciating behavior “as such”
by others (turtles, humans, whatever). That is, they might recognize some pattern of behavior but do not recognize it as rational goal pursuit.⁴

If one is skeptical about intentional psychology residing with these simpler animals, we can see them as having proto-beliefs (representations of their environment) and proto-desires (drives). There are many forms of animal signaling among such simple creatures that impacts their social behavior and perhaps their underlying psychology. A bird’s song, poison dart frog’s coloration, gazelle’s stotting, sparrow’s badge, and a rattlesnake’s rattle are just a few examples of signals that communicate socially relevant information. Such examples show that behavioral prediction, shaping, trust, commitment, and coordination can be achieved absent a teleological stance or mindreading.

Among artificial systems, think of cars that process information about their environment—including cars and pedestrians—and make suitable adjustments, but do not recognize other cars as purposive agents. This is true of cars at levels 0–3 of automation, such as those capable of automated lane changing (Rasouli & Tsotsos, 2018). This is a “social” skill that requires predicting the behavior of other cars and coordinating with their movements, but there need not be any discernment of the goals of other cars or any attribution of intentions. Or consider a missile system that has belief-like and goal-like representations directed at the enemy, yet it does not recognize the purposiveness behind, say, the enemy’s missile defense system. Simple radio signals might let it know which missiles to “trust” and which ones to destroy.

In many cases there can be an illusion of mindreading. Animals are hard-wired with certain instincts which make it look like they have a teleological understanding or even a theory of mind when they do not. Think of the broken wing display of a killdeer. There are also many learned cues—e.g., by the “mathematical horse” Clever Hans or your average dog. If behavior is very simple to simulate, then it is less likely that it is truly rational or guided by belief-desire pairs that cause purposive behavior (e.g., bird flock behavior that can be simulated with simple rules) (Reynolds, 1987). Even points of view (e.g., visual perspectives) can be appreciated without attributing beliefs.⁵

3 Goal pursuit with behavior-reading

3.1 Cognitive abilities

The next level of cognitive sophistication we will consider introduces “behavior-reading” or the teleological stance, meaning that there is some detection of behavior “as such.” Agents are now able to learn to appreciate the teleological nature of the behavior of others, but without invoking mentalistic explanations. The behavior of other agents is predicted and anticipated, and not simply by detecting physical cues

---

⁴ I am not committing myself to any particular claim about turtle psychology—they are simply being used as a placeholder or stipulated example.

⁵ There are several complications for us to remember, including that mindreading is (1) multi-faceted (gaze, intention, etc.) and not always well-integrated (e.g., in the very young), (2) limited in scope (applied only to conspecifics, predators), and (3) inconsistently present (only in competitive environments, say).
like those we found with solipsistic goal pursuit. Rather, others are treated as teleo-
logical systems—with perspectives (or limited information) and purposes—that are 
engaged in rational behavior. Some ethologists call this “mindreading,” but this skill 
do not require positing any mental states.

Ethologists, psychologists, cognitive scientists, and philosophers have recognized,
albeit in slightly different ways, the fruitfulness in discerning lower-level “mindread-
ing” abilities that (despite the label) do not demand the attribution of mental states. 
Krebs and Dawkins (1984, p. 386), ethologists writing about animal signaling, define 
mindreading as an animal’s use of statistical laws to predict future animal behavior. 
This can occur by attending to cues in the manner of our solipsistic agents described 
in the previous section. Such a skill is hardly worthy of the appellation mindread-
ing. Krebs and Dawkins (1984) describe a special case, however, which gets closer to 
genuine behavior-reading.

“For an animal that has any kind of social life, or that is a predator or is preyed 
upon, these probable consequences will depend crucially on the internal moti-
vational state and probable future behaviour of other animals—rivals, mates, 
parents, offspring, prey, predators, parasites, hosts.” (386)

But now they have swung to other end of the pendulum—rather than talking about 
external cues, they invoke internal mental states.

There is a conceptual middle ground: agents that can discern purposive behavior 
(teleology) but without conceptualizing or attributing internal mental states as driving 
that behavior. This concept is well-articulated by cognitive scientists Gergely and 
Csibra:

This raises the possibility of a representationally less sophisticated organism 
that—although unable to represent intentional mental states—could nevertheless 
have evolved a reality-based interpretational strategy to represent goal-directed 
actions. This ‘mindblind’ creature could represent its normative assumptions 
about the essentially teleological nature of actions in terms of the very same 
inferential principle of rational action that forms the central component of 
the mentalistic stance. The organism could then evaluate the efficiency of an 
observed action as a means to a goal by applying the rationality principle to the 
relevant aspects of reality states, as they are specified by the organism’s own 
representations formed while perceiving the action unfold. (Gergely & Csibra, 
2003, p. 290)

Note that such creatures detect goals, but, when employing the teleological stance, they 
predict behavior by looking directly at reality rather than by taking on the epistemic 
perspective of the other. So, there are no belief attributions. As such, this approach 
to predicting behavior fails when the interpreted subject has (from the interpreter’s 
perspective) false beliefs.

Our notion of behavior-reading is close to Gergely and Csibra’s teleological stance, 
except that we can allow for some appreciation of the perspective of the other. (If 
this appreciation reaches the level of outright belief attribution, then it has passed 
into our third category.) This appreciation of perspective includes gaze tracking and 
an understanding of gesture (e.g., responsiveness to pointing). With the teleological
stance, an animal appreciates the many different ways that the goal-directed action of another can be pursued, as well as the different ways in which the same purpose can be revealed. This “multi-track sensitivity” distinguishes the teleological stance from the more simplistic stimulus–response abilities of solipsistic animals (Bermúdez, 2009).6 In light of this, behavior-readers can anticipate the adjustments that would be made were certain obstacles put in place. Still, merely behavior-reading animals do not reach the heights of mentalistic mindreading. They fail to posit internal causes for behavior (specifically, anything like propositional attitudes) or to consider the constraints of holism and internal rationality (Zawidzki, 2013, pp. 13–14). They can only perform more modest tasks that involve appreciating the perspective of another, such as tracking gaze or pointing gestures.

3.2 Effects

Rather than simply seeing patterns in movement (color, sound, etc.), these teleological agents can recognize the various forms that action can take when others are pursuing goals. Recognizing this flexibility or multiple realizability of action allows for better prediction, especially as situations change and require the selection of different means to goal achievement. No simple, non-teleological pattern detection can rival the predictive power of the teleological stance. It is not hard to see how this greater predictive power could have enhanced the reproductive success of behavior-readers in recurring social situations (e.g., mate selection, in-group competition, violent encounters, predator–prey interactions). Recognizing the teleology (goal-pursuit) in the actions of others clearly furthers coordinative activity. Under friendly conditions, one can cooperate with others to help achieve joint goals, and one can align one’s own projects under the assumption that certain goal-directed tasks are being completed by others (i.e., division of labor). When less friendly intentions are detected, behavior-readers can learn, without having to risk a physical confrontation, whom to avoid. At its core, the teleological stance involves a recognition of intention.7 Once intentions can be recognized, novel ways of conveying commitment and trust (or their opposites) arise. Specific intentions can be conveyed through attention to environmental cues—say, through gaze or pointing—and not just through hard-wired animal signals.

The “behavioral feedback” depicted in Fig. 2—the critical new ingredient of reading behavior as teleological—goes in two directions. First, the purposiveness we read into the behavior of others can cause us to change our mental states and subsequent behavior, as we just discussed. Second, this behavior-reading ability can also have consequences for those whose behavior is being read as teleological. This is the point at which novel signaling—intention signaling—can emerge. We can exert selective pressure on their mental states to the extent that we differentially respond to the detected purposive behavior that they generate. Their mental states can be more robustly shaped

6 Bermúdez refers to this as minimal mindreading. His minimal mindreading is a bit different than our teleological stance, however, as it simply requires behavioral responses that covary with the psychological states of others, though not necessarily by any recognition of purposefulness. Solipsistic agents could meet this standard.

7 Think of these intentions as being either 1) non-propositional (e.g., a simple goal like “to open the door”), or 2) an exception to the rule that mere behavior-readers do not detect propositional attitudes.
by the fact that we behavior-readers have a greater appreciation of their purposiveness and perspective. 8 Alternatively, they sometimes will have incentives to mislead us behavior-readers about their perspective and purpose (dishonest signaling). This is the social pressure at work, which moves us from agential solipsism. This second kind of behavioral feedback can produce changes to the mindread subject’s psychology comparable to the changes in the underlying conditions that give rise to canonical

---

8 This is a previously unrecognized level of signaling that gets us halfway to the belief signaling described in Funkhouser (2017).
animal signals (tail feathers, camouflage, coloration, etc.). In the case of animal signaling, these underlying conditions (e.g., the chemical basis for coloration) are not themselves signals, because they are not detected by other animals in the relevant sense. Likewise, we do not have belief signaling with mere behavior-reading, which can alter mental states (via our differential response to teleological behavior) without detecting those mental states.

3.3 Examples in natural and artificial systems

We are looking for examples of animals that can detect perspective, purposiveness, and rational action in the behavior of others, but without recognizing cognitive states like belief and desire. Some corvids recache their food if they see that another bird was watching them the first time they hid it. This shows some recognition of the perspective and goals of another. But can the detected goals themselves become signals, such that the *purposiveness* of a behavior—not just the behavior itself—exists so as to be recognized for communicative purposes?

When an animal tracks the eye movements of another as a means to focus on an external object, the eye gaze conveys intentions. The gaze tracking animal is aware that the gazing animal has directed its eyes for a purpose. Gaze monitoring can be a means to achieve joint attention and better enable cooperative or coordinated action. The purposes of a gaze are often twofold: to draw attention to an object and to suggest some action in light of that attention. To give just one example, Emery (2000, p. 592) describes studies showing that gorillas can use eye contact to draw an experimenter’s attention to a latch (purpose #1) so as to convey the gorilla’s goal that it be opened (purpose #2). He concludes:

“The capacity to follow another individual’s gaze onto specific objects, events and individuals in the environment is accomplished by some species of macaques and apes, suggesting that eye gaze may be an important and complex component of learning about the environment from conspecifics, which does not require a more sophisticated ‘theory of mind’ mechanism.” (Emery, 2000, p. 599)

Things are a bit different for human beings, who can mindread at a young age. But even before mindreading emerges, at a mere 6 months of age, human infants appear to go through a teleological stage where we detect communicative signals and gaze monitor (Senju & Csibra, 2008). This is likely critical to our language development, as the young child tracks her mother’s gaze to discern an object (purpose #1) to which the word should be associated (purpose #2) (Emery, 2000, p. 588). Pointing can also be used to signal these kinds of intentions, often in conjunction with gaze alternation (i.e., when the signaler looks from the gaze tracker’s eyes to the external object). Indeed, Leavens and Hopkins (1998) found that chimpanzees paired pointing gestures with gaze alternation to communicate their intentions (e.g., to get a banana).

Computers, robots, and self-driving cars are now being built with this ability to both detect and signal intentions. Robots, like animals, become more intelligent through social learning. Imitation plays a prominent role in this process, but how does a robot know which behavior is to be imitated? They can do so by employing the teleological
stance and detecting intentions—the behavior to be imitated is determined by the goal being pursued or the focus of attention (indicated by pointing or gaze) (Breazeal & Scassellati, 2002). Once robots acquire the cognitive ability to track goals, signaling functions on the modeler’s side can arise to shape robot behavior. Either their fellow robots or their human handlers acquire intentions so as to manipulate the behavior of the intention-detecting robot. Current attempts at machine learning fare poorly (compared to humans) at tasks like playing video games that require temporal planning and involve multiple sub-goals, since they tend to lack the ability to see actions as teleological (Lake et al., 2017). Although, there have been recent attempts to formalize teleological vocabulary in terms of a “utility calculus” (discovered by working backwards from observed behavior) that could, in principle, be programmed into artificial intelligence (Jara-Ettinger et al., 2015). This procedure involves identifying a goal and computing the rational means to goal satisfaction given constraints, in line with the teleological stance of Gergely and Csibra.

Perhaps nowhere is the need to program devices with teleological reasoning more pressing than in the development of autonomous vehicles. In order to respond quickly and safely to its environment, an autonomous vehicle must be able to track the intentions of other cars and pedestrians. This is accomplished by attending to features such as blinkers (car) or head orientation (human). These directly perceived features are often recognized as signals, but the intention itself can function as a signal as well—here we have a signaling chain (Funkhouser, 2020). The recognized intention is distinct from the signal (e.g., blinker) for it, and the intention can communicate information as well. Some engineers, in an effort to make autonomous driving more human-like, find inspiration in the biological and psychological literature on joint attention, teleology, and theory of mind (Rasouli & Tsotsos, 2018).

Notice that when animals or machines can detect purposes, they can introduce novel signals for changing environments. Whereas much animal signaling is hardwired for a very particular task—such as acquiring a mate or deterring an attacker—a teleological signal, like a gaze, can be recruited to draw attention to any object in one’s environment and to communicate many different intended behaviors (e.g., open the latch). The eye movements themselves, seen as mere physical phenomena, cannot signal in this way. The teleology itself must be detected—it serves as the signal. Since the information is not communicated unless the intention is detected, the intention is a signal.

4 Goal pursuit with mentalistic mindreading

4.1 Cognitive abilities

Finally, consider agents that track mental states treated as internal causes of behavior which fit into a rational structure. Here we have mentalistic mindreading: recognition of belief “as such.” This level of detection qualifies belief as a signaling candidate (Funkhouser, 2018). Mentalistic mindreading need not involve conscious or explicit psychological theorizing, but it does require at least the functional recognition of false belief and rationality—e.g., a child passing the “Sally-Anne” false belief test (Baron-Cohen et al., 1985), an appropriate violation-of-expectation response (Onishi
& Baillargeon, 2005), and so on. Full-fledged belief attribution is what most separates mentalistic mindreading from behavior-reading, with the latter only recognizing goals and perspectives (e.g., visual orientation). Mentalistic mindreading might require some functional appreciation of both the holism of belief (e.g., attributing beliefs contingent upon other the assumption of other beliefs) and its combinatorial potential (e.g., generating action in combination with desire). Assertion and other expressive behaviors rise in importance, as these are especially common and easy ways to reveal internal states that would remain hidden to one with a limited exposure to our behavioral history (Fig. 3).

![Fig. 3 Goal pursuit with mentalistic mindreading](image-url)
4.2 Effects

Recall the five recurring social situations discussed earlier: prediction, stabilization, trust, commitment, and coordination. Mentalistic mindreading frequently produces new receiver-side benefits of each type. Indeed, many of the tests for mindreading operationalize mentalistic mindreading in terms of behavioral prediction (or reaction to surprising behavior). This predictive ability has obvious payoffs in virtue of anticipating behavior across contexts by positing persisting internal states (beliefs). It also allows for tactically deceptive actions as the Sally-Anne test suggests (e.g., moving a marble so that someone looks in the wrong place in accord with their false belief). To a limited extent, we see this ability with corvids who sometimes recache food when observed, but the human ability to deceive is much more robust in that it is domain general and can be iterated (e.g., “she believes that I believe…”). In addition to its greater predictive power, mindreading allows us to better monitor the minds of others and mold their beliefs in a self-serving manner. Such “mindshaping” produces homogeneity that can be beneficial to the individual and group (Zawidzki, 2013). Mindshaping allows us to use others as signaling tools (e.g., party leaders spread lies about a disease so that such beliefs can be used to discern who is a loyal party member), if their molded beliefs function as signals to others.

It is not difficult to see how mindreading furthers trust, commitment, and coordinative activity that can be mutually beneficial. Indeed, mindreading has been credited by some as the keystone to the unparalleled levels of cooperation (and deception!) among non-related individuals that we find with humans (Dunbar, 1998). Belief detection can benefit the receiver (mindreader) in these recurring social situations by:

1. Discerning factual beliefs and values that indicate general affinity—that we are fellow Republicans, Unitarians, Yankees fans or whatever. We have more justification to trust, commit to, and coordinate with the likeminded.
2. Discerning more specific beliefs about the individual’s values and good intentions which show them to be worthy of trust and commitment. We are more likely to trust and commit to somebody if we detect that they believe in honesty and reciprocation, say.
3. Revealing their longer-term or larger-scale projects, the assistance they need, and the help they can offer. If their projects overlap with ours, we can coordinate our actions to achieve a mutually desirable outcome that we could not achieve on our own.

In all these cases, the belief detection is valuable apart from its role in predicting their behavior. This suggests other functions for belief besides guiding action. Belief can also be used to solicit trust, commitment, and coordinated activity.

I have focused on receiver-side incentives to attend to these beliefs as potential signals, but we should see that signalers also can have incentives to use their beliefs as signals to achieve these social goals. The mindread agent’s mental states might be shaped or selected because of the effects they have on mindreaders once those states

---

It should be noted that Zawidzki argues that mindshaping precedes mindreading. I am pointing out a way in which the influence can go in the other direction.
are detected. Similar effects can happen with mere behavior-reading, but we do not have belief signaling without actual detection of the belief “as such.”

Significantly, when beliefs acquire functions beyond their customary action-guiding function, the norms that apply to them also change. Beliefs that function as social signals should produce social utility for the signaler; beliefs that function as signals to self should generate pragmatic benefits of their own (e.g., self-esteem, motivation). Beliefs are always governed by norms of rationality to some extent; reason imposes limits on what we can believe given the available evidence (Davidson, 1984; Dennett, 1987). But we need not accept the further claim that truth should be the standard of correctness for beliefs (contra Wedgwood, 2002). False beliefs could be very effective signals (e.g., the belief that the 2020 U.S. election results were fraudulent could signal loyalty to Trump). The signaling function brings directly pragmatic considerations to the forefront, to be balanced against traditional epistemic norms. Two questions are critical to this assessment: (1) Is the sacrifice of epistemic norms (and the bad consequences it induces) outweighed by the pragmatic benefits of signaling? (2) Practically speaking, to what extent can we successfully disregard epistemic norms for the sake of signaling?

Beliefs that are detected induce much greater manipulations in others—and therefore are themselves much more susceptible to selective pressure for signaling reasons—than do beliefs that are only indirectly modified in light of the detected behavioral changes to which they give rise, as depicted in Figs. 1 and 2. These manipulations become much more anticipatory, as the belief can be detected before it generates stereotypical, overt behavior. For example, a job candidate displays a belief in their domain competency, making them appealing to the employer even before their performance is observed. People react to the belief-as-signal rather than merely to the behavioral effects of the belief. Further, the belief does not simply function to generate action but plays other roles, such as social classification (e.g., indicating Republican party membership on the basis of climate change denial) or revealing personality type (Funkhouser, 2020). These are significant for forming alliances or social coordination. When interests in a larger project overlap, the benefits of coordinative activity are mutual. Trust and commitment can also be mutually beneficial because they lead to reciprocated altruism. Animals can benefit by signaling in advance their likelihood of reciprocating, using beliefs to convey this information. In the human case, this is often done by endorsing moral precepts such as “do unto others as you would have them do unto you.” Even if the signaler gains no reciprocated altruism, she might signal trust or commitment out of weakness so as to avoid greater conflict. We see this elsewhere in the animal world when an animal signals surrender in order to avoid an even more damaging outcome.

Our big theme has been that as detection powers improve, the candidate pool for signals expands. I have focused on improved psychological detection. (Improved physiological detection, clothing detection, general cultural detection, etc. suggest other directions for signaling.) The fact that mindreaders can be strongly incentivized to detect and respond to beliefs makes them a very logical option for signaler-side modifications to either better broadcast beliefs or distort them in ways that generate more favorable responses from others. Of course, these can also be incentives for deception. Signalers could indicate beliefs that they do not actually have, as in the case of a
religious hypocrite who merely fakes belief. Or people could genuinely possess those beliefs but lack the characteristics that such beliefs typically indicate—e.g., they are a climate change skeptic (which typically signals Republican partisanship), yet they are not a reliable Republican. Whether outright belief is incentivized will obviously vary by content and will need to be assessed on a case-by-case basis. But many socially important beliefs plausibly fit the bill: self-enhancing beliefs, commitments to group norms, and beliefs that reflect tribal identities. In addition to generating action, the beliefs signal information that boosts our value in social markets (jobs, dating, society life, etc.), shows our loyalty, or reveals us to be worthy of social-economic transactions—compare these to the ethnic signals of belief systems discussed by Richerson and Boyd (2005, pp. 211–213). Other times, beliefs that have positive social value (and thus are prima facie candidates for signaling) will be too personally harmful, and, thus, not be selected-for. Extreme benevolence, whether or not it is reciprocated, might be an example. In any signaling game, we should compare the benefits the sender and receiver expect to gain from signaling against the costs of production. For example, the belief that mask wearing is ineffective (or effective) at containing the spread of COVID-19 might function to communicate one’s political allegiances. While there are social benefits to communicating this affiliation to the likeminded, there are social costs (ostracizing others) and the great epidemiological costs of acting on an unjustified belief. That being said, things are not so simple as belief signaling occurring whenever the signaling benefits exceed the production costs.

It is not immediately clear how likely it is that mindreading will lead to belief signaling, since not every belief that can be detected by others and thereby used to manipulate their behavior comes to function as a signal. In fact, few beliefs do. But once beliefs are capable of being detected by mindreaders, they enter the pool of candidates for signaling. Pretty much any trait could function as a signal, so long as the trait can be detected and is capable of meeting any requisite honesty standards (e.g., by being costly to produce, difficult to fake, etc.). At a minimum, mindreading allows for beliefs to function as signals. Belief signals can then emerge in one of four ways: (1) the more prominent display of a previously existing belief (without altering it), (2) the skewing of a previously held belief so that it becomes more exaggerated and noticeable, (3) the reshuffling of believers such that (for the most part) only those who would wish to communicate the signaled information come to have the belief, or (4) the novel adoption of a belief solely for signaling purposes (Funkhouser, 2020). These are all possible ways for beliefs to solicit detection by others and manipulate their behavior.

A stronger claim is that beliefs are particularly good candidates for signaling, so once they can be detected by mindreaders we should expect belief signaling. The idea here is that beliefs are not just in the pool of candidates for signaling, but that they rise to the top. The suitability for signaling will vary by informational content, especially given the other candidates for signaling and their abilities to meet the requirements of efficacious and honest communication. Beliefs can be particularly good signaling candidates for some contents if it is not viable for other traits to efficaciously or

10 The issue of honesty mechanisms—like costly signaling—for belief signaling is a pressing and difficult problem. I will not address it here, however.
honestly convey that information. The fact that beliefs have propositional content that
mindreaders can detect and then incorporate into a worldview shows that beliefs are
inherently suited for conveying complex and nuanced information. A piece of clothing
or jewelry could be used to signal one’s good will, but it is not as reliable as a genuine
belief in the value of good will (or a genuine belief in solidarity, etc.). Anyone can put
on a shirt or necklace. Of course, anyone can assert something, too. For this reason,
belief signaling depends on mindreading vigilance on the receiver side rather than
passively accepting words at face value. Of course, people in the real world tend to
meet this standard.

The factors, then, that bear on whether beliefs are selected for signaling purposes
include:

1. Can the belief reliably and efficaciously communicate the desired information?
2. Are other candidate signals available that are more reliable, less costly, or otherwise
   preferable? Can one instead just wear a red hat or lie in order to signal political
   allegiance, rather than fully believe? Belief signaling will likely occur only if
easier, more credible, less costly, or more perspicuous methods are not at hand.
3. Are the epistemic costs—both theoretical and practical—of distorting belief for
   signaling purposes too great?
4. Even if the message is honestly and efficaciously broadcast, does the receiver have
   an incentive to detect the signal and allow for the manipulation? Otherwise, the
   receiver would be better served ignoring the signal altogether. In these situations,
   we should expect agents to eventually evolve toward this solution, leading to the
   signal fading from use.

4.3 Examples in natural and artificial systems

Human beings might be the only animals capable of mentalistic mindreading, 11 with
well-functioning individuals possessing this ability in their second year and integrating
it with their other cognitive abilities over the next few years (Carruthers, 2013). We
have turned to pro-social and tribal beliefs as prime areas for belief signaling, but
self-enhancement and other types of positive illusions could function as signals as
well. For example, people tend to have an overly positive self-conception (Taylor &
Brown, 1988). Such self-enhancement has typically been construed as for the sake of
self-esteem, which could be interpreted as a form of self-signaling (Levy, 2018). But
it could also be a form of social signaling so that others are more likely to hire, date, or
deer to us. Looking beyond the individual, social and corporate institutions also have
belief-like states that function as signals—e.g., policies, values, and manufacturing
projections. Corporate values and branding can signal a lifestyle which leads to sales;
corporate projections can signal confidence which boosts investment.

Artificial devices can evolve or be designed to mimic the mindreading functionality
found in humans. Self-driving cars could detect the internal representations—belief-
like and goal-like states—that generate the driving behavior in other cars, perhaps

---

11 Though, Kano et al. (2019) argue that other great apes can sometimes go beyond behavior-reading to
outright mentalistic mindreading.
because they share an operating system. Taking this into account, other cars might shape (through learning or programming) their representations of the environment in virtue of how these very representations are detected. These representations could be broadcast as signals so as to create homogeneity (e.g., an industry standard). The Open Automotive Alliance has this goal, pushing for the Android platform to be the industry standard for the automotive ecosystem.\textsuperscript{12} Or consider weapons systems that represent the enemy’s (human or machine) mental states, with mutual recognition of these representations. Given the high stakes, game-theoretic reasoning underlying them, the signaling pressure can be significant. We should expect these to be even more pronounced than for natural systems, as they can be strategically designed for these purposes or become such through machine learning. These processes can operate on a much faster timeframe than can natural selection (biological or cultural). As we head toward robust AI, we should expect ever more strategic social-psychological signaling in conjunction with mindreading by artificial systems (Scassellati, 2002).

Blockchains are excellent examples of the evolution of signaling systems in digital ecosystems. The represented information is displayed as a public ledger. This public display of information functions so as to communicate the transaction history and, further, to signal trust in the ledger—and in the case of cryptocurrencies, the coin—itself. With Bitcoin, the signal is judged to be honest because of the costliness of the work (proof-of-work). This trust is far from trivial—the blockchain signaling justifies the investment of billions of US dollars. For other cryptocurrencies, the costliness takes different forms—e.g., proof-of-stake. We should see the different blockchain consensus mechanisms as honesty mechanisms for a signaling game. The representations (the blockchain itself) exist so as to be advertised, and they are valuable only if they are accepted as being honest. Blockchains can be particularly suggestive for those of us seeking evolutionary explanations of honest signaling, since they are devices for honest signaling absent any centralized authority or preexisting trust among members. It very well could be the case that the honesty mechanisms for representations in human psychology and culture have analogues in blockchains.

5 Conclusion

In sum, we have analyzed three different levels of psychological development and belief-shaping due to social forces:

1. Rational navigation, in which beliefs are shaped only by their success at solipsistic goal pursuit. The selection pressure includes how well the beliefs fare given the (undetected!) beliefs and mental states possessed by others.
2. Rational navigation plus behavior-reading, in which beliefs are sometimes shaped by how they fare given that others detect the purposive behavior (“as such”) that is driven by those beliefs. Here, we find a previously unacknowledged level of signaling—intention signaling.

\textsuperscript{12} https://www.openautoalliance.net/#about.
3. Rational navigation plus mentalistic mindreading, in which beliefs are sometimes shaped by the fact that others detect them “as such.” Such beliefs function as signals which manipulate the behavior of others.

What could possibly be the next stages? The mindreading might be iterated with ever greater anticipatory effects, much like a sophisticated game of chess. We need not worry about that, though—there is enough room for exciting research with the social signaling that emerges at levels 2 and 3. The important point is that signaling pressure shapes the direction of thought. We should not expect the mind to be shaped simply for navigation, but also for signaling. The mind is differentially shaped as people become better at reading it.\(^\text{13}\)

Declarations

Conflict of interest None.

References

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a ‘theory of mind’? *Cognition, 21*(1), 37–46.

Bermúdez, J. L. (2009). Mindreading in the animal Kingdom? In R. Lurz (Ed.), *The philosophy of animal minds* (pp. 145–164). Cambridge: Cambridge University Press.

Breazeal, C., & Scassellati, B. (2002). Robots that imitate humans. *TRENDS in Cognitive Sciences, 6*(11), 481–487.

Bullock, J., Gerber, A., Hill, S., & Huber, G. (2015). Partisan bias in factual beliefs about politics. *Quarterly Journal of Political Science, 10*, 519–578.

Carruthers, P. (2013). Mindreading in infancy. *Mind & Language, 28*(2), 141–172.

Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science, 1*(2), 255–259.

Davidson, D. (1984). *Inquiries into truth and interpretation*. New York, NY: Oxford University Press.

Dennett, D. (1978). *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge, MA: The MIT Press.

Dennett, D. (1987). *The intentional stance*. Cambridge, MA: The MIT Press.

Dunbar, R. (1998). The social brain hypothesis. *Evolutionary Anthropology, 6*(5), 178–190.

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews, 24*, 581–604.

Frankish, K. (2004). *Mind and supermind*. Cambridge: Cambridge University Press.

Funkhouser, E. (2017). Beliefs as signals: A new function for belief. *Philosophical Psychology, 30*(6), 809–831.

Funkhouser, E. (2018). Detection, not perception: A reply to glazer. *Philosophical Psychology, 31*(7), 1120–1125.

Funkhouser, E. (2020). A tribal mind: Beliefs that signal group identity or commitment. *Mind & Language. https://doi.org/10.1111/mila.12326.*

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The Naïve theory of rational action. *TRENDS in Cognitive Sciences, 7*(7), 287–292.

Jara-Ettinger, J., Gweon, H., Tenenbaum, J., & Schulz, L. (2015). Children’s Understanding of the Costs and Rewards Underlying Rational Action. *Cognition, 140*, 14–23.

\(^{13}\) I would like to thank the two anonymous referees who provided me with very thorough and helpful comments on my original submission. I especially appreciate their patience with my scholarship as they encouraged work that they judged to be innovative.
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.

Krebs, John and Dawkins, Richard. 1984. “Animal Signals: Mind-reading and Manipulation,” in Behavioural Ecology: An Evolutionary Approach, eds. J.R. Krebs and N.B. Davies (Oxford: Blackwell Scientific Publications).

Lake, B., Ullman, T., Tenenbaum, J., & Gershman, S. (2017). Building machines that learn and think like people. Behavioral and Brain Sciences, 40, 1–72.

Leavens, D., & Hopkins, W. (1998). Intentional communication by chimpanzees: A cross-sectional study of the use of referential gestures. Developmental Psychology, 34(5), 813–822.

Levy, N. (2018). Showing our seams: A reply to Eric Funkhouser. Philosophical Psychology, 31(7), 991–1006.

Levy, N. (2020). Rationalization enables cooperation and cultural evolution. Behavioral and Brain Sciences, 43, 28–29.

Millikan, R. (1984). Language, thought, and other biological categories. Cambridge, MA: MIT Press.

Onishi, K., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? Science, 308(5719), 255–258.

Petersen, M., Osmundsen, M., & Tooby, J. (2021). The evolutionary psychology of conflict and the functions of falsehood. In D. Barker & E. Suhay (Eds.), The politics of truth in polarized America. Oxford: Oxford University Press.

Rasouli, A., & Tsotsos, J. (2018). Joint attention in driver-pedestrian interaction: From theory to practice. arXiv:1802.02522.

Reynolds, C. (1987). Flocks, herds, and schools: A distributed behavioral model. Computer Graphics, 21(4), 25–34.

Richerson, P., & Boyd, R. (2005). Not by genes alone: How culture transformed human evolution. Chicago, IL: University of Chicago Press.

Scassellati, B. (2002). Theory of mind for a humanoid robot. Autonomous Robots, 12, 13–24.

Schwitzgebel, E. (2019). Belief. In E. Zalta (Ed.), Stanford encyclopedia of philosophy, https://plato.stanford.edu/archives/fall2019/entries/belief/. Accessed 16 Sep 2021.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. Current Biology, 18, 668–671.

Spence, M. (1973). Job market signaling. The Quarterly Journal of Economics, 87(3), 355–374.

Sterelny, K. (2015). Content, control, and display: The natural origins of content. Philosophy, 43, 549–564.

Taylor, S., & Brown, J. (1988). Illusion and well-being: A social psychological perspective on mental health. Psychological Bulletin, 103(2), 193–210.

Tooby, J., & Cosmides, L. (2015). The theoretical foundations of evolutionary psychology. In D. M. Buss (Ed.), The Handbook of evolutionary psychology (2nd ed.). Hoboken, NJ: Wiley.

Wedgwood, R. (2002). The aim of belief. Philosophical Perspectives Language and Mind, 16, 267–297.

Williams, D. (2021). Socially adaptive belief. Mind and Language, 36(3), 333–354.

Zahavi, A. (1975). Mate selection—A selection for a handicap. Journal of Theoretical Biology, 53, 205–214.

Zawidzki, T. (2013). Mindshaping. Cambridge, MA: The MIT Press.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.