The Origins of Distinctively Human Mindreading: A Bio-Social-Technological Coevolutionary Account

Armin W. Schulz

Abstract

Humans stand out from other organisms in their ability to mentalize. They attribute mental states with high degrees of abstractness and complexity, and they do so frequently in situations of major importance. However, their mindreading abilities are also less than fully reliable and culturally variable. This paper proposes a novel account to explain these facts. Since human mindreading is cognitively very costly, cognitive and social tools are needed to make it work efficiently (or at all), even given an evolved psychological machinery for mindreading. The upshot of this is a feedback loop involving mindreading, cultural learning, and socio-technological development. In turn, acknowledging the existence of this feedback brings to view hitherto overlooked areas of investigation concerning human mindreading, including the impact of different forms of socio-cognitive technology on differences in human mindreading, and the development of new technologies to aid human mindreading.

1. Introduction

One of the ways in which humans stand out from other organisms is in their ability to figure out what others are thinking—i.e. to "mindread."[1] While somewhat culturally variable and not fully reliable, human mindreading is frequently used in situations of major adaptive importance, and can involve highly complex and abstract mental states.[2] What is less clear, though, is how to explain the origins of this ability. What factors underlie the evolution of distinctively human mindreading (Tomasello & Herrmann, [2010]; Penn et al., [2008]; Andrews, [2012])?

This paper develops a novel account that answers these questions. This account goes beyond the existing biologically-adaptationist (Carey & Spelke, [1996]; Carey, [2011]; Carruthers, [2006]; Nichols & Stich, [2003]) and cultural-learning based ones (Heyes, [2018]) and argues that, since human mindreading is cognitively very costly, further cognitive and social tools are needed to make it work efficiently (or at all). This implies that human mindreading is embedded in a positive feedback loop with cultural learning and socio-technological development—the acknowledgement of which brings to view overlooked areas of investigation.

[1] In this paper, I use the terms "mindreading" or "mentalizing" to refer to the ability to attribute mental states to others. To what extent this is a theory-like or simulationist process is left open here (which is why the term "theory of mind" is avoided). For more on this, see e.g. Nichols and Stich ([2003]); Goldman ([2006]); Saxe et al. ([2006]).
[2] The human ability to mindread may also underly another human uniqueness: religiosity (Boyer, [2001]; Atran, [2002]; Sperber, [1975], [1996]; Henrich, [2020]).

The paper is structured as follows. In section 2, the nature of human mindreading is laid out in more detail. Section 3 presents the key existing accounts of human mindreading and makes clear that these accounts leave several key issues unexplained. On this basis, section 4 develops a new account of the origins of distinctively human mindreading, and contrasts it with two further accounts of the evolution of human mindreading that also go beyond the biological-adaptationist and the cultural-learning based ones. Section 5 concludes.

2. Human Mindreading

What makes human mindreading unique? It is not that humans are the only organisms that can do any kind of mindreading. For example, there are good reasons to think that chimps can determine what others can see and that they can empathize with others (Tomasello et al., [2005]; Tomasello & Herrmann, [2010]; Andrews, [2012]), and something similar holds for dolphins, elephants, and magpies (Parker et al., [1994]; Suddendorf & Butler, [2013]; Toda & Platt, [2015]). It is true that these findings are somewhat controversial still (see e.g. Povinelli, [2003]; Penn et al., [2008]), but, for present purposes, it is sufficient to note the following two points.

First, it is very plausible that non-human animals can engage in many forms of non-conceptual or non-representational "low-level" mindreading, such as emotion attribution, behavioral mirroring, and other forms of implicit mental state tracking (Nichols & Stich, [2003]; Goldman, [2006]; Zawidzki, [2011]; Apperly & Butterfill, [2009]; Butterfill & Apperly, [2013]; Edwards & Low, [2017]; Heyes, [2018]).[3] For example, many non-human animals seem to show empathy or sympathy, and can at least track some perceptual and cognitive states (de Waal,

[3] The distinction between low- and high-level mindreading is far from clear (Goldman, [2006]; Schulz, [2011]). It is here only used for expository purposes, though, and is not central to the argument.

[2008]; Churchland, [2011]; Acebo & Thoman, [1995]; Bowlby, [1958]; MacLean, [1985]; Schulz, [2017]; Apperly & Butterfill, [2009]; Edwards & Low, [2017]).

However, these "sub-mentalizing" (Heyes, [2018]) forms of mindreading will not be central in what follows. This is not because they are not significant, or because there are no human / non-human differences whatsoever in them. Rather, it is because they deserve an inquiry of their own: the low-level mindreading system is sufficiently different from the high-level one—to be laid out in more detail momentarily—to call for an investigation of its own (Apperly & Butterfill, [2009]; Edwards & Low, [2017]). Given this, the focus here is solely on the high-level system: that is, the focus here is on the explicit representation of the mental states of others. This is an inherently interesting topic, and it also marks the strongest contrast between human and non-human forms of mindreading (Apperly & Butterfill, [2009]; Edwards & Low, [2017]; Heyes, [2018]).

Second, though, it is further important to note that it is possible to simply accept here that some non-human animals can engage in some forms of this "high-level"—i.e. conceptual and representational—mindreading as well. Apart from the fact that this assumption is empirically not implausible, it does not bias the discussion here. The point in what follows is to show that there is still something uniquely human when it comes to "high-level" mindreading; all of the arguments that follow will go through even if it were to turn out that non-human animals cannot engage in "high-level" mindreading after all. In particular, there are four key features of high-level human mindreading that makes it stand out from the mindreading abilities of other animals.

2.1. Attributive Sophistication

Humans can and do attribute mental states to others in ways that are cognitively sophisticated. There are several dimensions to this cognitive sophistication.

(a) High orders of intentionality

Humans frequently attribute mental states at the third or even fourth level of intentionality: for example, I might attribute to you the hope that I think that you think the treasure is hidden over there (because I think you are trying to deceive me), or I might attribute to you the fear that I think that you think that I am arrogant (due to a Jane Austen-like misunderstanding, say) (Sterelny, [2003]; Bennett, [1991]). Indeed, if something like the language of thought hypothesis turns out to be true, humans can attribute thoughts of arbitrary levels of intentionality (Fodor, [1975]; Fodor & Pylyshyn, [1988]). By contrast, it is not clear whether non-human animals even attribute second-order thoughts to others (Bennett, [1991]).

(b) Mental state concept complexity

Apart from perceptual states, belief states, and basic emotional states like fear and hunger, humans can *also* attribute more complex mental states that are composed out of other, not necessarily atomic, mental states ("he is just hangry;" "that's just wishful thinking on her part"). By contrast, non-human animals seem to generally attribute less complex mental states only (Hare et al., [2000]; Kanngiesser et al., [2020]; Tomasello et al., [2005]; Tomasello & Herrmann, [2010]; Andrews, [2015], [2012]).

(c) Mental state concept abstractness

Humans can and do attribute mental states that are highly abstract, and which lack straightforward links to behavior. So, humans can attribute states like *feeling slighted*, *nostalgia,* or *proud*. These are not states that are the outcomes of easily determinable circumstances (like perceptual states), or which are tied to specific behavioral patterns, as is true for the basic emotions (LeDoux, [2012]; Sauter et al., [2010]; Ekman, [1989]; Ekman & Rosenberg, [1997]). Non-human animals appear to mostly attribute less abstract mental states (Bennett, [1991]; Allen, [1999]).

2.2. Attributive Importance

Humans use mental states attributions very often in adaptively importantly decisions, and these attributions are often crucial factors determining the outcomes of these decisions.[4] It is part and parcel of the nature of human sociality that many of our interactions are based on our intentions and other mental states. We live in groups consisting of kin as well as non-kin and are utterly dependent on each other for our survival. Hence, successful social interaction requires being attuned to quickly shifting social conditions (alliances, etc.). In turn, this kind of social interaction requires or at least rewards an understanding of others' mental states (Warneken et al., [2011]; Tomasello, [1999]; Sterelny, [2003], [2012]; Henrich, [2015]; Henrich & McElreath, [2011], [2007]; Boyd & Richerson, [2005]; Cosmides & Tooby, [1992]).

Of course, this is not to say that mental state attributions are not important for non-human animals, too (Andrews, [2012]). On the contrary, what others can see, where they think food is hidden, or even what they think I can see or know about where food is hidden are important

[4] Note that claim is relativized to adaptively important decisions—the frequency of human mindreading per se is a point of contention in the literature (Apperly & Butterfill, [2009]; see also Gallagher, [2015]). That is, the claim here is just that human mindreading is unique, among other things, for often being a key element in adaptively important decisions. This is consistent with mindreading not being a major factor in human decision-making in general.

elements in animal decision making, too. However, the key point here is just that mindreading is a particularly central facet of adaptive human decision-making (Tomasello & Herrmann, [2010]; Warneken & Tomasello, [2006]).

2.3. Cultural Variability

Human mindreading is not culturally homogeneous. While mindreading seems to be a part of all human cultures (Henrich, [2020], pp. 67, 76, 129; Tomasello, [2021]), different cultures still differ somewhat over how to mindread others (Luhrmann et al., [2015]), in the emphasis they put on mindreading others (Henrich, [2020]), and in the development of mindreading (Shahaeian et al., [2011]; Mayer & Träuble, [2013]). Now, it is important to note that making empirically well-grounded cross-cultural comparisons of mindreading capacities and practices is very hard.[5] For example, the interpretation of linguistic data on belief attribution is made difficult by fact that the very concept of "belief" has different senses in different languages, some of which are psychologically richer than others (see e.g. Lavelle, [2021]). However, as matters stand, there is at least good reason to think that there is some significant cross-cultural variability in mindreading capacities and practices—though the details require further work and elucidation.

For example, in individualist ("WEIRD"—western, educated, industrialized, rich, and democratic) cultures, children seem to be able to attribute beliefs different from their own before they can determine what others know; children from more holistic (non-WEIRD) cultures show the opposite pattern (Shahaeian et al., [2011]). Similarly, in some cultures, it is more common to think of others as acting in line with social role expectations rather than stable character states, and in other cultures the opposite is the case (Henrich, [2020]). For example, in South Korea, a

[5] I thank an anonymous referee for useful discussion of this point.

young person contradicting an elder might be taken to be in an angry state of mind—given that this is a violation of social role expectations about deference to elders—whereas in the US, someone like that might be taken to be in a relaxed, comfortable, or playful state of mind—given social role expectations of people having open discussions with people they respect, value, and feel comfortable with, regardless of their age. An account of distinctively human mindreading thus needs to provide an explanation of the fact that there are *both* cross-cultural similarities *and* differences in human mindreading.[6]

2.4. Attributive Unreliability

Human mindreading should not be taken to be perfectly reliable (Andrews, [2018]; Nisbett & Wilson, [1977]; Spaulding, [2018b], [2018a]). This is evidenced by much of human art and history: the latter are not just full of instances of mindreading playing a major role in people's lives, but also of mindreading *failures*. (Consider the polyandrous marriage of Draupadi and the Pandava brothers in the Mahabharata: this rests on a misunderstanding of what the Pandava brothers' mother wanted—which is in turn due to *a misunderstanding* of what Arjuna was trying to tell her.) Theoretically, this is not surprising, since mindreading others is a non-trivial inference problem. As noted earlier, similar kinds of behaviors can result from very different mental states, so that figuring out what someone else is thinking from looking at their behavior contains many unknowns and uncertainties (Lurz, [2018]). This also explains why mindreading

[6] Note that this issue crosscuts the question of whether mindreading is simulationist or non-simulationist in nature (Nichols & Stich, [2003]; Goldman, [2006]). Even if we all used the same decision-making processes, and even if these same processes are key components in mental state attribution, there can be cultural variation in mindreading, as there may be cultural variation in what inputs we use for our simulations, as well as in how we interpret the outputs of that simulation (Rossi, [2014]). This point becomes important again in section III below.

is cognitively costly, and takes significant resources in terms of time, concentration, and attention (Schaller et al., [2007]).

This is not to say that we are hardly ever able to form reasonably accurate representations of the thoughts of others (see also Westra, [2020]). Rather, the point here is just that humans cannot be seen to *always* mindread others in a *completely* error-free manner. At times, we find others' behavior inscrutable and at other times we misattribute thoughts to them. This less than full reliability in mental state attributions needs to be kept in mind when formulating an account of the distinctive nature of human mindreading.[7]

3. Explaining the Evolution of Distinctively Human Mindreading: Biological Adaptations vs. Cultural Learning

There are several influential accounts that have been put forward to explain the evolution of distinctively human mindreading. Before laying out these accounts, though, it is important to stress again that these accounts focus on only one aspect of the human mindreading system: the one dedicated to the explicit representation of other's mental states. As noted earlier, this explicit, representational mindreading system should be seen as layered on top of an implicit mindreading system that humans may share with some non-human animals (see e.g. Apperly & Butterfill, [2009]; Edwards & Low, [2017]). What follows is thus restricted to the former system only. That is, what follows solely concerns mindreading with the four sets of features sketched in the previous section.

The first influential account of this kind of explicit, high-level distinctively human mindreading sees it as the result of the evolution of a dedicated neuropsychological mechanism

[7] This is something mindreading shares with many other mental capacities, from perception to episodic memory.

that has been selected for enabling humans to deal with the particular social environments in which they live. Several different versions of such an account can be distinguished.

The major one of these sees human mindreading abilities as the result of selection pressures surrounding the fact that humans needed to navigate a special kind of social environment (Byrne & Corp, [2004]; Whiten & Byrne, [1997]; Sterelny, [2012]; Humphrey, [1986]; Leslie, [1994]; Baron-Cohen, [1997]). As noted in earlier in section 2.2., adaptively accomplishing this navigation requires humans to make decisions that take into account what others want, think, and feel. Importantly, this is true even in relatively "tight" societies, in which human action is highly constrained by its social environment (Elster & Gelfand, [2021]; Henrich, [2020]). There is a myriad of factors that can influence human decisions. Human cultures are complex and allow for many different social roles that can change over time. This means that individual humans need to figure out how to balance the different roles and relationships they have, even if their decisions are mainly the result of the social roles they inhabit and the social relations they have (Hollis, [1994], chap. 8). The fact that person P was a reliable foraging partner yesterday does not mean that P will also be one such today: P may have obtained other social obligations or requirements in the meantime that imply that P will not be highly focused on this particular task. To deal with this kind of environment, it is therefore advantageous to be able to understand and predict the mental states of others. It is these mental states—intentions, emotions, expectations, etc.—that drive their behaviors, and which determine how they manage their different and dynamically changing roles.[8]

[8] It is noteworthy that it is not clear whether early human societies were looser or tighter. On the one hand, it seems that the "WEIRD" societies we encounter today are derived from tighter ancestral human cultures (Henrich, [2020]). However, since cultural institutions themselves culturally evolved, it is not clear what is true about earlier hominin (e.g. Australopithecus) societies. Also, chimp societies seem relatively looser (Byrne & Corp, [2004]; Whiten & Byrne, [1997]; Brosnan et al., [2008]). At any rate, as noted in the text, this does not materially affect the conclusions here reached.

According to this account of the evolution of human mindreading, it is precisely this advantageousness that has led to the evolution of a specific set of psychological expectations. These expectations include representations about (i) when to mindread (e.g. if the focus is the behavior of other people, rather than that of insects); (ii) what to consider when mindreading (e.g. where someone looks and what they did in the past), (iii) what some of the possible types of states are that could drive the behavior in question (e.g. beliefs, motivations, emotions), and (iv) some specific candidates of such states to consider (e.g. that the person believed that no two physical objects can occupy the same place at the same place and time, that parents are motivated to help their kids, or that the person wants to do something jointly with me) (Carey, [2011]; Carey & Spelke, [1996]; Carruthers, [2006]; Tooby et al., [2005]; Tomasello, [2021]; Tomasello et al., [2005]).

This evolved set of psychological expectations makes it possible for us to infer what others are thinking in a way that would not otherwise be possible. Without (i), we might be mindreading at the wrong times or fail to mindread at the right times—e.g. we might try to mindread a thunderstorm or fail to mindread our spouse.[9] Without (ii), we might be stuck considering too many possibilities as to what to pay attention to: are my brother's mental states a function of his eye color? Without (iii) and (iv), we might be unsure as to what sort of thing to attribute to someone else: was their behavior caused by a stack overflow error, a desire for drinking a can of paint, or a mistaken belief that this glass contained water? Flipping this around: given expectations like (i)-(iv), we narrow the inferential space when trying to predict and

[9] As suggested in note 2, mindreading expectations might underlie religious beliefs, and so we might interpret a thunderstorm as an expression of Zeus's wrath. However, it remains true that humans do not mindread indiscriminately.

explain the behavior of other people. In turn, this makes us better able to handle the pressures of social living.

A slightly different account of the evolution of human mindreading—though one that shares many key features with the above account—sees the latter abilities as the result of the fact that it was biologically advantageous for humans to be good communicators (Sperber, [1975], [1996]; Godfrey-Smith, [2002]). Determining what others are thinking, on this account, is important not so much because it enables us to anticipate their behavior, but because it makes it possible for us to influence each other's thoughts (and thus, their behavior) through communication.

Communication was key in the kinds of social environments humans evolved in, as it enables us to learn from each other and cooperate in the ways that we do—involving e.g. complex division of labor and trade. Importantly, though, it is not implausible that this kind of communication requires sophisticated mindreading abilities. An easy way to see this is by considering the account of meaning in Grice ([1989]). On this account (roughly), for a person A's utterance of *p* to mean "grass is green," A has to intend for another person B to form the belief that A is, in fact, intending B to form the belief that "grass is green" (and A and B have to know this about each other). To recognize you as a communicator, I need to recognize that your communicative behavior has a certain fixed relationship to the world, and that you *want* me to *think* it does.

Now, it needs to be acknowledged that this is not the only account of communication in the literature, and that other such accounts do not presuppose that human communicators must be sophisticated mindreaders (Millikan, [1984]; Gauker, [2003]; Geurts, [2019]; Moore, [2018]; Skyrms, [2010]; see also Brandom, [1994]).[10] Fortunately, it is not necessary to settle this debate

[10] Note while the account of Skyrms ([2010]) can be used to spell out the evolution of a Gricean picture of meaning, the former is broader than that, and does not require the latter (see also Moore, [2018]).

here. The point here is just that one influential attempt to ground the evolution of sophisticated human mindreading appeals to the view that humans needed to be successful communicators, and that successful communication requires the mutual recognition of our mental states. If this view of the nature of communication is denied, the problems below can be avoided, but only at the cost of finding another route towards the evolution of sophisticated human mindreading—such as the first account sketched above, or the cultural-learning-based account that will be sketched momentarily.[11]

In sum, therefore: what both versions of the standard account of the evolution of human mindreading abilities have in common is the idea that there existed biological selection pressures for these kinds of abilities. The psychological details of the machinery that is hypothesized to have evolved in this way will be made clearer in the next section below. For now, what is important to note is that this account does well in explaining the first two features of human mindreading noted in the previous section: *attributive sophistication* and *attributive importance*.

In particular, the hypothesis that there was selection for mindreading—whether for predicting the behaviors of others in one's culture or for communicating with them (or both)—predicts that human mindreading abilities would be (a) sophisticated, and (b) a key element of adaptively important decisions. Point (a) is a consequence of the very starting place of this kind of account: given that humans need to predict the behavior of and / or communicate with others that do the same thing, and who react to the world in otherwise complex ways, too, requires that they must attribute mental states of high orders of intentionality, complexity, and abstractness (see also Sperber & Wilson, [1986]). For example, we are not just making our behavior dependent on the weather—we are making our behavior dependent on what others think about the weather (or

[11] I thank an anonymous referee for useful discussion of this point.

even what others think about our attitudes towards the weather, as when I try to impress you by wearing shorts during a snowfall). Similarly, effective and efficient communication might require the use of metaphor: we might respond to someone saying "Robert is a bulldozer" by attributing to them the thought that Robert has a stubborn and forceful personality—not that he is literally a machine. Point (b) also follows straightforwardly from the present account of the evolution of mindreading: after all, on this account, helping us make adaptively important decisions well is precisely why these abilities are said to have evolved in the first place.

These points gain further support from the facts that, as far as we can tell, all human cultures mentalize in some form or other (Henrich, [2020], pp. 76-77), and that basic mindreading abilities seem to mature early (Baillargeon et al., [2010]). These kinds of similarities are exactly what we would expect if human mindreading were the result of biological selection pressures deep in the human lineage—for example, they match what is true about the ability to speak a language or to walk bipedally. That is to say, the assumption that mindreading has been selected for in human evolutionary history predicts that this is a trait that is deeply embedded in our pan-specific human psychology, and that it is therefore culturally universal and quite canalized in its development. The fact that the latter two are indeed the case thus lends credence to the hypothesis of selection for human mindreading (Forster & Sober, [2011]; Sober, [2008]).

However, this hypothesis also faces some challenges. In particular, the hypothesis does far less well when it comes to explaining the cultural variability of human mindreading. If human mindreading abilities are a selected response to the pressures of social living, it is not clear how to explain the existence of significant variation in human mindreading across cultures.

On the one hand, it is not plausible to see human mindreading practices as having evolved so as to be finely tuned to the exact details of how people in different cultures think. This does not

match the fact that culturally specific genetic-biological adaptations are rare. In general, cultures change too fast to make it possible for natural selection to lead to the evolution of traits that are specific responses to the features prevailing in a given culture (Boyd, [2018]; Boyd & Richerson, [2005], [1985]; Henrich, [2015], [2020]; Henrich & McElreath, [2011], [2007]). While exceptions to this do exist (Boyd & Richerson, [1985]; Zeberg & Pääbo, [2021]), these are uncommon and do not match the frequency of cultural variation in human mindreading.

On the other hand, the situation is not resolved by appealing to an analogy with natural languages. It is true that, even though there is considerable variation across cultures over *which* language a human speaks, the ability to speak *some* language plausibly has been selected for in human evolutionary history (Pinker & Bloom, [1990]). Could we not say the same thing about mindreading? In particular, could we not see culturally different mindreading practices as different mindreading "languages" that are however anchored in a common "universal mindreading grammar" that is the product of natural selection deeper in the human lineage?

The problem with this response is that, while perfectly coherent as such, it leaves open much of what needs to be explained here. How do people in different cultures determine the locally appropriate mindreading practices? The analogy with language is somewhat misleading here, in that children receive a lot of linguistic input with which to determine which linguistic culture they are in. With mindreading, though, this is less clear. What mindreading inputs do they receive that enable them to narrow down their mindreading culture? Note that, as will be made clearer in the next section, it is possible to answer these questions. The point here is just that pointing to the (supposed) existence of evolved mindreading expectations is (at best) part of the explanation of the cultural differences in human mindreading. Such evolved expectations are not

inconsistent with cultural variability in mindreading; by themselves, though, they do not provide an account of the latter. For this, something else is needed.

Given these challenges for the above account of the evolution of human mindreading, it is tempting to look for alternatives. A major such alternative is the second key account of distinctively human, high-level explicit mindreading: an account that sees the latter as the result of cultural learning and not of biological evolution (Heyes, [2012], [2018]; Heyes & Frith, [2014]; Moore, [2021]).[12] On this picture, distinctively human high-level mindreading is still crucially related to the social nature of human living, but in a different way from the sort of accounts sketched above: it is the result of the fact that humans are strong *cultural learners*. That is, on the cultural-learning-based account, it is not the case that human mindreading abilities are themselves biological adaptations; rather, they are learned from others in the culture a person lives in. As Heyes ([2018]) puts it: mindreading is like book reading. On this account, therefore, the explanation of human mindreading abilities lies in the fact that humans are standout cultural learners—there was no specific biological selection on human mindreading abilities. These abilities are simply culturally learned, like riding a bicycle.

It is important to note that this alternative account is not distinctive in seeing human mindreading as the product of *some kind of* learning. The standard, biologically-adaptationist account can allow for learning, too (Tooby & Cosmides, [1992]). In general, there will be many environmental influences on the maturation of any human abilities—mindreading ones included. Some of these environmental influences are likely to be learning-based ones. For example, the

[12] A quick point of terminology: since cultural learning has also evolved by biological evolution, and since biological and cultural evolution often affect various traits simultaneously, contrasting one with the other can seem somewhat problematic (Henrich & McElreath, [2011], [2007]; Boyd & Richerson, [2005]). However, here, the point is just terminological: of course, the cultural learning-based account is also biological in a wider sense. However, the point at issue here is whether human mindreading abilities are the result of a canalized biological maturation process akin to beard growth or language-acquisition, or whether they are the result of a cultural learning process akin to book reading—whatever the biological origins of the latter are.

evolved mental structures the above, biologically-adaptationist accounts posits as underlying mindreading could in fact mostly function to scaffold our learning in such a way that we end up acquiring the kinds of mental representations and processes that allow us to successfully mindread others (Tooby & Cosmides, [1992]; Carey & Spelke, [1996]). In this way, this picture of the evolution of mentalizing can be made consistent with the idea that children learn to mindread by testing and revising various "hypotheses" about how minds work (Gopnik, [1996]), or by learning to put themselves in the situation of others and then interpreting the results of their own simulated cognitions (Goldman, [2006]).[13] In particular, which hypotheses are tested, and how children determine what they are thinking, can be underwritten with innate mental structures (Nichols & Stich, [2003]; Schulz, [2011]).[14]

In contrast to this "scaffolded-learning" picture, the cultural-learning based account of human mindreading sees the latter ability as culturally learned *in the same way that* book-reading or bicycle-riding are learned: through a *domain-general* learning process (Heyes, [2018], [2016]; Henrich, [2015]; Csibra & Gergely, [2011]). That is, on this account, humans learn mindreading abilities by directly or indirectly observing others in their culture (who may even be explicitly engaged in teaching them the relevant skill). Their learning is not scaffolded by any specific mental structures, but rather it is the result of a process that underlies all other instances of cultural learning as well.

It is furthermore noteworthy that this account is perfectly consistent with the facts that human mindreading may be susceptible to being selectively damaged (Baron-Cohen, [1997]) or that it

[13] It is also possible to interpret either of these views as being based on *purely* on individual learning, without any innate priors (or the like) grounding this learning. In that case, though, it is not clear why there is so much cultural and even cross-cultural homogeneity in human mindreading. This makes this kind of view less compelling than either an innately grounded one or a cultural-learning based one.

[14] Again, the ability to speak a language may be a good example of this sort of picture (Pinker & Bloom, [1990]): the biological adaptation underlying this ability is precisely a set of psychological expectations and processes that make the learning of language possible. See also note 12.

may be underwritten by a specific set of neural structures (Goldman, [2006]; Saxe & Powell, [2006]). Cultural learning can lead to the development of specific brain structures that underlie the learned activities, and which can thus be selectively damaged. Indeed, the same is true for book reading (Henrich, [2020], pp. 3-7; Heyes, [2018]).

It is also clear that this account does well at explaining features 3 and 4 of human mindreading, and that it does so better than the biologically-adaptationist account does. In the first place, the cultural learning-based account straightforwardly *predicts* the existence of cultural variation in human mindreading (Heyes, [2012], [2018]). On the assumption that mindreading skills are learned from our culture, general cultural differences are likely to translate into mindreading differences as well. More generally, it is known that cultural learning is adaptive in circumstances where the relevant environments change too fast (spatially or temporally) to make purely genetic adaptation feasible, but not so fast that individual learning alone or no learning is required (Boyd & Richerson, [2005], [1985]; Henrich & McElreath, [2011], [2007]). Since people in different cultures think slightly differently (Henrich, [2020]; Nisbett et al., [2001]), this could thus make for precisely one of the situations in which it would be adaptive for humans to culturally learn how to mindread (especially since, as noted above, it is not plausible to posit genetic adaptations to quickly changing environmental conditions).

Second, the cultural-learning-based account also seems to sit well with the lack of full reliability in human mindreading. On the assumption that mentalizing is a culturally learned trait, many people might not end up being expert mindreaders. This matches what is true about book reading: not all humans can read equally well—some can plough through *Anna Karenina* in a few days, whereas for others, this is hard work taking years. Indeed, many humans cannot read at all. Similarly, since cultural transmission is consistent with the existence of various kinds of

biases in what is learned (Boyd, [2018]; Boyd & Richerson, [2005]), it is possible that humans learn to mindread in a slightly unreliable manner. For example, the cultural transmission of *pedagogical techniques* to teach children to learn to read can be stable, despite being in fact unreliable (i.e. not very good at teaching children to read). Mindreading could be like this, too.

However, this does not mean that there are no problems for this alternative account (see also Jacob & Scott-Phillips, [2021]; Morin, [2019]). In the main, these problems center on the fact that the cultural-learning based account seems to *overshoot* its target when it comes to feature 3 of human mindreading. As noted earlier, while there is some cultural variability in human mindreading, it is also true that human mindreading is found in all cultures (Peterson & Wellman, [2009]; Jacob & Scott-Phillips, [2021]). The contrast to book reading is actually quite telling here: relatively few human cultures are book-reading literate *at all*, but all cultures are mindreading literate. Compared to book-reading, mindreading appears highly culturally conserved. Indeed, book-reading takes many years of study (something it shares with learning how to make handaxes, bows, poison arrows, goulash, or maps: Sterelny, [2012]). While we may be able to observe some mindreading instruction in some cases—a point that is somewhat controversial (Jacob & Scott-Phillips, [2021])—we do not observe the kind of intensive learning and teaching found in other culturally learned abilities.[15]

All in all, what this means is that there is no major theoretical advantage for either account here. Both the biologically-adaptationist account and the cultural-learning-based one struggle to make sense of the full gamut of features surrounding human mindreading, but both accounts also seem to get important aspects of the evolution of human mindreading right. The obvious solution

[15] Of course, in principle, defenders of this view could argue that there are sufficient similarities in human environments that people in all cultures learn to mindread individually or with minimal instruction only. However, on the face of it, is not clear what these similarities would be, so that this response does not look greatly plausible on the face of it, at least as things stand.

is thus to combine these accounts somehow. The question is exactly how to do this. The next section proposes an answer.

4. The Bio-Social-Technological Evolution of Human Mindreading

At the heart of the present account of the evolution of human mindreading is a mutually reinforcing triad that comprises not just biological adaptation and cultural learning, but also socio-technological development. The first part of the account starts where the biological account above also started: with noting that either for social interaction-focused or for communication-and-cooperation-focused reasons (or both), mindreading abilities are highly adaptive. It is therefore plausible that we have evolved psychological expectations—about when to mindread, what sort of factors to pay attention to when doing so, and what some plausible options are for mental states to attribute—that make it easier for us to infer others' mental states from their observable behaviors and other features (Csibra & Gergely, [2011]; Luo & Baillargeon, [2007]).

However, the present account adds to this fact the further fact that, even given these kinds of expectations, mindreading of the kind shown by humans is quite costly, both in terms of time and the need for cognitive resources like attention and concentration.[16] Even knowing when to mindread, what sort of factors to pay attention to when doing so, and what some plausible options are for mental states to attribute, very many possibilities remain open. Human behavior is too variable to allow for mindreading expectations to be so fine-grained as to narrow the space of theoretical hypotheses to consider when attributing mental states to others to zero. To narrow

[16] Westra and Nagel ([2021]) also note that mindreading is cognitively costly, and therefore suggest that much of it is instead focused on "factive" states like what others *know* (see also Phillips et al., [2021]). However, as the below makes clear, their argument underplays the fact that cognitive technology and social roles can lessen these cognitive costs considerably.

down these possibilities further—beyond what is provided by our evolved mindreading expectations—the mindreader needs to consult their knowledge about the person in question and any general knowledge about how human minds work in this culture. Moreover, this might need to be done quite quickly if the mindreading inference is a crucial input into a decision.

To accomplish this, significant computational resources are needed. Applying the evolved mindreading expectations to the case at hand and supplementing them with specific information about the person and culture in question takes large stores of personal memory, as well as the ability to access the latter. It also requires highly flexible control over attention, concentration, and other executive functions to monitor the target and their interaction with their social, biological, and physical environment, potentially over significant periods of time.

Given that we have only finite amounts of such resources available to us and need to make decisions in real time (Lieder & Griffiths, [2020]; Gigerenzer, [2008]), this implies that mindreading is not a trivial task for us. In general, relying on abstract and complex mental representations when making decisions is typically computationally and cognitive-resource hungry (Lieder & Griffiths, [2020]; Schulz, [2018]). This is also why it is not compelling to see human mindreading as simply the result of more sophisticated (combinatorial) representational capacities in general: it may be the case that humans are able to represent more complex and abstract notions than non-human animals—however, the *use* of such representations in mindreading is costly, and cannot be taken for granted. Put differently: it needs not just be stated *that* humans have more sophisticated representational capacities (i.e. point (iii) of our evolved mindreading expectations); we must also explain *how* they can employ these capacities.[17]

[17] It may be thought that humans evolved not just the expectations and representations that *guide* mental state attributions—such representations about which situations call for mental state attributions and what features of the situation to pay attention to when making mental state attributions—but also the mental structures that help them *make* these inferences—such as dedicated neural processing space. However, the extent to which this is in fact a

With this in mind, the assumption that humans have evolved mindreading expectations to aid them in the navigation of their social environment—which, as noted earlier, is plausible as it narrows the inferential space—also implies that it is highly adaptively useful for humans to have access to the right kind of social or physical tools that help them lessen the costs of needing to rely on such expectations. Such tools can support the evolved psychological expectations with which humans make complex and adaptively important mental state attributions. There are two types of these kinds of tools that can be distinguished.

First, cognitive technology like books, myths, songs, and pictures can help sort out which of the large stock of evolved expectations concerning mental state attribution are especially important to focus on in the situation at hand. To deal with the complexity of human social living, humans are likely to need to rely on a large number of expectations about how to attribute mental states to others. Where someone is looking tells us something about what they were thinking; what someone did in the past may tell us something about what they are motivated to do now; who someone is related to may tell us something about what they feel like. We may have evolved to be able to make all sorts of mental state attributions; however, the ones that are most useful to do in the case at hand may be highly culturally specific, and their selection thus cognitively quite difficult. Stories, myths, and songs (among other things) can help make this selection. They model mental state attributions and tell us what features of the situation to pay particular attention to, so as to single out the mental states of particular importance here.

---

coherent view is not clear: mostly, it is thought that the way to help an organism *make* mental state attribution is by providing it with representations that *guide* the inference (Carruthers, [2006]; Fodor, [1983]; see also the overview of a related debate in linguistics by Chater & Christiansen, [2010]). At any rate, the proposal here is just that humans have evolved specialized expectations about how other minds work, but that these expectations still need to rely on the same cognitive and computational resources we use for other cognitive tasks (such as other kinds of representational decision making). Note also that, to the extent that the alternative, processing-focused proposal is found plausible, it continues to fall prey to the concerns for the biologically-adaptationist account noted in the previous section—without the ability to appeal to the kinds of considerations put forward in this section.

For example, in some cultures (such as a WEIRD one like the US), ascertaining someone else's character traits—e.g. whether they are conscientious—is useful for navigating social interactions: after all, in these (individualistic) cultures, people are encouraged to act on their character traits. In other cultures (such as a less WEIRD one like South Korea), ascertaining someone else's beliefs about their social ties—e.g. what status they have in relation to other members of their group—is more useful for navigating social interactions: after all, in these (less individualistic) cultures, people are encouraged to act in line with these beliefs about social ties, not their character traits per se (Henrich, [2020], pp. 32-33). In both cases, people may rely on the same set of evolved mindreading expectations; it is just that they "foreground" different ones.

Cognitive technology like stories, myths, and songs can help do exactly this. So, we may grow up listening to songs that say "I know that five years is a long time, and that times change; but I think that you'll find: people are basically the same" (Depeche Mode, *See You*), or we may grow up listening to stories like the Mahabharata, where you see that focusing on what Arjuna wants himself is not as helpful as focusing on his duties as a son to predict his actions. In both cases, the relevant cognitive technology aids us in determining the locally best uses of our otherwise similar mindreading capacities.[18] As Henrich ([2020], p. 33) puts it: "being consistent across relationships—"being yourself"—pays off more in America [than in Korea]." Songs, stories, and myths can help us figure this out quickly and easily.[19]

In this way, this kind of technology is crucial for helping humans be mindreaders. However, this is not because this technology is needed to teach humans to *become* mindreaders (as the cultural-learning-based account might have it: Heyes, [2018], p. 154). Rather, this technology

[18] It is also conceivable that this kind of technology adds some new, culturally specific mindreading expectations to the set of evolved ones (such as being "hangry" or feeling "ennui"). This is not so clear, though, so this will not be further considered here. It does not alter the main point in the text anyway.
[19] See also Apperly ([2011]), who spells this out in terms of mental scripts. See also Spaulding ([2018b], 2011]).

supports humans in *being* mindreaders. Cognitive technology can help humans make the locally best uses of the same set of evolved mental representations about other minds without breaking the bounds of their limited cognitive resources.

Second, complex and adaptively important mindreading is furthermore supported by the existence of social institutions that restrict the scope of possibilities for what an agent may be motivated by or thinking about in any given situation. In turn, this narrows down the option set of which mental state attributions to foreground in the case at hand, and thus makes mental state attribution much easier.

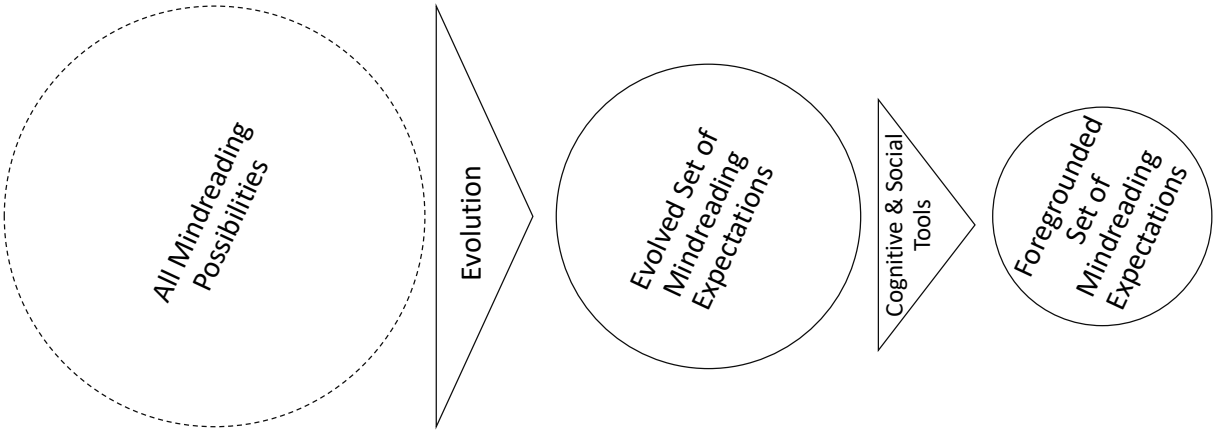
For example, if I know that you are the *lead hunter*, this gives me clues as to what you might be thinking about when there is discussion about hunting parties being put together. Where you are looking might provide me with information about whether you are hungry and your history of social interactions with person S might tell me something about your motivations for marriage. However, since I also know that you are the lead hunter, I know that the key motivational state to attribute to you here is to organize a hunting a party with the greatest chance of success.

As with cognitive technology like stories, songs, and myths, the existence of social roles allows humans to quickly tailor their mindreading to their local environment. It further narrows human mindreading expectations beyond their evolved set to the culturally most relevant ones, and thus brings the cognitive labor of mentalizing within the confines of what is possible given the limited cognitive resources open to humans.

However, also as with cognitive technology, it is important to understand this point correctly. It is sometimes thought that social role interpretation and prediction is an *alternative* to genuine mindreading (see e.g. Sterelny, [2003]). However, this is not the point here made. It may be that,

at times, the appeal to social roles makes mindreading redundant, as we can just go straight to the prediction of behavior.[20] However, at other times, the existence of social institutions can also *support* human mindreading—by making complex and locally adaptive mental state attributions feasible with limited cognitive resources.

Summarizing, then: if humans have evolved expectations that can guide mindreading inferences, then having access to the right kinds of cognitive technology and social institutions becomes adaptively valuable, too. This can further narrow the inferential space for attributing mental states to others, and thus brings it within the realm of feasibility for cognitively limited organisms like humans (see figure 1).



[Figure 1: Successive narrowing of the inferential space for mindreading]

How, though, do humans get access to these technological and social tools? It is here where the core idea of the cultural-learning-based account of human mindreading comes to its own.

[20] Social roles also have a normative dimension: they tell us what we *ought* to do and thereby provide a standard with which our behavior can be criticized. Indeed, this can extend to a meta-level: they might tell us that our behavior ought to stem from our mental states (see also McGeer, [2007]). I return to this point when discussing the work of Zawidzki below.

To see this, note that it is widely accepted that complex tools and social institutions develop by cultural learning. Building complex technology and institutions from scratch every generation is virtually impossible (Boyd & Richerson, [2005]; Boyd et al., [2011]; Henrich, [2015]). Even relatively simple tools like handaxes have been shown to have been refined over generations: figuring out, from scratch, how to hold, sharpen and shape a rock to make it useful as an axe is highly unlikely. This point holds even more strongly for highly complex cognitive technology like writing, dances, songs, myths, or stories. It also holds for social institutions: coming up with a stable social institutions like hierarchical stratification and religious authorities—especially ones that can be maintained in large societies—is very difficult (Henrich, [2020]). In this way, the present account takes on key insights from the cultural-learning-based account of human mindreading—but, unlike that account, it does not see mindreading as directly evolving by cultural learning. Rather, it views cultural learning as the process that builds up the tools and institutions that *support* mindreading.

This does not exhaust what needs to be said about the place of cultural learning in the evolution of human mindreading, though. It is not just the case that cultural learning leads to the existence of the kind of physical and social technology that can support mindreading. It is also the case that:

(a) Mindreading increases the scope for cultural learning.

(b) Mindreading increases the scope for tool use and social institutional navigation.

(c) Physical and social technology can support cultural learning.

These points are often separately recognized in the literature, but they are rarely put together and synthesized. This, though, is crucial to truly understand the evolution of human mindreading.

As far as point (a) is concerned, to learn complex behaviors from others, it is often necessary to be able to determine what the intention is behind the behavior. Otherwise, the student is unable to figure out which of the myriad of highly specific physical movements the teacher is going through are part of the behavior to be learned, and which are not (Sterelny, [2012]; Csibra & Gergely, [2011]; Tomasello, [1999]; Henrich, [2020], p. 129; Phillips et al., [2021]). In other words, determining what the teacher is intending is a key element in many learning processes (viz., those that go beyond cases where simple and direct behavioral copying is possible or useful).

Similarly, as far as point (b) is concerned, an understanding of what a designer intended a tool to be used for can help others use that tool more effectively. This holds for physical tools like screwdrivers (knowing it is intended to screw rather than hammer is helpful for using it effectively) as much as for cognitive and social tools. My understanding of what a story or song is about can be increased by an understanding of the intentions of the author.[21] Knowing that *The Threepenny Opera* is intended as a socialist critique of capitalist society can help decode its lessons. Equally, an understanding of the motives and thought processes of the occupants of various social roles can help me maneuver through my institutional landscape better: to avoid falling afoul of a social prescription, it can help to know what the motives are of town officials (say). This is especially important as these institutions get more complex, and it gets less clear what they prescribe. It is one thing to know about incest taboos in my culture; it is another to

[21] Of course, as is commonly noted in literary theory, the intentions of an author need not be crucial for an understanding of the (or a) meaning of the piece. The point here, though, is merely that there are circumstances where an understanding of the author's intentions improves the understanding of the meaning. This is uncontroversial—e.g. when it comes to instruction manuals.

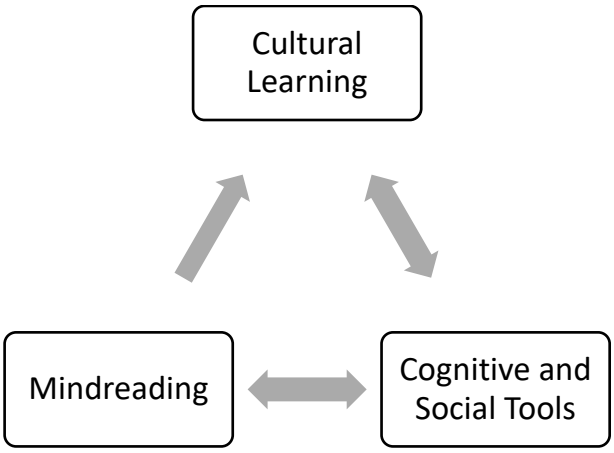
know what is required of merchants in foreign cultures (Henrich, [2020], pp. 318-319). Mindreading can help determine this.

Both (a) and (b) thus make explicit that there is a reverse side to the relationship between cultural learning, cognitive technology, and social institutions on the one hand, and mindreading on the other. It is not just the case that the former can guide the latter; it is also the case that mindreading can guide cultural learning and the interpretation and use of cognitive technology and social institutions.

Point (c), finally, is slightly less commonly noted, but also very important. Stories, myths, and songs and other cognitive artifacts enable teaching and learning on a scale that is drastically different from what is available otherwise (Henrich, [2020], pp. 436-437; Muthukrishna & Henrich, [2016]). By consulting these artifacts, I can learn from many more models—including ones from different times and places—by relying on information stored in long term, easily accessible media. I can also go over lessons frequently, in my time, and at my own pace. Social institutions can further prioritize my learning in important ways. This is obvious with formal schooling, but also applies to more informal toleration and encouragement of apprenticeship schemes (Sterelny, [2012]). In fact, it also holds for social institutions more generally: certain social institutions can structure society in such a way that it becomes possible to focus learning on different high-status individuals in different circumstances (Boyd & Richerson, [2005]). It is not just that we need to look to the village elder for all our instruction; we can look to the expert hunter for guidelines about how to hunt, the expert smith for guidelines about metallurgy, etc.[22]

[22] Note that the claim is not that *all* social institutions or all forms of cognitive technology can do this—the point is just that some can. I return to this below.

The upshot of all of these different causal relationships is a complex, interlocking set of feedback loops (see figure 2).[23]



[Figure 2: An interlocking set of feedback loops surrounding human mindreading]

The key thing to note about this figure is that it makes clear that cultural learning, technological and social tools, and mindreading support each other. Cultural learning allows for the development of complex cognitive tools and social institutions, which in turn support innate mindreading representational expectations, as well as more efficient and wide-ranging cultural learning. In turn, technologically and socially enhanced innate mindreading expectations support more efficient cultural learning, as well as better use of cognitive and social technology. In this way, cultural learning, technological and social tools, and mindreading can build each other up to higher and higher levels of complexity and sophistication.

[23] Henrich ([2020], pp. 299, 412) also hints at such a feedback loop, but does not spell it out in the way it is done here.

What makes this account more compelling than either of the other two accounts of mindreading is that it avoids the problems of the other two accounts above, while still being able to keep their advantages.

In the first place, it is due to the strength of this feedback loop that human mindreading can involve high levels of intentionality, highly abstract and complex mental state concepts, and be used in adaptively valuable situations. While doing so would generally be costly and time-consuming even for organisms that had the required representational resources, this use is supported by the existence of cognitive and social technology that helps human mindreading go smoothly.

Importantly, this also explains why non-human animals have not evolved mindreading abilities to rival those of humans, despite the fact that they sometimes faced some similar ecologies. Animals may have had rudiments of cultural learning, tool use, and mindreading expectations—but these did not have the strengths and connections to kickstart the feedback loop in figure 1. So, there is no groundbreaking, either / or difference between human and non-human animals here; rather, there are sufficiently many differences of degree that a larger difference emerges as a result of the feedback loop in figure 1 (or its absence).[24]

However, this feedback loop also explains the other two features of human mindreading: its cultural variability and lack of full reliability. Starting with the latter: it is precisely because mindreading is hard that it needs to be supported with the right kinds of cognitive tools and social institutions. Importantly, though, while these tools can *ease* our use of our evolved mindreading expectations, they cannot guarantee its success (Henrich, [2020]; Boyd & Richerson, [2005]). Importantly also, we do not always have access to these tools, and where

[24] This point is also supported by the differences in the mindreading capacities in domesticated and non-domesticated animals; see also Lurz ([2011]).

they are not available, mindreading efficiency will drop off. This explains why, apart from close social partners, our mindreading abilities are at their best when supported by well-functioning social and physical technology. It is easier to determine what occupants of social roles in our culture are thinking than otherwise similar random strangers (Henrich, [2020], pp. 318-319). However, the present theory can still make sense of the fact that mindreading is central to human social living: since technology and mindreading capacities culturally and biologically co-evolved, humans did often—though as just noted—not always have access to the right kinds of tools to enable their mindreading capacities.

When it comes to the cultural variability of mindreading, the key thing to note is that figure 1 shows how differences in the details of the cultural environment—the tools it makes available—can lead to differences in the nature of the resultant mindreading. However, at the heart of the account here is still the fact that all cultures mindread, and that they do so in ways that are often quite similar. All humans rely on what is a largely the same set of evolved mindreading expectations; it is just that the support system surrounding these mindreading expectations differs, leading to some differences in the way the mindreading is actually accomplished in different cultures. In this way, it is possible to explain how human mindreading practices can be finely tuned to how people in different cultures think (Henrich, [2020]; Nisbett et al., [2001]) without needing to appeal *just* to genetically driven mindreading adaptations or to *fully* culturally evolved mindreading abilities.

More than that: the present account brings to view novel areas of investigation. In particular, the present account points to the need for a more detailed picture of which socio-cognitive tools support which forms of mindreading. Myths emphasizing individual decision-making may favor a character-based rather than a situational form of mindreading, but what about different forms

of music and dance? Do more improvisational—and thus interpersonally-dependent—forms of music like jazz enhance mindreading abilities more than structured, individualistic forms of music like Western orchestral music? Do visual forms of story-telling (like comic books) lead to different forms mindreading than written forms (like traditional novels)? As yet, we do not know the answers to these questions, but, as the present account makes clear, they can greatly advance the cross-cultural understanding of human mindreading. The present account also aids in the development of tools to aid mutual human understanding. For example, it has become clear that social media does not promote social understanding: it appears to largely lead to the reinforcement of one's own perspective, rather than the appreciation of others. However, other forms of technology—such as neutral media sources—may do better here—though exactly which ones is not yet known.

Before concluding, it is finally useful to contrast this account with some alternative accounts that also seek to go beyond the biological-adaptationist and the cultural-learning-based perspectives sketched earlier. The goal in this is not an in-depth critique of these accounts, but rather to make clearer the unique aspects of the present account, so as to make explicit what it adds to the existing literature.

The first of these accounts is that of Tomasello (see e.g. Tomasello et al., [2005]; Tomasello, [1999], [2021], [2022]). According to Tomasello, at the core of human cognitive uniqueness—mindreading and cultural learning included—is the fact that humans have evolved the ability to jointly attend to something, and that they can therefore form collective intentions to accomplish a goal. That is, humans can not just track what others are seeing, they can represent the fact that they and others are attending towards the same thing with the mutual awareness that they all do

so. In turn, this allows humans to intend to do something together in the deeper sense of not just doing it simultaneously, but as a genuine collective agent (Tomasello, [2022]).

While Tomasello's account has some features in common with the present one—both of these accounts combine cultural learning (Tomasello, [1999]) and biological adaptations (see e.g. Tomasello, [2021], p. 242; chap. 11)—the account defended here goes beyond Tomasello's in several respects (see also Tomasello, [2021], p. 339). In particular, while it may be true that our evolved mindreading expectations contain the ability for joint attention and collective intention, as such, this leaves open *how* this ability can lead to distinctively human mindreading, given the difficulties of relying on mental representations. As also noted earlier, this is something that needs to be explained, and cannot just be taken for granted. Indeed, Tomasello's account does not lay out in detail the many different roles that technology plays in the bio-cultural evolution of human mindreading, as is done in using figures 1 and 2 here. The present account also diverges from Tomasello in not seeing non-human mindreading as mostly competitive in nature, and human mindreading especially cooperative. This is beneficial, since there is also much competition in human cultures (Whiten & Byrne, [1997]), and much cooperation—especially among kin—in non-human animals (see also Boesch, [2005]). Indeed, there is some work suggesting that complex forms of mindreading may be particularly helpful in competitive situations (Devaine et al., [2014]).

The other account that deserves mention here is that of Zawidzki (see especially Zawidzki, [2013]; Zawidzki, [2011]; Fenici & Zawidzki, [2021]; Zawidzki, [2018]). According to Zawidzki, at the heart of the human-non-human divergence in mindreading is the human ability to *shape* other minds. That is, we do not just take other minds as given, and try to determine what states they are in; rather, we actively *regulate* the thinking of others—through social tools like

myths and social norms—so as to narrow the space of possibilities of what others might be thinking (Godfrey-Smith, [2002]). In this way, we overcome the "holism" problem, according to which any particular form of behavior is consistent with indefinitely many mental states as causes. In turn, the existence of these social tools is made possible by the lower-level, implicit mindreading system, which is just sufficient to accomplish these more basic communicative and coordinative tasks (Renner & Zawidzki, [2018]).

Again, this account shares several key features with the present one—including, most importantly, the emphasis on social tools. However, there are again several important divergences here. First, there is little on physical technology like books, physical models, and images in Zawidzki's account, though as noted earlier, they play a key role here. Second, the present account does not limit social technology to "mindshaping." Rather, it shows that complex mindreading can be accomplished by using culturally acquired social and cognitive tools that *narrow* an evolved set of mindreading expectations, as in figure 1. Finally, the present account points out the existence of the feedback loop in figure 2: simple forms of cultural learning lead to basic forms of technology (Brown, [2018]); in turn, these support more complex mindreading—and thus more complex cultural learning.

In this way, the present account shows that while the picture sketched by Zawidzki is compelling as far as it goes—the kind of "mind-shaping" he sketches does indeed plausibly occur in some cases—it is not *all* that is going on here. Indeed, the present account expands Zawidzki's theory along two dimensions. First, some instances of mindreading are shown not involve significant mindshaping at all: they are based on a kind of "inferential narrowing" in mental state attributions instead. Second, even where mindshaping occurs, the latter is to be seen as embedded in the kind of ratchetting effect sketched in this paper: mindshaping influences and

is influenced by innate representational expectations, forms of cultural learning, and social and cognitive technology.

5. Conclusion

Human mindreading, though lacking in full reliability, is distinctive for its complexity and adaptive importance. It is found in all cultures, but also sees significant cultural variation in the way it is accomplished. This paper has argued the two major existing accounts—the biological one and the cultural-leaning based one—fail to explain all of these facts. To improve this situation, the paper has then presented a new account, according to which human mindreading is the result of selected for psychological mechanisms that are however supported by a ream of technological and social tools, including images, stories, and social institutions that guide mental state attribution. These tools and institutions are culturally developed. These three components also influence each other, thereby generating a positive feedback loop.

Acknowledgements

I would like to thank the referees of this journal for extremely helpful feedback on previous drafts of this paper.

Department of Philosophy

University of Kansas

Lawrence, KS, USA

awschulz@ku.edu

References

Acebo, C., & Thoman, E. B. [1995]. 'Role of infant crying in the early mother-infant dialogue'. *Physiol. Behav*, **57**, pp. 541-547.

Allen, C. [1999]. 'Animal concepts revisited: The use of self-monitoring as an empirical approach'. *Erkenntnis*, **51**(1), pp. 537-544.

Andrews, K. [2012]. *Do Apes Read Minds? Toward a New Folk Psychology*. Cambridge, MA: MIT Press.

Andrews, K. [2015]. *The Animal Mind. An Introduction to the Philosophy of Animal Cognition*. London: Routledge.

Andrews, K. [2018]. 'Do Chimpanzees Reason about Belief?'. In K. Andrews & J. Beck (Eds.), *The Routledge Handbook of Philosophy of Animal Minds* London: Routledge, pp. 258-268.

Apperly, I. A. [2011]. *Mindreaders: The Cognitive Basis of "Theory of Mind"*. London: Routledge.

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

Atran, S. [2002]. *In Gods We Trust: The Evolutionary Landscape of Religion*. Oxford: Oxford University Press.

Baillargeon, R., Scott, R. M., & He, Z. [2010]. 'False-belief understanding in infants'. *Trends Cogn Sci*, **14**(3), pp. 110-118. https://doi.org/10.1016/j.tics.2009.12.006

Baron-Cohen, S. [1997]. *Mindblindness: An essay on autism and theory of mind*. MIT press.

Bennett, J. [1991]. 'How to Read Minds in Behavior: A Suggestion from a Philosopher'. In A. Whiten (Ed.), *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mind Reading* Oxford: Blackwell Publishing, pp. 97-108.

Boesch, C. [2005]. 'Joint cooperative hunting among wild chimpanzees: Taking natural observations seriously'. *Behavioral and Brain Sciences*, **28**(5), pp. 692-693. https://doi.org/doi:10.1017/S0140525X05230121

Bowlby, J. [1958]. 'The nature of the child's tie to his mother'. *International Journal of Psychoanalysis*, **39**, pp. 350-373.

Boyd, R. [2018]. *A Different Kind of Animal: How Culture Transformed Our Species*. Princeton: Princeton University Press.

Boyd, R., & Richerson, P. [1985]. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. [2005]. *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.

Boyd, R., Richerson, P., & Henrich, J. [2011]. 'The cultural niche: Why social learning is essential for human adaptation'. *Proceedings of the National Academy of Sciences*, **108**(Supplement 2), pp. 10918-10925. https://doi.org/10.1073/pnas.1100290108

Boyer, P. [2001]. *Religion Explained: The Evolutionary Foundations of Religious Belief*. New York: Basic Books.

Brandom, R. [1994]. *Making it Explicit*. Cambridge, MA: Harvard University Press.

Brosnan, S. F., Grady, M. F., Lambeth, S. P., Schapiro, S. J., & Beran, M. J. [2008]. 'Chimpanzee autarky'. *PloS one*, **3**, pp. e1518.

Brown, R. L. [2018]. 'Animal traditions: what they are, and why they matter'. In K. Andrews & J. Beck (Eds.), *The Routledge Handbook of Philosophy of Animal Minds* New York: Routledge, pp. 362-371.

Butterfill, S. A., & Apperly, I. A. [2013]. 'How to Construct a Minimal Theory of Mind'. *Mind & Language*, **28**(5), pp. 606-637. https://doi.org/https://doi.org/10.1111/mila.12036

Byrne, R. W., & Corp, N. [2004]. 'Neocortex size predicts deception rate in primates'. *Proc Biol Sci*, **271**(1549), pp. 1693-1699. https://doi.org/10.1098/rspb.2004.2780

Carey, S. [2011]. *The Origin of Concepts*. Oxford: Oxford University Press.

Carey, S., & Spelke, E. [1996]. 'Science and Core Knowledge'. *Philosophy of Science*, **63**, pp. 515-533.

Carruthers, P. [2006]. *The Architecture of the Mind*. Oxford: Oxford University Press.

Chater, N., & Christiansen, M. H. [2010]. 'Language Acquisition Meets Language Evolution'. *Cognitive Science*, **34**(7), pp. 1131-1157. https://doi.org/https://doi.org/10.1111/j.1551-6709.2009.01049.x

Churchland, P. [2011]. *Braintrust: What Neuroscience Tells Us About Morality*. Princeton: Princeton University Press.

Cosmides, L., & Tooby, J. [1992]. 'Cognitive adaptations for social exchange'. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* Oxford: Oxford University Press, pp. 163-228.

Csibra, G., & Gergely, G. [2011]. 'Natural pedagogy as evolutionary adaptation'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**(1567), pp. 1149-1157. https://doi.org/doi:10.1098/rstb.2010.0319

de Waal, F. B. [2008]. 'Putting the altruism back into altruism: the evolution of empathy'. *Annu Rev Psychol*, **59**, pp. 279-300. https://doi.org/10.1146/annurev.psych.59.103006.093625

Devaine, M., Hollard, G., & Daunizeau, J. [2014]. 'Theory of Mind: Did Evolution Fool Us?'. *PloS one*, **9**(e87619), pp.

Edwards, K., & Low, J. [2017]. 'Reaction time profiles of adults' action prediction reveal two mindreading systems'. *Cognition*, **160**, pp. 1-16. https://doi.org/10.1016/j.cognition.2016.12.004

Ekman, P. [1989]. 'The Argument and Evidence About Universals in Facial Expressions of Emotion'. In H. Wagner & A. Manstead (Eds.), *Handbook of Social Psychophysiology* New York: Wiley, pp. 143-164.

Ekman, P., & Rosenberg, E. (Eds.). (1997). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford University Press.

Elster, A., & Gelfand, M. J. [2021]. 'When guiding principles do not guide: The moderating effects of cultural tightness on value-behavior links'. *Journal of Personality*, **89**(2), pp. 325-337. https://doi.org/https://doi.org/10.1111/jopy.12584

Fenici, M., & Zawidzki, T. W. [2021]. 'The origins of mindreading: how interpretive socio-cognitive practices get off the ground'. *Synthese*, **198**(9), pp. 8365-8387.

Fodor, J. [1975]. *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fodor, J. [1983]. *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Fodor, J., & Pylyshyn, Z. [1988]. 'Connectionism and Cognitive Architecture: a Critical Analysis'. *Cognition*, **28**, pp. 3-71.

Forster, M., & Sober, E. [2011]. 'AIC Scores as Evidence—A Bayesian Interpretation'. In M. Forster & P. S. Bandyopadhyay (Eds.), *The Philosophy of Statistics* Dordrecht: Kluwer, pp. 535-549.

Gallagher, S. [2015]. 'The new hybrids: Continuing debates on social perception'. *Consciousness and Cognition*, **36**, pp. 452-465. https://doi.org/https://doi.org/10.1016/j.concog.2015.04.002

Gauker, C. [2003]. *Words without Meaning*. Cambridge, MA: MIT Press.

Geurts, B. [2019]. 'Communication as commitment sharing: speech acts, implicatures, common ground'. *Theoretical Linguistics*, **45**(1-2), pp. 1-30.

Gigerenzer, G. [2008]. *Rationality for Mortals*. Oxford: Oxford University Press.

Godfrey-Smith, P. [2002]. 'On the Evolution of Representational and Interpretive Capacities'. *The Monist*, **85**, pp. 50-69.

Goldman, A. [2006]. *Simulating Minds*. Oxford: Oxford University Press.

Gopnik, A. [1996]. 'The child as scientist'. *Philosophy of Science*, **63**(4), pp. 485-514.

Grice, P. [1989]. *Studies in the Way of Words*. Cambridge MA: Harvard University Press.

Hare, B., Call, J., Agnetta, B., & Tomasello, M. [2000]. 'Chimpanzees know what conspecifics do and do not see'. *Animal Behaviour*, **59**, pp. 771-786.

Henrich, J. [2015]. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton, NJ: Princeton University Press.

Henrich, J. [2020]. *The WEIRDest People in the World*. New York: Farrar, Straus and Giroux.

Henrich, J., & McElreath, R. [2007]. 'Dual-Inheritance Theory: The Evolution of Human Cultural Capacities and Cultural Evolution'. In R. Dunbar & L. Barrett (Eds.), *The*

Oxford Handbook of Evolutionary Psychology Oxford: Oxford University Press, pp. 555-570.

Henrich, J., & McElreath, R. [2011]. 'The Evolution of Cultural Evolution'. *Evolutionary Anthropology*, **12**, pp. 123-135.

Heyes, C., & Frith, C. [2014]. 'The cultural evolution of mindreading'. *Science*, **344**(6190), pp. 1243091.

Heyes, C. M. [2012]. 'Grist and mills: on the cultural origins of cultural learning'. *Phil. Trans. R. Soc. B*, **367**, pp. 2181-2191.

Heyes, C. M. [2016]. 'Born Pupils? Natural Pedagogy and Cultural Pedagogy'. *Perspectives on Psychological Science*, **11**(2), pp. 280-295.

Heyes, C. M. [2018]. *Cognitive Gadgets: The Cultural Evolution of Thinking*. Cambridge, MA: Harvard University Press.

Hollis, M. [1994]. *The Philosophy of Social Science: An Introduction*. Cambridge: Cambridge University Press.

Humphrey, N. [1986]. *The Inner Eye: Social Intelligence in Evolution*. Oxford: Oxford University Press.

Jacob, P., & Scott-Phillips, T. [2021]. 'The Cultural Evolution of Human Social Cogntion: Is mindreading a gadget?'. *Synthese* **199**, pp. 1-27. https://doi.org/https://doi.org/10.1007/s11229-020-02620-4

Kanngiesser, P., Rossano, F., Frickel, R., Tomm, A., & Tomasello, M. [2020]. 'Children, but not great apes, respect ownership'. *Developmental Science*, **23**(1), pp. e12842. https://doi.org/https://doi.org/10.1111/desc.12842

Lavelle, J. S. [2021]. 'The impact of culture on mindreading'. *Synthese*, **198**, pp. 6351-6374.

LeDoux, J. E. [2012]. 'Evolution of human emotion: a view through fear'. *Progress in brain research*, **195**, pp. 431-442. https://doi.org/10.1016/B978-0-444-53860-4.00021-0

Leslie, A. M. [1994]. 'Pretending and believing: issues in the theory of ToMM'. *Cognition*, **50**, pp. 211-238.

Lieder, F., & Griffiths, T. L. [2020]. 'Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources'. *Behav Brain Sci*, **43**(e1), pp. 1-60. https://doi.org/doi:10.1017/S0140525X1900061X

Luhrmann, T. M., Padmavati, R., Tharoor, H., & Osei, A. [2015]. 'Hearing Voices in Different Cultures: A Social Kindling Hypothesis'. *Top Cogn Sci*, **7**(4), pp. 646-663. https://doi.org/10.1111/tops.12158

Luo, Y., & Baillargeon, R. [2007]. 'Do 12.5-month-old infants consider what objects others can see when interpreting their actions?'. *Cognition*, **105**(3), pp. 489-512. https://doi.org/10.1016/j.cognition.2006.10.007

Lurz, R. [2011]. *Mindreading Animals: The Debate over What Animals Know about Other Minds*. Cambridge, MA: MIT Press.

Lurz, R. [2018]. 'Animal Mindreading: The Problem and How it Can Be Solved'. In K. Andrews & J. Beck (Eds.), *The Routledge Handbook of Philosophy of Animal Minds* London: Routledge, pp. 229-237.

MacLean, P. D. [1985]. 'Brain evolution relating to family, play, and the separation call'. *Arch. Gen. Psychiatry*, **42**, pp. 405-417.

Mayer, A., & Träuble, B. E. [2013]. 'Synchrony in the onset of mental state understanding across cultures? A study among children in Samoa'. *International Journal of Behavioral Development*, **37**, pp. 21-28.

McGeer, V. [2007]. 'The Regulative Dimension of Folk Psychology'. In D. Hutto & M. Ratcliffe (Eds.), *Folk Psychology Re-Assessed* Dordrecht: Springer, pp. 137-156.

Millikan, R. [1984]. *Language, Thought, and Other Biological Categories*.

Moore, R. [2018]. 'Gricean communication, joint action, and the evolution of cooperation'. *Topoi*, **37**(2), pp. 329-341.

Moore, R. [2021]. 'The cultural evolution of mind-modelling'. *Synthese*, **199**(1), pp. 1751-1776. https://doi.org/10.1007/s11229-020-02853-3

Morin, O. [2019]. 'Did social cognition evolve by cultural group selection?'. *Mind & Language*, **34**(4), pp. 530-539. https://doi.org/https://doi.org/10.1111/mila.12252

Muthukrishna, M., & Henrich, J. [2016]. 'Innovation in the collective brain'. *Phil. Trans. R. Soc. B*, **371**(20150192), pp. https://doi.org/10.1098/rstb.2015.0192

Nichols, S., & Stich, S. [2003]. *Mindreading. An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds* (Vol. 114). Oxford University Press.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. [2001]. 'Culture and Systems of Thought'. *Psychol Rev*, **108**(2), pp. 291-310.

Nisbett, R. E., & Wilson, T. D. [1977]. 'Telling more than we can know: Verbal reports on mental processes'. *Psychol Rev*, **84**, pp. 231-259.

Parker, S. T., Mitchell, R. W., & Boccia, M. L. [1994]. *Self-awareness in Animals and Humans*. Cambridge: Cambridge University Press.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. [2008]. 'Darwin's mistake: explaining the discontinuity between human and nonhuman minds'. *Behav Brain Sci*, **31**(2), pp. 109-130; discussion 130-178. https://doi.org/10.1017/s0140525x08003543

Peterson, C., & Wellman, H. [2009]. 'From fancy to reason: Scaling deaf and hearing children's understanding of theory of mind and pretence'. *British Journal of Developmental Psychology*, **27**, pp. 297-310.

Phillips, J., Buckwalter, W., Cushman, F., Friedman, O., Martin, A., Turri, J., Santos, L., & Knobe, J. [2021]. 'Knowledge before belief'. *Behavioral and Brain Sciences*, **44**, pp. e140, Article e140. https://doi.org/10.1017/S0140525X20000618

Pinker, S., & Bloom, P. [1990]. 'Natural Language and Natural Selection'. *Behavioral and Brain Sciences*, **13**(4), pp. 707-784.

Povinelli, D. [2003]. *Folk physics for apes: The chimpanzee's theory of how the world works*. Oxford: Oxford University Press.

Renner, E., & Zawidzki, T. W. [2018]. 'Minimal Cognitive Preconditions on the Ratchet'. In L. D. Di Paolo, F. Di Vincenzo, & F. De Petrillo (Eds.), *Evolution of Primate Social Cognition* Cham: Springer International Publishing, pp. 249-265.

Rossi, M. [2014]. 'Simulation theory and interpersonal utility comparisons reconsidered'. *Synthese*, **191**(6), pp. 1185-1210. http://www.jstor.org/stable/24021575

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. [2010]. 'Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations'. *Proceedings of the National Academy of Sciences*, **107**(6), pp. 2408-2412. https://doi.org/10.1073/pnas.0908239106

Saxe, R., & Powell, L. J. [2006]. 'It's the Thought That Counts: Specific Brain Regions for One Component of Theory of Mind'. *Psychol Sci*, **17**(8), pp. 692-699.

Saxe, R., Schulz, L. E., & Jiang, Y. V. [2006]. 'Reading minds versus following rules: dissociating theory of mind and executive control in the brain'. *Soc Neurosci*, **1**(3-4), pp. 284-298. https://doi.org/10.1080/17470910601000446

Schaller, M., Park, J. H., & Kenrick, D. T. [2007]. 'Human evolution and social cognition'. In L. Barrett & R. Dunbar (Eds.), *Oxford Handbook of Evolutionary Psychology* Oxford: Oxford University Press, pp. 491-504.

Schulz, A. [2011]. 'Simulation, simplicity, and selection: an evolutionary perspective on high-level mindreading'. *Philosophical Studies*, **152**(2), pp. 271-285. https://doi.org/10.1007/s11098-009-9476-5

Schulz, A. [2017]. 'The Evolution of Empathy'. In H. Maibom (Ed.), *The Routledge Handbook of Philosophy of Empathy* London: Routledge, pp. 64-73.

Schulz, A. [2018]. *Efficient Cognition: The Evolution of Representational Decision Making*. Cambridge, MA: MIT Press.

Shahaeian, A., Peterson, C., Slaughter, V., & Wellman, H. [2011]. 'Culture and the sequence of steps in theory of mind development'. *Developmental Psychology*, **47**, pp. 1239-1247.

Skyrms, B. [2010]. *Signals: Evolution, Learning, and Information*. Oxford: Oxford University Press.

Sober, E. [2008]. *Evidence and Evolution*. Cambridge: Cambridge University Press.

Spaulding, S. [2011]. 'Embodied Social Cognition'. *Philosophical Topics*, **39**(1), pp. 141-162.

Spaulding, S. [2018a]. 'Do you see what i see? How social differences influence mindreading'. *Synthese*, **195**(9), pp. 4009-4030.

Spaulding, S. [2018b]. *How We Understand Others: Philosophy and Social Cognition*. London: Routledge.

Sperber, D. [1975]. *Rethinking symbolism*. Cambridge: Cambridge University Press.

Sperber, D. [1996]. *Explaining culture: A naturalistic approach*. Oxford: Blackwell Publishing.

Sperber, D., & Wilson, D. [1986]. *Relevance: Communication and Cognition*. Oxford: Blackwell.

Sterelny, K. [2003]. *Thought in a Hostile World: The Evolution of Human Cognition*. Oxford: Wiley-Blackwell.

Sterelny, K. [2012]. *The Evolved Apprentice: How Evolution Made Humans Unique*. Cambridge, MA: MIT Press.

Suddendorf, T., & Butler, D. L. [2013]. 'The nature of visual self-recognition'. *Trends Cogn. Sci.*, **17**, pp. 121-127.

Toda, K., & Platt, Michael L. [2015]. 'Animal Cognition: Monkeys Pass the Mirror Test'. *Current Biology*, **25**(2), pp. R64-R66. https://doi.org/https://doi.org/10.1016/j.cub.2014.12.005

Tomasello, M. [1999]. *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. [2021]. *Becoming Human: A Theory of Ontogeny*. Cambridge, MA: Harvard University Press.

Tomasello, M. [2022]. *The Evolution of Agency*. Cambridge, MA: MIT Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. [2005]. 'Understanding and sharing intentions: The origins of cultural cognition'. *Behavioral & Brain Sciences*, **28**, pp. 675-735.

Tomasello, M., & Herrmann, E. [2010]. 'Ape and Human Cognition: What's the Difference?'. *Current Directions in Psychological Science*, **19**(1), pp. 3-8.

Tooby, J., & Cosmides, L. [1992]. 'The Psychological Foundations of Culture'. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* Oxford: Oxford University Press, pp. 19-136.

Tooby, J., Cosmides, L., & Barrett, H. C. [2005]. 'Resolving the Debate on Innate Ideas: Learnability Constraints and the Evolved Interpenetration of Motivational and Conceptual Functions'. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The Innate Mind: Structure and Contents* Oxford: Oxford University Press, pp. 305-337.

Warneken, F., Lohse, K., Melis, A. P., & Tomasello, M. [2011]. 'Young Children Share the Spoils After Collaboration'. *Psychol Sci*, **22**(2), pp. 267-273. https://doi.org/10.1177/0956797610395392

Warneken, F., & Tomasello, M. [2006]. 'Altruistic helping in human infants and young chimpanzees'. *Science*, **311**(5765), pp. 1301-1303. https://doi.org/10.1126/science.1121448

Westra, E. [2020]. 'When is mindreading accurate? A commentary on Shannon Spaulding's How We Understand Others: Philosophy and Social Cognition'. *Philosophical Psychology*, **33**(6), pp. 868-882.

Westra, E., & Nagel, J. [2021]. 'Mindreading in conversation'. *Cognition*, **210**, pp. 104618.

Whiten, A., & Byrne, R. W. (Eds.). (1997). *Machiavellian Intelligence II: Extensions and Evaluations*. Cambridge University Press.

Zawidzki, T. W. [2011]. 'How to interpret infant socio-cognitive competence'. *Review of Philosophy and Psychology*, **2**(3), pp. 483-497.

Zawidzki, T. W. [2013]. *Mindshaping: A New Framework for Understanding Human Social Cognition*. Cambridge, MA: MIT Press.

Zawidzki, T. W. [2018]. ‘Mindshaping’. In S. Gallagher, A. Newen, & L. D. Bruin (Eds.), *The Oxford Handbook of 4E Cognition* Oxford: Oxford University Press, pp. 735-754.

Zeberg, H., & Pääbo, S. [2021]. ‘A genomic region associated with protection against severe COVID-19 is inherited from Neandertals’. *Proceedings of the National Academy of Sciences*, **118**(9), pp. e2026309118. https://doi.org/doi:10.1073/pnas.2026309118