Synthese (2022) 200:150
https://doi.org/10.1007/s11229-022-03662-6

ORIGINAL RESEARCH

Minimal theory of mind – a Millikanian Approach

Nimra Asif[1]

Received: 23 July 2021 / Accepted: 17 March 2022

Abstract

Minimal theory of mind (ToM) is presented in the theory of mind literature as a middle ground between full-blown ToM and mere behavior-reading. Minimal ToM seems to be a useful construct for studying and understanding the minds of nonhuman animals and infants. However, providing an account of minimal ToM on which minimal mindreading is significantly less demanding than full-blown mindreading yet more than just a behavior-reading process is a challenge. In this paper, I argue that to address this challenge, we need to depart from the traditional framework of mindreading in more radical ways than offered by current minimal theory of mind accounts. First, I explain the traditional view of mindreading on which mental state attribution is treated as essential for mindreading and analyze the general respects in which it makes mindreading demanding for the mindreader, such as requiring the mindreader to have concepts of mental states, engage in inferential reasoning processes involving mental states, and form meta-representations. Then I discuss and critically evaluate two accounts of minimal ToM and argue that these accounts either do not depart sufficiently from the demanding requirements of traditional mindreading or risk becoming re-descriptions of behavior-reading accounts. Finally, I present an alternative Millikanian account of minimal ToM that avoids this risk while departing more radically from the traditional view of mindreading by providing a way for minimal mindreaders to represent the mental states of others and respond to them without engaging in conceptual mental state attribution.

Keywords Full-blown theory of mind · Minimal theory of mind · Behavior-reading · Millikanian Minimal Mindreading

✉ Nimra Asif
nimra.asif@uconn.edu

[1] Department of Philosophy, University of Connecticut, 06269-1054 Storrs, CT, United

Published online: 15 April 2022

1 Introduction

'Mindreading' or 'Theory of Mind' refers to the ability to attribute mental states, such as intentions, beliefs, and desires to oneself and others (Premack & Woodruff, 1978). As Premack and Woodruff explain, "A system of inferences of this kind is properly viewed as a theory because such states are not directly observable, and the system can be used to make predictions about the behavior of others" (515). According to the theory-theory, a prominent account of how we understand other minds, we come to know other minds using a folk-psychological theory (Churchland, 1991; Carruthers, 1996; Gopnik & Wellman, 1992). The theory-theory proposes that our understanding of other minds is the result of postulating mental states understood as "abstract unobservable entities" to best explain observable behavior and predict future behavior (Gopnik & Wellman, 1992, 148). So, our understanding of others' minds involves "theoretical constructs," such as perceptions, desires, and beliefs, that "go beyond the focal evidential phenomena" (Gopnik & Wellman, 1992, 153). This posited theory of how we come to understand other minds also requires attributing to mindreaders an understanding of coherent law-like relations between theoretical constructs of mental states and sensory input from the external world, overt behavior, and other mental states (Churchland, 1991; Gopnik & Wellman, 1992; Carruthers, 1996), such as "If S desires a certain outcome G and S believes that by performing a certain action A she will obtain G, then *ceteris paribus* S will decide to perform A" (Barlassina & Gordon, 2017). The general idea is that knowing or understanding other minds is a matter of theoretical reasoning that yields mental state attributions, which enable one to predict and explain the behavior of others (Avramides, 2020).

Even when a theory is not seen as required for mindreading, mental state attribution remains essential on traditional views of mindreading. Simulation theorists, for example, argue that we understand other minds by simulating them using our own minds as a model of others' minds, instead of using a folk psychological theory (Goldman, 1989; Gordon, 1986; Heal, 1998). These researchers argue that we "exploit the fact that we are or have minds" to understand what is going on in the minds of other people (Heal, 1998, 84). We put ourselves in the shoes of another person based on the input information we have about the person and their situation, and then, "offline," we run our mental processes (what we would feel, think, do) as if we were that person. The end result of the process is used to form predictions and explanations of the other person's behavior (Avramides, 2020). So, A pretends to be in B's situation and then A attributes to B the mental state that A is in during this process of imagination.

These traditional views highlight some of the different reasons that the act of mindreading is taken to require mental state attribution and inference and is viewed as cognitively demanding. First, mental states are seen as hidden or unobservable phenomena, so their existence must be inferred from subjects' observable behaviors or from other mental processes and then attributed to them. Second, theory of mind (ToM) is viewed by some as a genuine theory that requires the possession of theoretical mental state concepts and the ability to make inferences using these concepts and an understanding of the relationship between mental states, agents' behavior, and their environmental conditions. The end products of such theorizing (posited mental states) then need to be attributed to others to predict or explain their behavior.

Thirdly, even if ToM is not viewed as a genuine theory, it is thought to require representing the mental states of others *as such* – i.e. as beliefs, desires, and perceptions. (The attributer has to think of e.g. Amy as *believing that* the toy is in the yellow box, or Max as *knowing that* Jack is hiding his toy car.) This means that individuals who make mental state attributions need to have some significant conceptual understanding of the nature and essential properties of the attributed mental states as a precondition of attributing them to others.

The requirement of mental state attribution for ToM makes traditional theory of mind cognitively demanding in many ways. Mental state attributions require the mindreader to (1) conceptually represent mental states[1], such as beliefs, desires, intentions, or perceptions, (2) represent the mental states with their intentional content or "that which they are 'about' or directed toward" and (3) understand "by some means, the relations between an agent's mental states, their environmental conditions, and their behavior" which enables the mindreader "to make predictions about others' behaviors and to explain those behaviors" (Hutto et al., 2011).

Since, on the traditional view of mindreading, mindreading consists in attributing mental states to others, the conceptual requirements that the mindreader needs to meet to engage in mindreading are quite demanding. To attribute propositional attitudes to others, the mindreader must herself have the ability to form metarepresentations: beliefs *about* others' mental states. This, in turn, requires the mindreader to have concepts of mental states and an understanding of their relation to the agent's environment and behavior, as well as engage in inferences that result in mental state attributions.

Recently, there have been attempts to characterize minimal forms of mindreading that are less demanding than the traditional account of theory of mind that are motivated by empirical findings concerning the theory of mind abilities of infants and non-human animals (Onishi & Baillargeon, 2005; Butterfill & Apperly, 2013; Low & Perner, 2012; Gomez, 1996, 2017; Whiten, 2013). The goal of minimal theory of mind accounts is to identify a way that those with limited cognitive resources could engage in mindreading. However, I argue that current accounts of minimal ToM fail to identify genuinely minimal forms of mindreading, because they do not sufficiently depart from the traditional frameworks for understanding mindreading. These accounts focus on taking *what* infants and non-human animals represent to be more minimal; but they fail to identify ways in which *how* minimal mindreaders understand the minds of others is more minimal. This is because the process of mindreading on these views continues to be as cognitively demanding as on traditional accounts of mindreading, insofar as it requires mental state attribution (understood as above).

In this paper, I begin to develop an account of minimal ToM that departs from traditional accounts of mindreading more radically. On the account I offer, individuals can represent mental states and respond to them without engaging in conceptual attribution of mental states to others or having any theoretical knowledge of mental states and of other minds. Section II presents a sketch of the relevant empirical literature

[1] In order to conceptually represent mental states, the attributor must have the full range of mental state concepts.

that motivates the search for minimal forms of mindreading, and briefly discusses the importance of providing an account of minimal theory of mind. In Section III, I critically evaluate contemporary accounts of minimal theory of mind and argue that these accounts fail to identify sufficiently minimal forms of minimal mindreading. Section IV provides an alternative account of minimal mindreading using Ruth Millikan's framework of mental representation. On this account, infants and non-human animals can have practical knowledge of mental states, which consists in forming representations of affordances that allow them to deal with the mental states of others without having any conceptual understanding of mental states or engaging in mental state attribution. Section V briefly discusses a few empirical and practical advantages of this account and raises further research questions that need to be addressed.

2 Empirical debate on theory of mind

Traditionally, the test for ToM abilities has been the false-belief task that tests children's ability to attribute false beliefs to others. Passing the test is supposed to demonstrate that children understand that individuals can believe things about the world that are different from reality (Wimmer & Perner, 1983). In the standard false-belief task, children are shown two dolls, Sally & Anne, playing with a toy (Wimmer & Perner, 1983; Wellman et al., 2001). Then Sally puts the toy in her basket and leaves the room. After Sally leaves the room, Anne takes the toy from the basket and hides it in her box. Then Sally returns to the room and the child is asked where Sally will look for her toy. Empirical studies with false-belief tasks for children have found that 4-year-olds respond that Sally will look in the basket, whereas 3- and 2-year-old children point to the box where the toy is actually located (Wellman et al., 2001). These studies have been interpreted as evidence that ToM abilities develop around 4 years of age and children younger than 4 do not have these abilities and cannot understand mental states (Wimmer & Perner, 1983; Wellman et al., 2001). Some researchers have explained the late development of ToM by arguing that ToM development is dependent on language (Rakoczy, 2017; O'madagain & Tomasello, 2019; Moore 2020).

Other researchers argue that children younger than 4 can understand mental states, but they fail standard ToM tests because these tests are cognitively demanding and require language comprehension and executive functioning (Bloom & German, 2000; Leslie, 1987). In the past 17 years, researchers have started using non-standard and implicit measures to test the ToM abilities of younger children. Whereas the standard tests involve cognitively demanding verbal measures that require children to report how agents with certain beliefs will act, the implicit non-standard measures use non-verbal looking time tests that rely on children's looking behavior to see if they anticipate how agents with certain beliefs will act (Onishi & Baillargeon, 2005; Southgate et al., 2007; Surian et al., 2007; Kovacs et al., 2010).

The studies that use non-standard measures to test for ToM understanding have found positive results. For example, Onishi & Baillargeon (2005) found that 15-month-old infants expected an actor to search for an object based on the actor's belief about its location rather than its actual location. Surian et al., (2007) showed

that 13-month-old infants expected an animated caterpillar in a video to change its behavior based on the information (true belief or false belief) it had about an object's location. Non-verbal and non-standard measures have also been applied to testing the ToM abilities of non-human animals, such as apes, monkeys, and corvids; these tests, too, have found positive results suggesting that these non-human animals can understand others' mental states, such as perceptions, goals, and false beliefs (Hare et al., 2000, 2001; Emery & Clayton, 2001, 2004; Krupenye et al., 2016). For example, in the Hare and colleagues' study (2000), a dominant chimpanzee and a subordinate chimpanzee competed for pieces of food placed in the central arena between their rooms. In the study, one piece of food was placed in the center of the arena and visible to both the dominant and the subordinate chimpanzee, but the second piece of food was visible to only the subordinate because it was placed on the subordinate's side of an opaque barrier. When the doors of their rooms were opened and the chimpanzees were released into the central arena, the researchers found that the subordinate chimpanzees had a clear preference for taking the food that the dominant could not see and leaving the food the dominant could see alone. Another study by Krupenye and colleagues (2016) found that great apes understand false beliefs and pass false belief tests when they are tested using anticipatory looking time measures.

The interpretation of these positive results has been a topic of intense debate in both the infant and the non-human animal literature (Van der Vaart & Hemelrijki, 2014; Beaudoin et al., 2020; Moore, 2020). Some researchers argue that these findings suggest that infants do understand others' mental states and can thus be credited with a ToM (Powell et al., 2018; Carruthers, 2013; Baillargeon et al., 2010). The same argument is made by certain researchers about non-human animals (Hare et al., 2000, 2001; Emery & Clayton, 2008; Tomasello & Call, 2006). Others find these positive results unconvincing and think that the behavior of infants and nonhuman animals can be explained by taking the subjects to be using behavior-reading, which involves prediction or anticipation based on simple behavioral rules, such as "People look for an object in the last place they encountered it" (Ruffman & Perner, 2005; Heyes, 2014; Penn & Povinelli, 2007; Lurz, 2011; Penn et al., 2008).

The behavior-reading and the mindreading positions require attributing to the nonhuman animals and infants very different abilities. If we adopt the mindreading position, we have to credit nonhuman animals and infants with complex social-cognitive abilities that are involved in predicting the behavior of others by positing underlying mental states. These abilities include understanding that others have mental states, inferring the presence of such states and representing them, and using attributed mental states to predict or make sense of their behavior (Hutto et al., 2011; Tomasello, 2018). At the very least, we must suppose that the mindreaders have the ability to attribute such states to others. On the other hand, the behavior-reading position only requires us to credit animals and infants with a simple predictive strategy, according to which they use statistical correlations or associative rules connecting observable cues (eye gaze, vocalizations, body posture) and behaviors to predict another agent's future actions (Heyes, 2018; Penn & Povinelli, 2007). This predictive strategy does not rely on complex social-cognitive and inferential abilities; it does not even require the attribution of mental states to others.

A middle ground between these two positions (mindreading vs. behavior-reading) has been advocated by some researchers. They put forward the view that, although infants and nonhuman animals do not engage in the traditional form of mindreading, their behavior is more sophisticated than mere behavior-reading and involves "*implicit"* theory of mind, or 'minimal theory of mind' (Onishi & Baillargeon, 2005; Butterfill & Apperly, 2013; Low & Perner, 2012; Gomez, 1996, 2017; Whiten, 2013). 'Explicit ToM' or 'full-blown ToM' are both used to refer to the traditional or standard view of mindreading, whereas 'implicit ToM' and 'minimal ToM' are terms that are given for ways of establishing a middle-ground between traditional mind-reading and behavior-reading. Since there does not seem to be a consistent and clear use of the term 'implicit ToM,'[2] I will use the term 'minimal ToM' which was coined by Butterfill & Apperly (2013) to describe a less demanding way of tracking mental states than traditional mindreading that could be used by those with limited cognitive resources. This notion of minimal theory of mind presents a promising way of understanding the behavior of infants and non-human animals, since it does not appear to require crediting them with the linguistic or complex social-cognitive abilities that are required for full-blown ToM of the sort adult humans have. At the same time, it can be recognized that infants' and non-human animals' understanding of others goes beyond the mere use of predictive behavioral rules. The notion of minimal ToM or minimal mindreading can also be used to provide an account of the origins and the evolution of mindreading on which there is a psychological continuity between the mindreading abilities of nonhuman animals and infants, on the one hand, and adult humans, on the other. But providing an account of minimal theory of mind on which minimal mindreading presents a genuine middle ground between mindreading and behavior-reading remains a challenge. In the next section, I will discuss current accounts of minimal theory of mind and argue that these accounts either do not depart sufficiently from the traditional view of mindreading or risk falling into the behavior-reading 'camp.'

In what follows, I will use the term Reader to refer to the individual that is confronted with the task of interpreting and handling the action or behavior of another individual, the Agent. ("Reader" is intended to be neutral between using behavior-reading, mindreading, or something else to interpret the behavior of other agents.) On a minimal ToM account, the Reader would have to interpret and handle the behavior of another agent using a strategy that is significantly less demanding than traditional mindreading but more demanding than the use of correlations between observable cues and behaviors posited by the behavior-reading account.

3 Minimalist Accounts of Mindreading

Stephen Butterfill and Ian Apperly (B&A) (2013) characterize a form of mindreading that does not require the Reader to represent propositional attitudes, such as perceptions, intentions, beliefs, or knowledge *as such*. Instead, it would only require the Reader to represent the "simpler, relational mental states" of individuals that can be

[2] See Low & Perner (2012) for a detailed discussion of this point.

used to track propositional attitudes, such as beliefs and perceptions (B&A, 2013, 607). On B & A's account, there are a few main principles that are needed to engage in minimal mindreading. First, *representing goals* does not require representing intentions as such; instead, it requires linking bodily movements to outcomes, since "bodily movements form units which are directed to goals" (B&A, 615). But an individual cannot perform goal-directed behavior towards an object without encountering it. *Encountering* is defined as a relation between an agent and an object in the agent's field. And determining whether or not an object is within an agent's field is dependent on spatial proximity as well as the agent's orientation and posture (B&A, 2013). Since representing encountering consists in representing relations between an agent and an object, it contrasts with representing the mental state of perception. The last principle required on B&A's account includes the notion of *registration*, which is defined as a relation between an individual, an object, and a location. So, an individual registers an object at a location if the individual most recently encountered the object in that location. A Reader that represents registrations can represent an agent's correct or incorrect registration of an object based on whether the object is still in that location where the agent last encountered it. For B&A, representing registration is similar to belief representation, but it also contrasts with belief representation, since it does not require representing belief (true or false) as such (B&A, 2013).

An example that can be used to illustrate B&A's account is the Hare et al., (2001) study. In the study, a subordinate and a dominant chimpanzee competed for food. The food was placed on the subordinate's side of the opaque barrier, so the subordinate always saw the baiting process and whether the dominant was watching or not. There were three experimental conditions: (1) the control - the dominant watched the food being hidden (2) the uninformed - the dominant did not watch the food being hidden and (3) the misinformed - the dominant watched the food being hidden, but then it was moved while the dominant was not watching. The researchers found that the subordinate chimpanzees chose to approach and retrieve the food significantly more often in the uninformed and misinformed conditions than in the control condition.

According to B&A's minimal theory of mind account, the subordinate chimpanzee represents the dominant chimpanzee as being in a state of correct registration in the control condition in which he encountered the piece of food at location B, and it is in location B. In the uninformed condition, the dominant has not encountered the food, so he cannot register it. While in the misinformed condition, the subordinate represents the dominant as incorrectly registering the location of the food as A when it is in location B. The subordinate understands that "correct registration is a condition" for "successful goal-directed action," so the dominant cannot get the food in the uninformed and misinformed conditions (B&A, 2013, 618). This minimal form of mindreading can explain why the subordinate approaches and retrieves the food only in those conditions. B&A think that their account presents a middle ground between full-blown mindreading and behavior-reading, since it describes a way one can track mental states without representing propositional attitudes, such as perceptions, intentions, and beliefs.

One of the major difficulties with B&A's account has to do with their notion of registration and how it is different from belief. Is registration a mental state that intervenes between perception and action? Is it both represented by the minimal

mindreader and attributed to other agents? If we take registration not to be a mental state, then it becomes difficult to distinguish B&A's minimal ToM account from behavior-reading accounts that only credit the Reader with behavior-reading abilities. In her commentary on B&A's paper, Shannon Spaulding raises this criticism and argues that B&A's account would not be substantially different from behavior-reading accounts if registration is understood as merely a relational state (Butterfill et al., 2013). For example, behavior-reading accounts claim that Readers deal with other minds or Agents by invoking behavioral rules, such as 'people will look for objects where they last left them or encountered them.' These behavioral rules do not require the Reader to have a capacity to attribute mental states, possess concepts of mental states, or use meta-representations. They merely involve representing a relation that obtains between the Agent, an object or event, and location. Based on this interpretation, B&A's minimal theory of mind account would not seem to be substantially different from behavior-reading accounts because it would also require mindreaders only to represent such ('external') relations (Butterfill et al., 2013). And proponents of behavior-reading accounts would have no problem accepting B&A's relational principles because they are consistent with their framework. Spaulding challenges B&A to explain how their account of minimal theory of mind differs from behavior-reading.

In "Replies to Three Commentaries on Minimal Theory of Mind," B&A take on board Spaulding's challenge and attempt to explain how their account can be distinguished from a behavior-reading account. B&A begin by discussing the question of what is a mental state and whether registration is a mental state. They explain that "mental states involve subjects having attitudes toward contents" and there are debates about "how the attitudes should be characterized (e.g., whether belief is characterized by appeal to norms or not) and how the contents of beliefs should be distinguished (e.g., by appeal to one or another kind of proposition)" (Butterfill et al., 2013, 8). According to B&A, registration could be a mental state on some philosophers' views, albeit a simpler one, given that its functional role is less complex than belief, and specifying its content only requires appealing to relations between objects and locations. On the other hand, some philosophers would deny that registration is a mental state, since they require normativity for mental states. B&A "duck these controversies" by arguing that empirical tests for ToM understanding do not test for normative aspects of mental states, so "on the notion of a mental state implicit in much theory of mind research, registrations are mental states" (Butterfill et al., 2013, 9). And since registration is a mental state, on B&A's view, minimal mindreading requires the mindreader to attribute it to other agents. They write: "So how does minimal theory of mind differ from the use of behavioral rules? It involves ascribing registration, which is a mental state or something very like one" (Butterfill et al., 2013, 9). Thus, whereas in their original paper B&A only refer to mindreaders' "tracking" of mental states, in their reply to Spaulding's challenge, they appear to acknowledge that their view requires mindreaders to attribute mental states to others.

This suggests that B&A's account would differ from behavior-reading accounts precisely because it would require the Reader to attribute to the Agent mental states (albeit simpler ones, such as registration). But if this is so, then the account may still portray minimal mindreading as quite demanding. This becomes evident when we

consider a crucial distinction between *what* the Reader is attributing to the Agent, which is the mental state *attributed* and its content, on the one hand, and the process of attribution, which is *how* the mental state is being attributed or represented by the Reader.[3] On B&A's account, the Reader is attributing simpler mental states to others, instead of full-blown propositional attitudes. But the process of attribution that the Reader would have to engage in is still a meta-representational, inferential process as on the traditional mindreading account. The Reader would still herself need to have metarepresentational *beliefs about* the mental representations of others (registrations), such as whether the dominant has correctly registered the location of the food or incorrectly registered it. Forming beliefs about the mental states of others requires conceptual meta-representational capacities: it requires deploying the concept of the relevant mental state and its content (in this case – registration and its content). (And it would also appear to involve an inference about the other's mind; see below.) So, even if what the Reader is representing are simpler mental states, the process of understanding the mental states of others remains as cognitively demanding as it is on the full-blown mindreading account.

Moreover, it is not clear what the difference between tracking and 'representing as such' is on B&A's account. In his commentary on B&A's account, Tad Zawidzki argues that on some theories of mental representation, this distinction between representing and tracking cannot be drawn. On these theories, the following inference is valid:

1. S can track beliefs by virtue of representing registrations.

Therefore:

2. S's representations of registrations are representations of beliefs. (Butterfill et al., 2013, 10).

Absent a positive theory of mental representation, it isn't clear how this inference would be blocked on B&A's account. Consequently, there would seem to be no way for them to distinguish tracking from representing, or explain what more is needed to represent beliefs 'as such,' beyond representing registrations.

Furthermore, it seems as though B&A do not entirely reject the view that mental states are hidden and need to be inferred from behavior; for registrations, they say, "are intermediate variables and play a subset of the causal roles characteristic of belief" (2013, 618). As they explain, "registration is only defeasibly linked to encountering" because "someone using minimal theory of mind may infer that an individual registers an object at a location even though that individual has never encountered the object at that location" (Butterfill et al., 2013, 11). And registration is not "reducible to action" because it is possible to register an object at a location after encountering it there and not perform any goal-directed actions on it (11). So, registration cannot be "identified with either encountering or acting" and "intervenes

[3] This distinction between how and what in mental state attribution and representation is due to Dorit Bar-On.

between encounters and actions" (Butterfill et al., 2013, 12). Since registration is an intermediate variable between encountering and goal-directed action, and the Reader is required to attribute it *as* an intermediate variable on this account, then the Reader must infer that the Agent is in a state of correct registration based on the observation that the Agent has encountered the object, and can now perform goal-directed action towards the object because the Agent is in the state of correct registration. And to make this inference, the Reader must have concepts of these mental states as well as an understanding of the principles that connect them to behavior, such as the general principle that the state of correct registration is a necessary condition for successful goal-directed action. Thus, B&A's minimal mindreading account would still be demanding along the dimensions I mentioned earlier, insofar as it places conceptual and cognitive requirements on mindreaders that are not significantly less demanding than those placed by traditional accounts.

Another minimal theory of mind account by philosopher Laura Danón (2016) takes a first step in the right direction; however, it too falls short of providing a genuinely minimal ToM account by not departing sufficiently from the traditional view of mindreading. Danón (2016) relies on Ruth Millikan's theory of primitive mental representations to explain how animals can engage in minimal forms of mindreading. I briefly explain Millikan's general theory of mental representation before going on to discuss the role that her theory plays in Danón's account of minimal theory of mind.

Millikan (1989; 1995; 2004) uses evolutionary theory to provide a naturalistic account of intentionality. She explains the simpler representational abilities of biological organisms, such as bacteria and bees, and the more complex representational abilities of humans using a unified naturalistic model. Millikan (1995) offers an account of primitive mental representations, and then shows how organisms evolve more complex mental representations by "modifying less differentiated multipurpose structures into more differentiated dedicated ones" (192). Millikan's framework is interesting for the purposes of offering an account of minimal ToM because of the promise it holds for explaining how, beginning with primitive mental representations, nonhuman animals and infants might acquire the capacity to represent the mental states of others, before possessing language or highly complex cognitive capacities.

On Millikan's account, there are three main types of representations - descriptive, directive, and pushmi-pullyu representations (Millikan, 1995). Descriptive representations depict what is going on in the world – representing facts – while directive representations guide action by showing what to do – representing goals. For example, a shopping list can be directive or descriptive.[4] When the list is a directive, it gives directions about what to purchase, and when it is descriptive, it indicates what was purchased. For Millikan, using descriptive and directive representations requires the ability to make mediated inferences, since directive representations that show what to do have to be combined with appropriate representations of facts or what is the case in the world to be applied. Similarly, representations that show what is the case in the world have no action-guiding utility unless combined with directive representations of goals (Millikan, 1995).

[4] The example is due to Anscombe (1957).

Pushmi-pullyu representations, however, are simpler and more primitive since they do not require this process of mediating inference. Pushmi-pullyu representations (PPRs) have both the descriptive and directive function in a single representation that provides information about what is the case in the world and what the organism should do about it (Millikan, 1995). For example, a bee dance can be described as a PPR, since it gives descriptive information about the location of the nectar and directs the receivers to go retrieve the nectar. In contrast to purely descriptive and directive representations, a process of practical inference is not required since the descriptive information in the PPR is immediately linked to action. According to Millikan's theory, pushmi-pullyu representations are the most primitive; as a representational system advances, the descriptive aspects begin to come apart from the directive aspects, and then turn into detached descriptive and directive representations (2004).

On Danón's minimal ToM account, full-blown mindreaders attribute detached mental states to others that either have purely descriptive or directive content but not both in order to explain and predict their behavior. Since the mental states being attributed are detached, the mindreader must understand how they are being linked to guide the other agent's behavior, so the mindreader needs to reason about the inferential processes of others (Danón, 2016). The minimal mindreader does not have to engage in this reasoning because the minimal mindreader attributes pushmi-pullyu representations to others to predict and explain their behavior. Since the mental state that the mindreader attributes already has both the descriptive and directive content combined together, the mindreader does not have to figure out the inferential links between them. For this reason, attributing PPRs is simpler or more minimal, because it requires the minimal mindreader to attribute a single mental state rather than attributing multiple mental states and reasoning about their inferential links to predict and explain the behavior of others.

On Danón's account, there are minimal forms of mindreading that are more or less sophisticated depending on the type of PPs[5]that the mindreader attributes to others. According to Danón, the most basic mindreaders attribute pure pushmi-pullyu representations (PPs) to others. The basic mindreader that attributes pure PPs is representing one mental state that has both perceptual ('food there') and prescriptive content ('grab and eat it'). And this mindreader has to understand the relation between the agent's mental states, the environment, and its behavior, because the mindreader uses the mental state (how the agent perceives the environment/how it should act) to predict the behavioral response of the agent (Danón, 2016, 216). Danón further argues that more sophisticated minimal mindreaders can be described as attributing more complex PPs to others. The attribution of such PPs to others gives a mindreader more flexibility and predictive power than the attribution of basic PPs. Here, again, Danón follows Millikan's lead. Millikan describes more advanced types of PPs. There are at least two such types: pushmi-possible pullyu representations (PPPs) and pushmi-multiple pullyu representations (PMPs) (Millikan, 2004, Chap. 17). PPPs are more complex than pure PPs because "there is not an imperative content compelling the animal to act no matter what, but rather a 'potentially directive face'" that will guide

[5] Millikan calls these representations "PPRs" and Danón calls them "PPs," but both refer to pushmi-pullyu representations.

the animal's behavior *if* certain psychological conditions are met (Danón, 217). For example, an animal might see food nearby, but will only go get it if it is hungry. Pushmi-multiple pullyu representations (PMPs) represent various possibilities of action. Danón explains that for Millikan, PMPs allow animals to "represent the same object as affording different possibilities of action in different situations, and consequently compel them to act in different ways toward the object" (219).

Danón suggests that the minimal mindreader M that attributes PPPs to a creature, C, will attribute to C a descriptive/potentially directive mental state (222). The mental state indicates "what it is out there in the world" and "what C should do about it" – but only "if C happens to be appropriately motivated" (222). For example, if M is trying to predict the behavior of a rival creature C that is standing in front of a piece of food, M will attribute to C the following PPP: Food there/ (if motivated) grab it and eat it (222). The ability to attribute more complex PPPs will allow M to distinguish two types of situations: (1) ones in which C being motivated (hungry) will grab and eat the food and (2) ones in which C, not being motivated (not hungry), will not grab the food and eat it. The mindreader that attributes PPPs needs to represent one mental state with perceptual and prescriptive intentional content as the most basic minimal mindreader does, but this mindreader also needs to represent or track another motivational mental state that activates the prescriptive content. For the prescriptive content of the PPP to be activated, the environmental and psychological conditions need to be right. As a result, the mindreader must keep track of the environmental context and the motivational states of the agent, such as hunger or thirst, to predict accurately whether the animal will act as the PPP directs. For example, if the agent is near a source of food, the mindreader must figure out whether he is hungry using behavioral or other contextual clues to predict whether he will grab the food as the PPP directs (217).

The minimal mindreader M that can attribute PMPs will be able to attribute to C the perception of different possibilities of action embodied in an object O. To predict which possibility is guiding the creature's behavior, M would need to attend to C's motivational states and the environmental context in which C is acting. For example, if C is playing, C will use the tree to climb to do stunts, but if there is a predator nearby and C is scared, then C will use the tree as a hiding place (220). Thus, if M can attribute PMPs, M can use different PMPs to predict and explain C's different behavior in different contexts. This minimal mindreader that attributes PMPs is the most sophisticated minimal mindreader because it can represent a range of different perceptions of the same object and a range of different internal goals towards the same object. The PMP mindreader can attribute different mental states to an agent in different contexts. In addition, the PMP mindreader has the most sophisticated understanding of the relation between the mental states of the agent and its environment and behavior. To decide which perception of the object and which goal towards the object the agent C has, the mindreader must understand C's environmental conditions, necessities, and psychological conditions (220). For example, to determine whether C will use the tree as a hiding place or use it to play, the mindreader needs to understand how C perceives the tree based on whether C is scared or being playful and whether there is a predator nearby or not. This understanding appears to require a high level of sophistication on the part of M.

As noted earlier, Danón takes a step in the right direction, by not requiring Readers to engage in reasoning about the inferential processes of others. However, her account of minimal ToM faces a similar difficulty to the one faced by B&A's account. On Danón's account, minimal mindreading is not significantly different from traditional or full-blown mindreading, since it is demanding along the same dimensions as full-blown mindreading. Even though *what* the Readers are attributing is simpler because the Readers only attribute to others pushmi-pullyu representations rather than detached mental states, such as beliefs and desires, the *process* of attribution is not significantly simpler. The Readers still need to form beliefs about the mental representations of others, such as whether the Agent is in the mental state of a PP, PPP, or a particular PMP. In addition, Danón's Reader may also have to infer PPs that they attribute to others from their observable behavior, since these PPs are internal representations. The Readers that attribute more complex PPs to others, such as PMPs also need to use the psychological conditions of Agents, such as their motivations, to determine which mental state they are in or which pushmi-pullyu representation to attribute to them. And to engage in this process of mindreading, the Reader needs to have meta-representational capacities as well as a conceptual understanding of these simpler mental states and their relation to the Agent's behavior and external world to make accurate predictions. Given all this, Danón's account does not significantly reduce the demanding requirements of full-blown or traditional ToM.

If we are to articulate a genuine middle-ground between behavior-reading and full-blown mindreading, we should focus on making the entire process of mindreading minimal for the Reader and not 'go minimal' only on *what* the Reader is attributing to others. An adequate account of minimal mindreading would be minimal along all the relevant dimensions that I mentioned earlier yet credit the Reader with more than mere behavior-reading. In the next section, I argue that providing such an account requires departing more radically than existing accounts do from the traditional framework of ToM. It requires making a distinction between what the Reader represents during mindreading and how the Reader represents it, and then providing a way of making the how or process of mindreading less cognitively demanding for the Reader by moving away from mental state attribution and inference.

4 An alternative Millikanian Mindreading Account

In the last section, I gave a brief overview of Millikan's framework for understanding mental representation, which applies to both primitive and more sophisticated mental representations. Although Millikan only comments on theory of mind in passing and does not address minimal theory of mind at all, I think Millikan's framework can be used to provide a promising account of minimal ToM, since her theory gives us a way of explaining how non-human animals and infants could *represent* mental states without deploying complex social cognitive abilities. The general idea of my alternative Millikanian account of minimal mindreading is that mental state *representation* can come apart from mental state *attribution*. *What* the Reader represents will be determined by the proper function of the representation; *how* the Reader represents will also depend on the Reader's cognitive capacities and the kind of representations

that *the Reader* can form. A conceptually equipped Reader that can engage in full-blown mindreading will represent mental states differently from a more cognitively 'naïve' Reader that can only engage in minimal mindreading. I will begin by discussing the relevant aspects of Millikan's framework and then use them to illustrate my account of minimal mindreading and distinguish it from both full-blown mindreading and behavior-reading accounts.

In *Beyond Concepts* (2017), Millikan briefly explains how animals can represent mental states without having any theoretical understanding or concepts of mental states (as traditionally understood). Millikan points out that it has been traditionally assumed that representing a mental state requires having a 'factic' thought[6] that involves descriptive or factual knowledge of mental states. And having such thoughts has been presumed to require having some conceptual understanding of what mental states are, how they work, or how they guide the Agent's actions and behavior. But, according to Millikan, animals can represent mental states through *affording* knowledge, as opposed to *knowledge of facts about* mental states (2017). She says, "one needs only to be able to recognize it [the mental state] in some way or ways so as to either collect information about it or learn how to deal with it"[7] (2017, 104). For instance, a dog can recognize the squirrel's intention to escape up a tree by recognizing the squirrel's goal-directed behavior and react to prevent the squirrel from escaping. Millikan explains that the *dog's* perception of the squirrel's intention is *itself* a pushmi-pullyu representation (PPR) - the descriptive side represents the squirrel's intention to escape while the directive side guides the dog to respond so as to block its escape (2017).[8]

The general framework of teleosemantics can be used to explain mental state representation on this alternative Millikanian account. According to teleosemantics, "representation" is a function term[9] and the content of a representation is determined by its *proper function*, which is "a kind of function that can result from selection processes operating on any level" (2017, 6).[10] Proper functions, Millikan explains, "are effects of devices that have, to speak strictly, been retained (not designed) despite selection pressures and that continue to be duplicated or reproduced because they are producing these effects" (6).[11] So, the content of a representation is not determined by

[6] A factic thought is a purely descriptive propositional attitude (Millikan, 2017, 67), such as the belief that snow is white. An example of a factic thought about someone else's mental state would be the descriptive belief that Kelly knows/believes that the banana is in the box.

[7] Millikan (1995) is relying on the Gibsonian idea that "in perception we perceive certain affordances (opportunities for action)" which she argues suggests that "perceptual representations are PPRs" (191).

[8] Plausibly, as Danón suggests, the squirrel's motivational state may itself only be a PPR (or perhaps a PPP/PMP), rather than a full-blown (human-like) intention – in which case what the dog represents is only an intention-like PPR.

[9] Millikan (2020) provides an overview of this framework of representation and discusses the convergence between the understanding of the notion of representation on this framework and the understanding of representation in neuroscience.

[10] For a detailed discussion of teleosemantics, see Millikan (1989), (1995), (2004) and, most recently, (2020) in which she defends the claim that representation is a function term, arguing that "[i]f it were not there would be no standard by reference to which representations would be correct or incorrect, true or false, fulfilled or unfulfilled, satisfied or unsatisfied" (2).

[11] See Chaps. 1 and 2 of Millikan (1984) for a precise explanation of the notion of "proper function."

something particular the subject knows about the extension of the representation or by the subject's conceptual understanding of the nature of the extension (2017, 72). Instead, the content of the representation is fixed by its proper function: "What an inner representation represents is just whatever state of affairs needs to correspond to it if its interpreters' designed responses to it are to serve their functions in a Normal way" (2017, 105).[12]

Thus, consider again the squirrel and dog example. The dog's PPR must be a *representation of* the squirrel's intention because "it will produce its designed effects in accordance with a Normal explanation only if it indeed corresponds to such an intention" (2017, 106). The main idea that is relevant here is that, to have a representation of an Agent's mental state, the Reader does not need to have conceptual understanding or informed beliefs about it. Nor does the Reader need to make inferences to or from the represented mental state. Instead, the Reader can be said to have a PPR of the Agent's mental state because the Reader has an internal state that is a representation whose content is fixed by the proper function of the representation – a content that needs to correspond to the mental states of the Agent to produce its designed effects in a Normal way. On Millikan's view, then, "the dog no more needs to grasp the true nature of squirrel intentions in order to represent and take account of them than you or he needs to grasp the true nature of water – that it consists of two hydrogen and one oxygen atom bound together, that a hydrogen atom consists of… and so forth – in order to represent it" (2017, 106).

If it is possible to represent a mental state without having any conceptual understanding of it, we can articulate a form of minimal mindreading that avoids the overly intellectualized traditional conception of mindreading, without committing to the idea that minimal mindreaders *only* represent Agents' behavior. On this account, the *function* of mindreading for the Reader is not to make sense of, explain, or even explicitly predict the behavior of Agents. Rather, the function of mindreading is to guide the Reader's own actions or behavior relative to other Agents in the social environment. The Reader is sensitive to others' minds and mental states because they *matter* to her; but this concern is a *practical* one: it is limited to knowing how to deal with others' minds or how to respond to them.

Since the function of mindreading on the present Millikanian proposal is to guide practical action, the *process* of mindreading can be genuinely minimal on this account as well. The Reader represents mental states using perceptions of affordances. The idea is that perception consists in non-inferential recognition of affordances, which are "aspects of the environment that afford the possibility of various activities" for the perceiving animal (Millikan, 2004, 159). It consists in the recognition of abstract patterns that guide action (Millikan, 2004). These perceptions of affordances are themselves pushmi-pullyu representations. So, the Reader would perceive mental states and directly respond to them using pushmi-pullyu representations whether they are pure PPs or complex PPs. As a result, the Reader would not be engaging in infer-

[12] Millikan clarifies: "What is normal (common, average) is by no means always Normal, not always of a kind that has helped to proliferate the mechanisms involved during the historical selection processes that formed them" (2017, 85). A detailed analysis of this special use of "Normal" is given in Millikan (1984) Chaps. 1–2.

ential reasoning that culminates in the attribution to Agents of propositional attitudes *or* PPs, where these are understood as 'mediating variables' between behavior and action. The process of mindreading would be perceptual rather than inferential.

In order to understand how the process of mentalizing could be perceptual rather than inferential, we need to consider the way Millikan's view goes against the traditional idea that only what is 'concrete' can be perceived. Millikan's overall model contrasts with traditional inferential models of perception on which perception requires a form of abductive inference or inference to the best explanation, so that even perceptual judgments of the most ordinary things involve inferential processes that require having and employing concepts (Millikan, 2017, 102). By contrast, according to Millikan, perceptual judgments can be made based on experience or signs of things without using concepts of them or even relying on what Millikan calls "unicepts". In Millikan's framework, *unitrackers* and *unicepts* replace traditional concepts. A *unitracker* is a "same-tracking mechanism that tracks information, affording or factual, that concerns some one particular thing (or kind, etc.), its target" using many same-tracking techniques and by recognizing its target in many ways (2017, 42). A unitracker collects this information over time and stores it as information concerning the same thing by linking it to a single unicept. A *unicept* "takes various items of information about one and the same thing and holds them together so that they can be used together" (44). There are two types of unicepts: affording unicepts and factic unicepts. Affording unicepts store affording knowledge that allows individuals to deal with the stuff in the world, while factic unicepts store factual knowledge that includes judgments about the properties and attributes of things in the world (2017, 62–68). Consider the following example of how perceptual judgments can be made without even unicepts: an individual can identify lemons using their color without relying on the unicept for yellow or going through an inference that has 'This is yellow' as a premise (2017, 103). The way an individual can make such perceptual judgments can be explained by Millikan's theory of infosigns and perception as sign reading.

On Millikan's framework, natural information is carried by 'infosigns' – a category that includes natural signs, such as smoke and rashes, but also some intentional signs, such as linguistic symbols and maps. Infosigns are "signs of states of affairs" which are "chunks of the world" or states that are embedded in the actual world (2017, 110). Millikan takes a position on infosigns that is a middle ground between Dretske's and Peirce's views; information, on her view, is objectively found in the world, but what constitutes information is dependent on interpreters (2017, 146). Information is best thought of as "food for cognition": what counts as food depends on animals' particular digestive system; similarly, what can serve as a sign depends on the peculiarities of animals' cognitive systems and their environment (146). Since information is food for cognition, perception is the consumption of information, which, in turn, requires some form of digestion: the "translation of the natural informational content carried by patterns in sensory data into inner intentional signs, either beliefs or representations of affordances" (2017, 185).

Millikan further explains that it is traditionally assumed that you cannot simply perceive or "see that Peter is unhappy, smell that bacon is frying, hear that wind is blowing or that the engine is missing on a cylinder" (2017, 194). The traditional

idea is that these examples do not constitute mere perceptual knowledge and require inference, since "only concrete individuals are given in sensory experience" (2017, 200). Millikan argues, however, that there is no principled restriction on the range of things that can be perceived. If we take "perception to be a process of translation from infosigns into thoughts, perception can be of causally distal affairs, it can be highly fallible, it can be layered, capable of focusing at different depths, it can interpret indirect as well as direct signs" (201).[13] On this model of perception then, it is possible to "hear that someone is practicing trumpet, to taste that there is oregano in the spaghetti sauce, to smell that Grandma is frying bacon" because perceptual processing is a "silent process" of translating the natural information in sensory data into representations without inference (2017, 192–195).

Millikan's account of perception as reading infosigns can be applied to mindreading so as to show how the Reader can represent mental states through perception non-inferentially, even if it is denied that mental states are exhausted by behavioral patterns.[14] Using Millikan's framework, we would reject the traditional dichotomy between the observable and unobservable in which only bodies and their behavior are observable and mental states are not. On this alternative account, mental states would not be necessarily hidden or unobservable states that must be *inferred from* observable behaviors or other mental processes. Instead, individuals can directly perceive mental states just as they can perceive concrete things *or* abstract perceptual objects in the world. For example, a wasp can perceive the angry state of another other wasp by observing its face. During this perceptual process, the information about anger that is carried by abstract patterns in sensory data will be turned ('translated') into inner intentional signs like PPRs or representations of affordances without any inferential processes or conceptual understanding of mental states involved. The descriptive side of the PPR would represent the perceived state of affair (anger) while the directive side would guide the perceiver's action or how to respond to such a state. Thus, on this account, mental states can be directly perceived by Readers and representation of mental states does not require inference or concepts of mental states.

Moreover, representation of mental states does not require mental state attribution on this account. The reason that the Reader does not need to engage in mental state attribution on this view can be explained by the nature of their representations of mental states. The Reader that engages in minimal mindreading on this view does not have factic inner representations of others' mental states that require assigning or attributing properties to their subjects. The Reader has PPRs or representations of affordances. Rather than "attributing properties to situations," these "serve to potentiate actions or ways of being guided by perceived conditions towards certain ends" (Millikan, 2017, 65). The idea is that the Reader does not distinguish between ontological categories, such as "properties, individuals, materials, kinds, and so forth" or engage in making factual judgments about the properties and nature of the mental states of individuals (2017, 62). The Reader merely recognizes them as the same

[13] For a detailed analysis of this view of perception, see Chap. 14 in *Beyond Concepts* especially pgs. 190–202.

[14] I am grateful to an anonymous reviewer for suggesting that I incorporate Millikan's theory of infosigns and perception as sign reading into my account of minimal mindreading.

"stuff-again" for the specific purpose of "immediate practical manipulation and action" (2017, 64). For example, a child can know how to recognize and immediately deal with a certain mood that the parent sometimes gets in without understanding the nature of the mood, categorizing the mood as a property of the parent, or attributing it to the parent (64).

The Millikanian account I am proposing, then, departs much more radically than other accounts of minimal mindreading from the traditional view, since it does not take the Reader to engage in any explicit attribution of mental states to Agents[15]. Whether or not this account can be read as positing *implicit* mental state attribution would depend on how 'implicit' is understood. If implicit attribution requires possession and employment of *concepts* of mental states or (at least) factic unicepts, then the present account rejects the idea that mental state representation involves even implicit attribution of mental states, since it allows that individuals can represent mental states despite lacking such concepts or factic unicepts altogether. I take this to be an advantage of the account, given that there is no clear understanding of what implicit mindreading requires currently on offer (Low & Perner, 2012; Gomez, 2017). Thus, on my proposed minimal mindreading account, the Reader does not need to employ *concepts* of mental states and have factual knowledge of mental states and their relation to the Agent's behavior and environment to engage in mindreading – either explicitly *or* implicitly. All she needs is a practical understanding of others' mental states that enables her to recognize them in order to deal with them. The Millikanian framework gives us a way of showing how the Reader can have *affording (practical) knowledge* rather than factual (theoretical) knowledge of mental states, where affording knowledge, to repeat, is not conceptual knowledge and does not involve factic beliefs about its subject-matter.

Although minimal mindreading on my proposed account would require a practical understanding of mental states that involves representing and dealing with them via representations of affordances, the conceptual requirements that are needed to engage in minimal mindreading would depend on where we determine that minimal mindreading 'proper' begins. The account makes room for different ways of representing mental states. Thus, consider Millikan's distinction between *uni*tracking, on the one hand, and *merely* tracking something, or *same-tracking* it.[16] There are "non-uniceptual same-tracking mechanisms, mechanisms that same-track not in order to implement storage of information about their targets, but merely as an aid to the identification of further things" (2017, 56). In general, same-tracking mechanisms do not require unitrackers or unicepts. For example, young children same-track the

[15] An anonymous reviewer raises the worry that PPRs might themselves be attributive because they are presented as having indicative and directive contents or as belief-desire pairs that are bound up together. But it is important to keep in mind that a Millikanian PPR does not amount to the *conjunction* of a pure descriptive representation and a pure directive representation. It is more primitive than either type of representation (Millikan, 1995). Thus, PPRs should not be conceptualized as belief-desire pairs with contents in the form of conjunctions. And as I explained earlier, the descriptive aspect of PPRs does not involve attributing properties to situations or individuals. For this reason, PPRs of mental states need not involve attributing mental states as psychological properties of agents.

[16] See Chap. 4 of *Beyond Concepts* for an overview of the distinction between same-tracking and unitracking. Same-tracking something or merely tracking it in the environment does not require individuals to have unitrackers or unicepts of that thing.

phonemes that make up linguistic strings for the purpose of understanding speech, but they do not have either unitrackers or unicepts *for phonemes*; in fact, they have "difficulty in acquiring 'phonological awareness,' difficulty in acquiring ideas of phonemes as identifiable individual units so as to match them with letters" (2017, 58). Another example is a young child who can identify a banana by its color without knowing any words for colors or using a unicept of yellow; adults probably identify bananas in the same way (58). Millikan explains that children "have not had reason to collect information about or to develop skills regarding these properties" of colors (or of shapes) and "their abilities to same-track them have been used only in the service of same-tracking individuals and kinds that displayed those properties" (58). The overall idea is that one can track something without having any way of collecting and storing knowledge about that thing.

Going back to mindreading: a Reader that represents mental states only via same-tracking mechanisms would be able to recognize and respond to certain mental states (using specific PPRs), but does not have unicepts for mental states to store information *about* those mental states. Since the Reader does not store information about mental states, the Reader cannot learn from experience new ways of recognizing the same mental states or modify its responses to those mental states over time. A Reader that has affording unitrackers and affording unicepts for mental states, on the other hand, has the mechanisms to store information about the represented mental states and develop their abilities to recognize and practically deal with mental states based on what they learned from their past experiences. Both Readers are representing mental states, but do they both count as minimal mindreaders? It could be that in order to count as a minimal mindreader, the Reader needs to go beyond mere same-tracking and have unicepts of mental states that allow the Reader to have a more flexible practical understanding of mental states. But perhaps there are different types of minimal mindreaders – some more primitive, others more sophisticated. This question requires further theorizing, so I will leave open where minimal mindreading 'proper' begins for now, but in the next section, I present one possibility that can be explored further.

This alternative Millikanian minimal ToM account minimizes the requirements on the *Reader's* process of mindreading, unlike the accounts discussed earlier. Danón's account, for example, minimizes the nature of the mental states that are represented and attributed to others by the Reader. But it does not portray the nature of the representations that the Reader needs to form to engage in minimal mindreading as minimal. On Danón's account, the Reader's representations of the primitive mental states of others are still meta-representations that are caught up in mental state attributions. B&A's account also focuses on minimizing the mental states represented by the Reader, but the process of minimal mindreading seems to remain inferential and requires mental state attribution and thus the use of concepts of (perhaps more primitive) mental states. As I argued earlier, to provide an account of mindreading on which the process of mindreading is minimal, these accounts need to depart from the traditional view of mindreading more radically. They need to question – and reject – the assumption that conceptually-informed, inferentially-driven mental state attribution is required for mindreading.

This is what my alternative Millikanian account does, by explaining how the Reader can recognize and respond to mental states without engaging in conceptual-inferential mental state attribution. Using the Millikanian framework of perception as sign reading, the account proposes that mental states can be directly perceived by the minimal mindreader, despite such states not being merely behavioral patterns. On my Millikanian account, to engage in mindreading, the Reader does not need to have a proper theory of other minds that requires factic understanding or concepts of mental states and employing them in inferences that result in attributions of mental states to others. Most importantly, this alternative account deploys a separation between the representation of mental states and their attribution. It portrays minimal mindreaders as representing mental states using affording knowledge, which does not require (factic) recognition of agents' psychological attributes. Thus, on this alternative minimal ToM account, the process of mindreading is not as cognitively demanding as on other accounts I have considered (both traditional and more minimal).

Furthermore, my alternative Millikanian account is not a re-description of a behavior-reading account. A Reader who engages in behavior-reading will only represent the behaviors of others and the correlations between different observable cues and those behaviors, but the Reader that engages in Millikanian minimal mindreading represents mental states – though not in the traditional way. In a way, this account calls into question the standard dichotomy between behavior-reading and mindreading, insofar as it rests on the traditional distinction between observation and theory. On the standard view, to repeat, since those with limited cognitive resources cannot infer or postulate unobservable states, their 'mindreading' can at best amount to the tracking (as well as prediction) of behavioral patterns. By contrast, on the account proposed here, there can be a certain continuity between more minimal and more sophisticated mindreaders: both represent mental states, but those with limited cognitive resources represent them in a practical rather than theoretical way. In the next and final section, I discuss how we can distinguish different types of minimal mindreaders using this account. I also briefly discuss the empirical and practical advantages of this account and raise some questions to be addressed by future research.

5 Applications & advantages of alternative account and further research questions

The Millikanian mindreading account proposed here can be used to distinguish types of minimal mindreaders that engage in more or less sophisticated kinds of mindreading. It could be that the most minimal mindreaders only same-track the mental states of others without collecting any information about those states or developing skills that have to do with mental states. These primitive minimal mindreaders would simply track mental states for the purpose of doing other things as children track phonemes in their understanding of speech. For example, it might be that the wasp is a primitive minimal mindreader that can same-track the aggressive states of other wasps as a way of avoiding or dealing with aggressive wasps but does not have any mechanisms for collecting and storing information about the mental state of aggression itself. A more sophisticated minimal mindreader would move beyond

mere tracking of mental states by having affording unitrackers and unicepts for mental states. This minimal mindreader would collect and store information about mental states themselves by way of developing their practical skills regarding them. The reason that a minimal mindreader that has affording unicepts for mental states would be more sophisticated than the primitive minimal mindreader is that this minimal mindreader would be able to learn through experience new ways to recognize mental states and be able to modify its responses to those mental states over time, which the more primitive minimal mindreader cannot do. For instance, if the dog from the example earlier collects and stores information about squirrel intentions, in order to guide its own behavior, the dog can learn to recognize these intentions in new ways and modify its response to them accordingly. Then the dog would be a more sophisticated minimal mindreader than the wasp because the dog can develop its ability or skill to practically deal with mental states over time.[17] The sophisticated minimal mindreader, however, would still be distinguished from a full-blown mindreader that has unitrackers and unicepts for collecting and storing factual knowledge about mental states and their properties; this mindreader can use her factic knowledge of mental states for theoretical purposes, such as explaining and understanding the behavior of other agents.

This helps us see how my proposed account of minimal mindreading can allow us to distinguish between different types of mindreading. In addition, the account may also be useful for empirically investigating ToM abilities in human infants and non-human animals, as well as for interpreting the results of some empirical studies. For example, consider how the behavior of the chimpanzees in the Hare et al., (2001) study discussed earlier could be interpreted using this account. It is possible that the subordinate chimpanzees have unitrackers and unicepts for the perceptual states of the dominant chimpanzee that allow the subordinates to collect and store affording knowledge about these mental states. This affording knowledge can allow the subordinate to recognize the dominant's different perceptual states in different contexts – e.g. the informed, uninformed, and misinformed conditions – and then guide the subordinate to respond in different ways.[18] For instance, when the subordinate recognizes that the dominant has seen the food, it would react by not grabbing the piece of food and when the subordinate does not represent any perceptual awareness of the food's location from the dominant or represents the dominant's false-beliefs about the food's location, the subordinate would react by grabbing the food. This recognition of – and responses to – the presence and character of mental states can occur without any conceptual attribution of mental states or theoretical understanding of others' minds.

A similar analysis can be provided for the results of Hare et al., (2000), Kaminski et al., (2008), and Karg et al., (2015) that also involve competitive tasks. The behavior of infants and chimpanzees in false-belief tests that use an interactive helping paradigm (Buttelmann et al., 2009, 2017), in which infants and non-human animals

[17] Whether the dog is actually a sophisticated minimal mindreader would be an empirical question that cannot be answered here.

[18] The subordinate chimpanzee could use pushmi-multiple pullyu representations (PMPs) to represent the dominant's different perceptual states and different responses to those states.

have to be sensitive to the false-beliefs of other agents in order to successfully help them achieve their goals, can also be explained using my alternative account of minimal mindreading. It could be that the presence of minimal mindreading abilities or a minimal, practical understanding of others' mental states can be more easily measured using these interactive paradigms that involve or require behavioral responses from the subjects to demonstrate their sensitivity to others' mental states, such as the false beliefs of other agents. The fact that infants and non-human animals do well on these interactive and non-standard tasks, even though they do not do well on standard false-belief tasks, might suggest the presence of a minimal form of mindreading that allows infants and non-human animals to recognize and respond to mental states using their affording knowledge of mental states.

Furthermore, even adults can rely on this more minimal form of mindreading when under cognitive load or in everyday situations. For example, in daily life, human adults probably do not spend a lot of time theorizing about the mental states of others or explicitly explaining and understanding their behavior using full-blown mindreading. In these everyday interactions with other agents, adults can simply rely on their affording knowledge to recognize and immediately respond to the mental states of others. The reliance on practical knowledge in these situations would make sense because it would save time and precious cognitive resources that can be reserved for complex social situations.

The details of this alternative Millikanian account, however, would need to be worked out. For example, the relationship between pushmi-pullyu representations and affording unicepts needs to be investigated and explained further. Minimal mindreaders on this account treat mental states as affordances that potentiate actions, where grasping a current affordance would involve inner PPRs. But doesn't forming these PPRs, in turn, require the Reader to have affording unitrackers and affording unicepts? If not, then at what point are unicepts of mental states required for mindreading? It could be that unicepts are required for more sophisticated forms of mindreading that require the ability to form complex PPs, such as PMPs, that allow the Reader to recognize mental states in multiple ways and respond to them in different ways. It might be that unicepts are needed for developing general abilities that allow the Reader to interact with the mental states of others more flexibly than a Reader that can only form pure PPs that are inflexible (there is only one specific response). But this possibility would need to be further explored and a more detailed explanation will have to be provided for how unicepts aid the development of general abilities for recognizing and appropriately dealing with different mental states. Moreover, PPRs are themselves under-theorized, so a full development of the Millikanian account of minimal mindreading will need to address questions about their nature.[19]

[19] Artiga (2014), for example, is skeptical about the very existence of PPRs and has raised problems for teleosemantic theories that endorse PPRs. Bauer (2020) offers some responses to Artiga's objections and provides some support for the role of PPRs in explaining primitive systems. Millikan herself has acknowledged that the "very possibility of pushmi-pullyu representations requires defense" and has offered such a defense and responses to Artiga's objections in (2021). However, analyzing the objections and evaluating possible responses to them goes beyond the scope of this paper.

6 Conclusions

In this paper, I presented the traditional view of mindreading and the cognitively demanding requirements that it places on mindreaders by making conceptual-inferential mental state attribution necessary for mindreading. I argued that providing a minimal account of mindreading on which the process of mindreading is genuinely minimal requires departing from the traditional framework in more radical ways than current minimalist accounts of ToM do. I proposed an alternative view of minimal mindreading that provides a way for minimal mindreaders (viz. infants and non-human animals) to represent mental states without engaging in mental state attribution, which requires concepts of mental states, harboring meta-representational beliefs and making inferences about them. Further developing this proposed account can help us not only identify minimal forms of mindreading but also explain the origins and the development of theory of mind abilities by establishing continuity between non-human animals, infants, and adult humans.

Acknowledgements I would like to thank Dorit Bar-On for helpful feedback, detailed comments, and many conversations throughout the various stages of writing this paper. I am grateful to Ruth Millikan for inspiring me with her work and discussing the ideas in this paper with me. I am also thankful to William Lycan and Richard Moore for their comments on an earlier draft of this paper, and to the anonymous reviewers for feedback that led to valuable improvements in my paper. An earlier version of this paper was presented to the Expression, Communication, and Origins of Meaning (ECOM) research group (December 2020); I wish to thank members of the group for their feedback. The writing of this paper was in part supported by an ECOM fellowship (summer 2020).

Funding Not applicable.

Availability of data and material Not applicable.

Code Availability Not applicable.

Declarations

Conflict of interest None.

References

Anscombe, G. E. M. (1957). *Intention*. Ithaca: Cornell University Press

Artiga, M. (2014). Teleosemantics and Pushmi-Pullyu Representations. *Erkenntnis*, 79, 545–566. https://doi.org/10.1007/s10670-013-9517-5

Avramides, A. (2020). Other Minds. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/other-minds/#NatuTurn

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in cognitive sciences*, 14(3), 110–118. https://doi.org/10.1016/j.tics.2009.12.006

Barlassina, L., & Gordon, R. M. (2017). Folk Psychology as Mental Simulation. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.https://plato.stanford.edu/archives/sum2017/entries/folkpsych-simulation

Bauer, M. (2020). The explanatory breadth of pushmi-pullyu representations. *Biol Philos*, 35, 35. https://doi.org/10.1007/s10539-020-09751-5

Beaudoin, C., Leblanc, É., Gagner, C., & Beauchamp, M. H. (2020). Systematic Review and Inventory of Theory of Mind Measures for Young Children. *Frontiers in psychology*, 10, 2905. https://doi.org/10.3389/fpsyg.2019.02905

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, 77(1), 25–31

Buttelmann, D., Buttelmann, F., Carpenter, M., Call, J., & Tomasello, M. (2017). Great apes distinguish true from false beliefs in an interactive helping task.PLoS One, 12(4), e0173793

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112(2), 337–342. https://doi.org/10.1016/j.cognition.2009.05.006

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

Butterfill, S. A., Apperly, I. A., Rakoczy, H., Spaulding, S., Zawidzki, T. W. Symposium on, Butterfill, S., & Apperly, I. (2013). "How to Construct a Minimal Theory of Mind." https://www.semanticscholar.org/paper/Symposium-on-S.-Butterfill-and-I.-Apperly%2 C-%22How-to-Butterfill-Apperly/a3b55ed9ceb0fb26d8cce75e438ac2ae85a82172

Carruthers, P. (1996). Simulation and self-knowledge: a defense of theory-theory. In P. Carruthers, & P. R. Smith (Eds.), *Theories of theories of mind* (pp. 22–38). Cambridge: Cambridge University Press

Carruthers, P. (2013). Mindreading in infancy. *Mind Lang*, 28, 141–172. doi:https://doi.org/10.1111/mila.12014

Churchland, P. M. (1991). Folk psychology and the explanation of human behavior. In J. D. Greenwood (Ed.), *The future of folk psychology* (pp. 51–69). Cambridge: Cambridge University Press

Danón, L. (2016). Pushmi-pullyu representations and mindreading in chimpanzees. *Studia Philosophica Estonica*, 9(1), 208–236

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

Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science*, 306, 1903–1907

Emery, N. J., & Clayton, N. S. (2008). How to build a scrub jay that reads minds. In S. Itakura, & K. Fujita (Eds.), *Origins of the social mind: Evolutionary and developmental perspectives* (pp. 65–98). Tokyo: Springer

Goldman, A. I. (1989). Interpretation psychologized. *Mind & Language*, 4, 161–185

Gomez, J. C. (1996). Non-human primate theories of (non-human primate) minds: Some issues concerning the origins of mindreading. In P. Carruthers, & P. Smith (Eds.), *Theories of Theories of Mind* (pp. 330–343). Cambridge: Cambridge University Press

Gomez, J. C. (2017). Knowing without knowing: implicit cognition and the minds of infants and animals. *Studies in Psychology*, 38(1), 37–62

Gopnik, A., & Wellman, H. M. (1992). Why the Child's Theory of Mind Really Is a Theory. *Mind & Language*, 7(1–2), 145–171

Gordon, R. M. (1986). Folk Psychology as Simulation. *Mind & Language*, 1(2), 158–171. doi:https://doi.org/10.1111/j.1468-0017.1986.tb00324

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behavior*, 61, 39–151

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

Heal, J. (1998). Understanding Other Minds from the Inside. *Royal Institute of Philosophy Supplement*, 43, 83–99

Heyes, C. (2014). False belief in infancy: a fresh look. *Dev Sci*, 17, 647–659. https://doi.org/10.1111/desc.12148

Heyes, C. (2018). *Cognitive Gadgets: The Cultural Evolution of Thinking*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press

Hutto, D., Herschbach, M., & Southgate, V. (2011). Social cognition: mindreading and alternatives. *Review of Philosophy and Psychology*, 2, 375–395

Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, 109(2), 224–234. https://doi.org/10.1016/j.cognition.2008.08.010

Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2015). Chimpanzees strategically manipulate what others can see. *Animal Cognition*, 18(5), 1069–1076. https://doi.org/10.1007/s10071-015-0875-z

Kovacs, A. M., Teglas, E., & Endress, A. D. (2010). The social sense: susceptibility to others' beliefs in human infants and adults. *Science*, 330(6012), 1830–1834

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, 354(6308), 110–114. doi: https://doi.org/10.1126/science.aaf8110

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind. *Psychological Review*, 94(4), 412–426

Low, J., & Perner, J. (2012). Implicit and explicit theory of mind: State of the art. *British Journal of Developmental Psychology*, 30, 1–13

Lurz, R. (2011). *Mindreading Animals: The Debate Over What Animals Know About Other Minds*. Cambridge, Massachusetts: The MIT Press

Millikan, R. G. (1984). *Language, thought, and other biological categories: New foundations for realism*. Cambridge MA: MIT Press

Millikan, R. G. (1989). Biosemantics. *Journal of Philosophy*, 86, 281–297

Millikan, R. G. (1995). Pushmi-pullyu representations. *Philosophical Perspectives*, 9, 185–200

Millikan, R. G. (2004). *Varieties of Meaning*. Cambridge, MA: The MIT Press

Millikan, R. G. (2017). *Beyond Concepts: Unicepts, Language, and Natural Information*. Oxford: Oxford University Press

Millikan, R. G. (2020). Neuroscience and teleosemantics. *Synthese*, 199, 2457–2465. https://doi.org/10.1007/s11229-020-02893-9

Millikan, R. G. (2021). Comment on Artiga's "Teleosemantics and Pushmi-Pullyu Representations. *Erkenntnis*. DOI:https://doi.org/10.1007/s10670-020-00354-w

Moore, R. (2020). The cultural evolution of mind-modelling. *Synthese*. https://doi.org/10.1007/s11229-020-02853-3

O'Madagain, C., & Tomasello, M. (2019). Joint attention to mental content and the social origin of reasoning. *Synthese*. Doi: https://doi.org/10.1007/s11229-019-02327-1

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255–258. doi: https://doi.org/10.1126/science.1107621

Penn, D., & Povinelli, D. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical transactions of the Royal Society of London*, 362(1480), 731–744

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, 31, 109–178

Powell, L. J., Hobbs, K., Bardis, A., Carey, S., & Saxe, R. (2018). Replications of implicit theory of mind tasks with varying representational demands. *Cogn. Dev*, 46, 40–50. doi: https://doi.org/10.1016/j.cogdev.2017.10.004

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci*, 1, 515–526. doi: https://doi.org/10.1017/S0140525X00076512

Rakoczy, H. (2017). In defense of a developmental dogma: Children acquire propositional attitude folk psychology around age 4. *Synthese*, 194(3), 689–707

Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? Response to Leslie. *Trends Cogn Sci*, 10, 462–463. doi: https://doi.org/10.1016/j.tics.2005.08.001

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychol. Sci*, 18, 587–592. doi: https://doi.org/10.1111/j.1467-9280.2007.01944.x

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, 18(7), 580–586

Tomasello, M. (2018). How children come to understand false beliefs: A shared intentionality account. *Proceedings of the National Academy of Sciences*, 115 (34), 8491–8498

Tomasello, M., & Call, J. (2006). Do chimpanzees know what others see—or only what they are looking at?. In S. Hurley, & M. Nudds (Eds.), *Rational animals* (pp. 371–384). Oxford: Oxford University Press

Van der Vaart, E., & Hemelrijk, C. (2014). 'Theory of mind' in animals: ways to make progress. *Synthese*, 191, Doi: https://doi.org/10.1007/s11229-012-0170-3

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, 72(3), 655–684

Whiten, A. (2013). Humans are not alone in computing how others see the world. *Animal Behavior*, 86, 213–221

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128