

# Pushmi-pullyu Representations and Mindreading in Chimpanzees

Laura Danón

Faculty of Philosophy and Humanities and Faculty of Psychology, National University of Córdoba

---

Lurz and Krachun (2011) propose a new experimental protocol designed to discriminate genuine mindreading animals from mere behavior-readers and to give evidence in favor of the claim that chimpanzees are capable of attributing internal goals to others. They suggest that chimpanzees' variety of "internal goal attribution" consists in attributing to others basic intentional representations, baptized by Millikan as "pushmi-pullyu representations" (PPs). Now, Millikan (1996; 2004b; 2004a) distinguishes what I propose to call 'pure' PPs from more complex varieties of PPs, which allow their owners to respond more flexibly to their environments. But, what would happen if we tried to differentiate, analogously, between more or less sophisticated mind-readers in virtue of the sorts of PPs that they could attribute to others? What would attributing complex PPs consist in and how would such capacity increase the predictive powers of chimpanzee mind-readers? This paper offers an answer to these questions. Based on Millikan's work, I differentiate two varieties of complex PPs. Then, I examine what a basic mind-reader, only capable of attributing 'pure' PPs, would be able to do. After that, I distinguish two more sophisticated varieties of mindreading, each consisting in the attribution of one of the complex PPs previously presented, and I show how the ability to attribute complex PPs to others comes with more potent and flexible capacities to anticipate their behavior. Finally, I offer some reasons to think that attributing complex PPs is still simpler than full-blown mindreading and I briefly evaluate the prospects of extending this proposal to infant social cognition.

*Keywords:* mindreading, animal cognition, pushmi-pullyu representations, the logical problem

---

According to many prominent researchers, nowadays we have good empirical evidence in favor of chimpanzees having a primitive understanding of

*Corresponding author's address:* Laura Danón, Centro de Investigaciones "María Saleme de Burnichon" Facultad Filosofía y Humanidades (CIFYH), Pabellón Agustín Tosco, 1º piso, Facultad Filosofía y Humanidades, Ciudad Universitaria, Universidad Nacional de Córdoba, Córdoba, cp: 5000, Argentina. Email: ldanon@gmail.com.

some basic mental states of others, like perceptions, intentions, and goals (Call and Tomasello 2008; Tomasello and Call 2006; Call 2007; Tomasello et al. 2003; Call and Jensen 2006). Nevertheless, other well-known researchers argue that the experimental protocols currently in use are not capable of effectively distinguishing genuine mind-readers from mere behavior-readers (Penn and Povinelli 2007; Povinelli and Vonk 2003; Povinelli and Vonk 2004). This methodological flaw has been labeled as “the logical problem.” According to it, the empirical results provided by extant studies do not give us good reasons to conclude that there are any non-human species capable of attributing mental states to others.

Lurz and Krachun (2011) propose a way to advance this debate, both conceptually and empirically, focusing on whether chimpanzees are capable of attributing to others a specific kind of mental state: internal goals (see also Lurz 2011). In order to move forward empirically, they propose a new protocol that can solve the logical problem and, eventually, give support to the claim that chimpanzees attribute internal goals to others. To move forward conceptually, they outline an alternative account of mind-reading—the appearance-reality mindreading theory (ARM)—and suggest that chimpanzees’ variety of “internal goal attribution” consists in attributing to others some basic intentional representations, baptized by Millikan as “pushmi-pullyu representations” (PPs).

In different papers, Millikan (1996; 2004b; 2004a) has characterized pushmi-pullyu representations (PPs) and has suggested a distinction between what I will call ‘pure PPs’ and more sophisticated varieties of PPs, which endow their owners with the capacity to respond to their environments in more flexible ways. By distinguishing kinds of PPs, it becomes possible to account for some differences in the cognitive and behavioral capacities of those non-human animals to which *we* attribute such representations. However, what would happen if we tried to differentiate, analogously, between more or less sophisticated mind-readers in virtue of the sort of PPs that *they* attribute to others? What would being able to attribute complex PPs consist in and how would such capacity increase the predictive powers of animal mind-readers?

In this paper, I will try to answer these questions by proposing a way to distinguish basic mind-readers, who only attribute pure PPs to others, from (slightly) more sophisticated mind-readers, capable of attributing more complex PPs. Even though Lurz and Krachun do not explicitly contemplate this possibility, I think it is compatible with their original proposal and contributes to enrich it. Based on Millikan’s work, I begin by delimiting two varieties of PPs, which are more sophisticated than pure PPs. Then I examine what a basic mind-reader, capable only of attributing pure PPs, would

be able to do. After that, I distinguish two more sophisticated varieties of mindreading, each consisting in the attribution of one of the complex PPs previously presented, and I show how the ability to attribute complex PPs to others comes with more potent and flexible capacities to anticipate their behavior. Finally, I present some reasons to think that the capacities to attribute complex PPs are still simpler than the full-blown mindreading of (adult) human animals and I discuss the potential benefits of extending this proposal to research on the mindreading abilities of human infants.

### **1. Lurz and Krachun's proposal**

In their paper, Lurz and Krachun (2011) focus on whether animals can represent the internal goals of others (understood as motivational mental states that may represent objects or states of affairs that an agent tries to obtain or realize). Following Povinelli and colleagues, they affirm that all current experimental paradigms used to test mindreading abilities in non-human animals are methodologically flawed and face the aforementioned logical problem.

Since our minds cannot be directly observed, Povinelli claims, our attributions of mental states to others must be based, inevitably, on some observable behaviors of the agent and/or on relevant environmental features. Now, current tests make the relevant cues available to the experimental subjects and presuppose that they notice them. But, then, we have two possibilities. On the one hand, a subject's predictions about another agent's future behavior can involve multiple inferential steps—at least a first one from the observable cues to the other's mental states and a second one from the other's mental states to her/his subsequent behavior. On the other hand, the behavioral predictions can be made directly in one step, based on knowledge of the contingent relationship between the relevant observable features and the agent's subsequent behavior. Thus, when some animals predict the behavior of others in a way which suggests they are mindreading, it is always possible to give an alternative explanation, according to which they are exclusively basing their predictions on "behavioral rules" that associate the available behavioral/environmental cues with the consecutive behavior of the agent (Povinelli and Vonk 2004; Povinelli and Vonk 2003; Penn and Povinelli 2007).

Now, Lurz and Krachun think it is possible to overcome the logical problem if we: (i) identify some predictions about the behavior of others that can be made if we attribute mental states to them, but which cannot be made on the basis of externally observable cues alone; and (ii) design an experiment which can test if chimpanzees are capable of making precisely these sort of predictions.

In their opinion, current experimental protocols do not satisfy (i) and (ii). However, they propose a new test designed to distinguish a mere behavior-reader from a chimpanzee capable of attributing internal goals to others. Their proposal is supported by a specific theory of mindreading: the appearance-reality mindreading theory (ARM). According to the ARM, the capacity to attribute mental states to others evolved in some species because, in certain settings, it enabled them to predict the behavior of other agents in a better way than their behavior-reading rivals. In particular, the mind-reader has leverage over behavior-readers in those situations in which an object appears to have properties that it does not really have. In such illusory settings, the way in which a distal object appears to an agent is a better predictor of its behavior than how the object really is. Now, the problem for behavior-reading animals is that they can only appeal to reality-based facts (like the agent's past behavior, its current physical position, etc.) in order to make their predictions. Mind-readers, on the contrary, would also be capable of identifying the subjective modes in which external objects appear to the agents and of employing this information to adequately predict their behavior on illusory settings.

It is important to remark that, since in illusory settings a mind-reader would have predictive capacities that a behavior-reader will necessarily lack, it becomes possible for researchers, at least in principle, to find a way to test such differences satisfying (i) and (ii). With this idea in mind, Lurz and Krachun elaborate a complex experiment to test if chimpanzees are capable of anticipating the actions of a human competitor by understanding how illusory perceptual states look like to him.

Very roughly, in their test, a chimpanzee and a human competitor are placed in adjoining rooms divided by a clear Plexiglass wall. Against the wall, on the competitor's side, there is a table with a sliding platform and two transparent glass containers with opaque lids. Above each of them, on the wall, there are two small holes through which the chimpanzee can stick a finger to choose one of the containers. In front of the competitor, there is an opaque backdrop with two curtained windows, aligned with the containers. The competitor can stick an arm through any of the windows to indicate his choice. An experimenter, present in the room, puts grapes on both containers and slides them within the reach of the contestants.

During the pre-test trials, the human and the chimpanzee take turns choosing the grapes from non-distorting glass containers. When the competitor is absent, the experimenter puts a big grape in one container and a small grape on the other, while the chimpanzee observes those actions, and closes the lids. After that, the competitor enters the room, sits in front of the table and looks at each grape through the glass containers before staring

straight ahead. At that moment, one of two bells rings (one bell indicates that the experimenter is going to slide the table over the competitor allowing him to choose first, the other one indicates that the table will be slid before the chimpanzee which will make the first choice). Then, the platform is slid over the appropriate contestant. When allowed to choose first, the competitor always prefers the big grape to the small one and the chimpanzee is expected to learn this during the pre-test trials.

During the proper test, the same procedure takes place with one exception: in some cases, while the chimpanzee is watching, the grapes are placed inside size-distorting containers with opaque lids. One of the containers magnifies the size of the grape and the other minimizes it. Since the competitor is outside while this happens, he cannot watch the change in the apparent sizes of the grapes once they are inside each container.

What should happen if chimpanzees could attribute perceptual states about how things look to others and make predictions about their future behavior based on such attributions? In such a case, Lurz and Krachun conjecture, chimpanzees should be able to anticipate that the competitor will select the container with the grape that looks bigger when choosing first. Consequently, they should look first to the window in front of the magnifying container after hearing the competitor's bell. Besides, since they understand the difference between how grapes look inside the containers and how they really are, when choosing first they should prefer the grapes that appear smaller, but are in fact the bigger ones.<sup>1</sup>

Now, if chimpanzees were only behavior-readers, they would not be able to replicate these responses. Being incapable of differentiating how things

<sup>1</sup> There is a well-known distinction between two ways in which young children may understand the perspective of others: level 1 and level 2 perspective-taking. The former refers to the basic capacity to understand *what* another person visually perceives or not (Flavell 1992). The latter, instead, refers to a more complex understanding of *how* someone else perceives the world. Now, in order to succeed at Lurz and Krachun's test, chimpanzees would need not only to identify what object—the grape—can be seen by the competitor, but also how the object looks or appears to him. It seems, then, that they would need to go beyond level 1 visual perspective-taking. Nevertheless, it should also be noted that this protocol is not as demanding as the typical tasks of level 2 perspective-taking, which usually require the tested subject to understand that a single object can be seen or conceptualized in different ways from multiple perspectives. Strictly speaking, chimpanzees do not need to understand that the same object—the grape—can be seen in multiple ways from different perspectives, nor do they need to contrast how things visually appear to a competitor with how they appear to them. The grape inside the size-distorting jar looks exactly the same to both the chimp and the competitor. Thus, it seems that chimps only need to distinguish how the grape visually appears (to them and to the competitor) with how it is (See Moll and Meltzoff 2011 for an interesting suggestion on the need to distinguish different kinds of level 2 visual perspective-taking which may help to explain the specific kinds of cognitive abilities involved in Lurz and Krachun's test).

are from how they perceptually appear to others, they would show, instead, one of the following behavioral patterns: (a) anticipate that the competitor will choose the fruit from the magnifying container first, and choose the same fruit themselves (because they also think this is the bigger one); or (b) choose the fruit from the minimizing container when they go first, but expect the same behavior from their rivals (because they think this is where the bigger grape in fact is). Thus, Lurz and Krachun claim that their test allows us to discriminate neatly between mind-readers and behavior-readers by predicting that each of them will give different patterns of responses. They seem to have found, then, an appropriate tool to overcome Povinelli's problem.

To sum up: Lurz and Krachun claim that chimpanzees can only succeed in their experimental protocol by attributing some mental states to their competitors. The reason why they think their design ensures that a mental attribution is taking place is that the task that chimpanzees face can only be solved if they are capable of contrasting how the grape inside the size-distorting jar appears to the naive competitor with how it really is. But, is this really so?

As far as I can see, the only way to challenge such a conclusion is to come up with a plausible alternative behavior-reading hypothesis, capable of explaining the successful behavior of the chimpanzees. To the best of my knowledge, only Andrews (2015) has presented an objection along these lines. According to her, chimpanzees may have learned, during the familiarization trials, that the human competitors choose the bigger grape. Then, during the proper test, the chimps can use that experience to expect the competitors to go on choosing the large looking grape. Nevertheless, the chimpanzees will choose the other grape because "they can use their own past experience to set up their own goals but they need not know why they were able to solve the task" (cf. Andrews 2015, 148).

Now, despite what she claims, Andrews' hypothesis does not seem to be a mere behavior-reading one. What chimpanzees have learned during familiarization trials is that human competitors usually choose the bigger grape. But, in the actual tests, the situation is relevantly different from those trials, because the chimps know, based on their previous experience, that the grape that looks bigger is in fact the smaller one. So they cannot merely project what they have learned during familiarization trials to the test situation (i.e., that there is a correlation between the real size of the grapes and the behavior of the competitors). What they need to do is to anticipate that the competitor will choose the grape that looks bigger, while integrating that information with their own knowledge about the real size of the grape in order to make their own better-informed choice. But this already implies being

capable of drawing the appearance/reality distinction and of understanding that their competitors will choose grapes based on how they appear to them and not on their real size. Something that, as Lurz and Krachun repeatedly point out, mere behavior readers cannot do. Whether chimpanzees can also understand that drawing such distinctions and making those attributions is why they are able to solve the task is not relevant here. The mere fact that they can draw the appearance/reality distinction and project it to others suffices to indicate that chimps are (some kind of) mind-readers. Of course, someone else may come up with a better behavior-reading hypothesis but, until that happens, we have pretty good reasons to think that Lurz and Krachun's experiment does solve the logical problem and is capable of providing solid evidence of chimpanzees' mindreading.

Moreover, as Lurz et al. have argued recently, there are reasons to be optimistic about chimpanzees being capable of succeeding at their test (Lurz et al. 2014). As they point out, we have evidence that chimps have many of the component skills involved in such a task. They are capable, for example, of engaging in competitive tasks over food with humans as well as other conspecifics (Hare and Tomasello 2004) and they can anticipate a competitor's action, adjusting their behavior accordingly (Hare et al. 2006; Schmelz et al. 2013). Finally, we also have some evidence that chimpanzees succeed in experiments designed to test whether they can distinguish appearance from reality (Krachun et al. 2009).<sup>2</sup> Clearly, the authors admit, none of this guarantees that chimpanzees will be also capable of using all these skills simultaneously in the kind of complex task designed by Lurz and Krachun (2011). This is something that will have to be tested. Nonetheless, it does show that it is not too far-fetched to think that this protocol can be implemented and, even more, that it can yield positive results.

## 2. Varieties of PPs and kinds of mindreading

### 2.1 Attributing pure PPs

At this point, one might ask: why should we think that Lurz and Krachun's new protocol does test whether chimpanzees are capable of internal goal at-

<sup>2</sup> It might be argued that chimpanzees' poor performance on false belief tasks (Hare et al. 2001; Kaminski et al. 2008) gives us reasons to doubt that they will be able to attribute false perceptual states to others, as this task requires. Against that objection, Lurz et al. (2014) respond that attributing false perceptions to others should be easier for chimps than attributing false beliefs. Succeeding on traditional false belief tasks requires representing the inferences that lead others from past or current evidential information to their false beliefs. Succeeding in false perception tasks, on the contrary, does not require representing the others' inferences, since their false perceptual states are caused by currently viewing an illusory object.

tribution? Doesn't it rather test their capacity to attribute perceptual states to others? The answer to such questions lies in Lurz and Krachun's suggestion that what chimpanzees would be attributing to their competitors are not traditional perceptual states but pushmi-pullyu representations (PPs). Millikan (1996; 2004b; 2004a) portrays PPs as the most simple and primitive intentional representations that a creature may have; representations whose distinctive feature is that they fulfill, at the same time, two inseparable functions: (i) a descriptive function to provide perceptual information about the external world ('this is a large grape'); and (ii) a directive function to indicate what the organism should do in such a situation ('grab it and eat it'). As a result, at least the simplest PPs—pure PPs—affect our actions directly, triggering a behavioral response in an automatic and inflexible way. Thus, if chimpanzees attribute PPs to others, they are simultaneously attributing some sort of perceptual content (i) and some kind of internal goal (ii).<sup>3</sup>

Lurz and Krachun add to these requirements that mindreading chimpanzees must be capable of understanding that the descriptive face of the PP may differ from reality. And, as we have seen before, this is something that a mere behavior-reader will be incapable of anticipating. It is important to notice, though, that even if, according to Millikan's theory, the descriptive content of a PP can be incorrect when things appear perceptually to its owner in a way that differs from how they really are, it does not follow from this that an animal which attributes PPs to others should be capable of grasping or appreciating that fact. The last one is an additional requirement added by Lurz and Krachun in order to be able to solve Povinelli's problem.<sup>4</sup>

<sup>3</sup> It seems to me that Lurz and Krachun are only offering here a plausible interpretation of the (hypothetical) positive results of their experiment. But their protocol does not conclusively favor the idea that chimpanzees are attributing pushmi-pullyus over other kinds of mental states like, for example, perceptual states. In section 3, I offer some general reasons in favor of the former interpretation. Briefly, I argue that the different varieties of "pushmi-pullyu mindreading" which I will examine in this paper are all psychologically simpler than full-blown mindreading and, as such, can provide us with valuable tools to account for the different capacities (and limitations) of infants and non-human primates in the domain of social cognition. If all this is sound, I think it can provide some additional support to Lurz and Krachun's suggestion.

<sup>4</sup> I will not question this requirement, at least in the case of animal mindreading. On the contrary, I will assume here that it must also apply to the more demanding varieties of animal mindreading which I will delineate below. To be very clear: from now on I will be proposing different varieties of mindreading, all of which require that mindreading chimpanzees, which attribute some complex varieties of PPs to others, must also be capable of distinguishing how things appear to others from how they really are. This is not a gratuitous or superficial decision, since two reasons justify it. Like Lurz and Krachun, I take the logical problem seriously and I think that adding this requirement to my proposal gives me a good way of overcoming it. Besides that, I think that it is also heuristically interesting, at least initially, to defend a model of chimpanzee mindreading, according to which



Clearly, the ability to attribute PPs is very different from the capacity to attribute standard psychological goals. Internal goals are usually thought to be directive psychological states. They refer to what their owner purports to do and can only guide the behavior of an agent when they are combined, through a process of practical inference, with descriptive mental states about how the world in fact is (Millikan 2004b). What follows from this is that attributing standard prescriptive internal goals should necessarily come with the additional capacity to attribute descriptive mental states, like perceptions and beliefs, and with some talent to represent—or somehow track—the inferential relations between these two varieties of mental states. Otherwise, our mind-reader would lack the means to make adequate predictions about how the others are going to behave.

Attributing pure PPs, however, is a less demanding cognitive activity, because it only requires the capacity to form a single representation that indicates at once how the agent purports to act based on how it perceives the environment. Still, such simplicity comes with a cost: the mind-reader who attributes PPs will always predict the same behavioral response from the agent: the one that figures in their prescriptive contents.

Nevertheless, in her work, Millikan distinguishes other more complex varieties of PPs which allow different non-human species to behave in more flexible ways (Millikan 1996; 2004b; 2004a). But, then, it seems legitimate to ask: what would happen if we tried to differentiate, analogously, between more or less sophisticated mind-readers in virtue of the sorts of PPs that they could attribute to others? What would attributing complex PPs consist in? And, in which ways would such capacities increase the flexibility and predictive powers of chimpanzee mind-readers? In the next section, I will try to offer an answer to these questions.

chimps would be very sophisticated in one sense—since they should be capable of drawing the appearance/reality distinction and of projecting it to others—and rather primitive in a different sense—since they will not need to represent the inferential links between the mental states of others (see section 3.1). One virtue of such a model, I think, is that it clearly distinguishes, and compels us to look for two different kinds of evidence: on the one hand, evidence that they are capable of detecting how things appear to others and contrasting this with how they really are (which indicates that they are not mere behavior-readers) and, on the other hand, evidence that chimps attribute different sorts of two-faced PPs (rather than traditional propositional attitudes). That being said, I admit that if we eventually find out that chimps do not pass Lurz and Krachun's test, we may choose to develop a model of chimpanzee mindreading that, while not demanding the capacity to distinguish appearances from reality, would still consist in the attribution of different kinds of PPs. With the appropriate modifications, much of what I propose here could still be useful for such a task. But, if that were the case, we would still need to find a good way to overcome or dismiss the logical problem.

## 2.2 Attributing complex pushmi-pullyu representations

Without pretending to be exhaustive, it is possible to distinguish at least two ways in which pushmi-pullyu representations became more complex. One important modification takes place, according to Millikan (2004b), when an animal not only acts following the directive face of a PP, but also modifies its responses according to (some of) its internal motivational states. Thus, an animal may perceive a prey nearby, but it will only follow the directive ‘go after the prey’, issued by its PP, when it is hungry. If, on the contrary, the animal feels satiated, it will not respond to the PP indicating the presence of the prey. These are not pure PPs, because there is not an imperative content compelling the animal to act no matter what, but rather a “potentially directive face” which will guide its behavior in a certain way if some additional psychological conditions are fulfilled. Following Duhau (2006), we can call them ‘pushmi-possible-pullyu’ representations (PPPs). Despite their greater complexity, PPPs are still relatively inflexible in the following sense: each of them only indicates one possibility of action. An animal may not act as indicated by its PPP when it lacks the adequate motivation, but it is not possible for this animal to represent the state of affairs which figures in the descriptive content of the PPP and still behave in a different way from the one prescribed by it.

However, what would happen with a mind-reader M capable of attributing PPPs? Firstly, M would have to be able not only to attribute perceptual-directive representations to others, but also to detect or track those motivational states that would lead them to act as the PPP indicates (or not). How can M accomplish such a thing? One option would be that M could explicitly represent the motivational states of another creature C, like hunger or thirst. There is, nevertheless, a second option: that M merely detects some behavioral and/or gestural signs reliably correlated with those states. Thus, for example, instead of perceiving C’s hunger, M may perceive its promptness to turn to the food, the way in which it leans its body, etc., as signs that indicate that C will try to fetch it.

Now, what would it take to establish whether an animal attributes PPPs to others? Roughly, I think we would need to have experimental protocols that allow us to test two things:

- (i) If a mind-reader M is capable of attributing to another creature C a pushmi-pullyu representation of an object O with a content in which O is represented as being different from how M knows that O really is. Testing this requires finding out if M can correctly predict the behavior that C would perform if C had an inadequate or illusory pushmi-pullyu.

- (ii) If M is capable of predicting the response of C to an inadequate or illusory pushmi-pullyu (C will choose the grape that looks big) only when M also detects that C is in the adequate motivational state (C is hungry), and of making a different prediction for cases in which C is not in such state (C is not hungry).

Now, we already know how to assess if the predictive capacities described in (i) are present, because this is exactly what Lurz and Krachun's original test was for. What we need is to find a way to test the capacities described in (ii). And this is something that requires designing a protocol which allows us to test whether M can attribute a PPP to C not only in virtue of how an object O appears to C but also in virtue of C's current motivational states.<sup>5</sup>

Let us turn now to another variety of complex pushmi-pullyu. The second important modification in the representational capacities of animals arises when they become capable of perceiving several different *affordances*, or possibilities of action, in the same object (Millikan 2004b). In such animals, which affordance they effectively follow will vary depending on their

<sup>5</sup> This is a rough sketch of what I have in mind. Assuming that chimpanzees have already succeeded at Lurz and Krachun's experiment and that they attribute PPs, we could run an experiment, using a violation of expectation method, to assess whether they also attribute PPPs. In the familiarization phase, chimpanzees will be shown a video of another chimp C being allowed by an experimenter to choose between a big grape, a small grape and an attractive toy. In the video, C chooses the big grape and the experimenter takes it away before C can eat it (this should allow the chimpanzees to learn that C is motivated to eat grapes and that C prefers the big grape rather than the small grape or the attractive toy).

Next, chimpanzees will be divided in two groups. *Group I* will watch a video of chimp C being fed with grapes which have been seasoned with an invisible and unpleasant flavor so that, after trying one of those, C rejects to continue eating (*bad flavor condition*). Then, C goes away from the room and the experimenter puts a big grape and a small one in size-distorting jars next to the attractive toy. In turn, *Group II* will watch a video in which the other chimp C conspicuously tries to obtain a grape in vain (*hunger condition*). Then C goes outside and the experimenter puts a big grape and a small grape in size-distorting jars next to the attractive toy. In both conditions, when C returns to the room, it is given to choose between these three things. Then the chimps will be shown three videos: in video 1 C chooses the toy; in video 2 C opts for the grape that looks smaller and eats it; in video 3 C chooses the grape that looks bigger and eats it.

If chimps were capable of attributing PPPs to others, the ones in *Group I* should look longer at C, in the bad flavor condition, when it chooses the big or the small grape instead of the toy (videos 2 and 3), while chimpanzees in the second group should look longer at hungry C when it chooses the toy (video 1) or when it chooses the grape that looks smaller over the one which looks bigger (video 2). If, on the contrary, chimpanzees were only PP mind-readers that rigidly linked certain perceptions to the same behavioral prescriptions, their expectations about the behavior of C should be the same ones for the two groups. I think that, with due polishing, this is a protocol that could help us distinguish mind-reading-chimpanzees which only attribute PPs from chimpanzees capable of attributing PPPs.

context, motivations and necessities at a given moment. Thus, chimpanzees can perceive and use a leaf as a sponge to catch water (Goodall 1968; McGrew 2004), as a napkin to wipe off bodily fluids or plant juices (Goodall 1968; McGrew 2004), or as a cushion to seat more comfortably while doing terrestrial tasks (Hirata et al. 1998; McGrew 2004).<sup>6</sup> I would like to suggest that these animals have what can be called ‘pushmi-multiple-pullyu representations’ (PMPs) which allow them to represent the same object as affording different possibilities of action in different situations, and consequently compel them to act in different ways toward that object.

Imagine now a mind-reader *M* capable of attributing to another creature *C* the perception of different affordances, in the same object, under different circumstances. Clearly, the cognitive capacities involved in such feat are more demanding than the ones required to attribute the other two varieties of pushmi-pullyu representations examined before. When all *M* attributes to *C* are pure PPs, or even PPPs, each of these representations indicates only one way in which *C* must act. As a consequence, *M* will always predict the same behavioral response from *C*: the one indicated by the pushmi-pullyu representation which *M* has attributed to *C*. At most, if *M* is attributing a PPP to *C*, then *M* will be able to abstain from making the relevant behavioral prediction when *C* is not in the adequate motivational state. In contrast, the mind-reader *M* capable of attributing to *C* the perception of different affordances in the same object *O* will have to determine which of those affordances is currently guiding *C*’s behavior in order to make a successful behavioral prediction. But, how can *M* do that? It seems reasonable to think

<sup>6</sup> One may wonder at this point: if *M* is attributing to *C* a PMP with a descriptive content in which an object is seen as affording a specific possibility of action, how does this differ from attributing traditional propositional contents? The answer depends on how one understands the notion of propositional content. Sometimes, philosophers assume that a mental state has propositional content if and only if it represents the world as being in a certain way (Searle 1983). It certainly seems that, in this minimal sense, the directive content of a PMP would be propositional. But it is also possible to think, in a more demanding way, that propositional contents must also have a sentence-like structure (Cf. Millikan 2004b, 92–93). This means, amongst other things, that propositional contents must be constituted by representational units (concepts) that can be freely re-combined to generate new thoughts. But the directive contents of PMPs need not be like that. A chimpanzee may come to have a limited set of contents like: ‘this leaf is usable as a hat’ and ‘this stone is usable as a hammer’ which represent different states of affairs. But these contents may not admit being decomposed and recombined in the same ways that propositional contents can. The chimp may not be able, for example, to have a thought like ‘this leaf is a hammer’ like we can do. This line of thinking allow us both to think that PPs may represent objects as being in certain ways, while at the same time denying that they collapse with propositional contents. (For a different proposal on how to understand animal contents, which is propositional in the non-demanding sense, and non-propositional in the more demanding one, see Bermúdez 2003).

that the affordances that guide C's action will depend on C's necessities, its motivational states, and its present context. If C is playing, it will probably climb a tree to make stunts; if C is scared because there is a predator nearby, it may use the tree as a hiding place, etc. Then, to determine which affordance is guiding C's behavior, M will need to attend not only to the object perceived by C, but also to C's motivational states and to the practical context in which C's actions are taking place.

Anyway, mindreading animals capable of attributing PMPs will have a more versatile capacity to predict the behaviors of others, since they will be able to:

- (i) Anticipate that, in certain circumstances or under certain motivational states, C will act in a way  $A_1$  towards object O.
- (ii) Refrain from predicting that C will act in the way  $A_1$  towards O when C is not in the appropriate circumstances or lacks the adequate motivational states.
- (iii) Anticipate that, in a different context and with different motivations, C will act in front of O in the way  $A_2$ .<sup>7</sup>

Finally, I think it is also possible to empirically distinguish between a mindreading animal which only attributes pure PPs or PPPs to others, from an animal capable of attributing to others the capacity to perceive multiple affordances in a given object. To test this difference, the following will be required: (i) putting M and C in front of the same object O while varying relevant contextual features and/or the motivational states of C that, presumably, will change the affordances that C perceives in O; and (ii) checking if M is capable of predicting the changes in C's responses to those different affordances. This is a complex task, but I think it is also a viable one, mainly because the theoretical rationale behind Lurz and Krachun's original experiment still holds in this case. After all, no matter which one of O's multiple affordances is guiding C's behavior, it is always possible that C perceives O in a way that differs from M's knowledge of how O really is. In such a case, what will be relevant to successfully predict C's internal goals and C's subsequent behaviors will be how things appear to it and not how they really are. And, as it happened in other cases, only mind-readers capable of grasping such a difference will be able to adequately predict C's behavior and to adjust their own behavior accordingly.

<sup>7</sup> It is interesting to compare these sets of abilities to the predictive powers of the mind-reader which can only attribute pure PPs to others, and to the mind-reader which can, at most, attribute PPPs to them. The first one is only capable of (i) and the second one of (i) and (ii).

### 3. Mindreading as an attribution of pushmi-pullyu representations: Further issues

Up to now, I have provided a rudimentary characterization of two new varieties of mindreading that non-human animals might have. Basically, they both consist in the capacity of attributing some complex types of pushmi-pullyu representations to others. Everything I have outlined so far is rather speculative and it seems that, until we have relevant empirical evidence on its favor, these models will not have earned its keep.<sup>8</sup> But, even lacking such evidence, there are two further issues which I would like to discuss in this section: what distinguishes the varieties of mindreading proposed so far from full-blown mindreading, and whether we can legitimately extend the present proposal to research on the social cognition abilities of non-human infants.

#### 3.1 Attributing complex pushmi-pullyu representations vs. full-blown mindreading

Firstly, I would like to focus on whether the more sophisticated varieties of pushmi-pullyu mindreading which I have proposed are genuinely different from, and simpler than, full-blown mindreading. According to Millikan (2004b, 174), as a representational system becomes more sophisticated, pushmi-pullyu representations progressively come apart, and descriptive representations detach and become independent from directive ones. Following this idea, we can use the notion of ‘detached mental states’ to refer to those intentional mental states which have a directive or a descriptive content, but not both at the same time. Traditional propositional attitudes would be our paradigmatic example of this kind of mental states.

Making use of this notion, it can be claimed that full-blown human mind-readers predict and explain the behavior of other creatures by attributing detached mental states to them. Some of these mental states have a descriptive function and others a directive one, and they can figure in different varieties of practical reasoning which guide the behavior of the interpreted agents. Presumably, then, the mind-reader who attributes this kind of men-

<sup>8</sup> It is interesting to point out, anyway, that if we assume that chimpanzees are mind-readers—ignoring the logical problem for a while—, there is some evidence hinting that they must be capable of attributing at least one variety of complex pushmi-pullyus: PPs. According to Call et al. (2004) chimpanzees respond differently (gesturing more and leaving the testing situation earlier) when an unwilling experimenter fails to give them a grape than when the same experimenter is unable to do so. This experiment could be interpreted, I suggest, as showing that chimps are attributing to both experimenters a PPP with a descriptive content ‘this is a grape’ and a potentially prescriptive content ‘give it to the chimp’ that will only be followed if the experimenter is motivated to do so.

tal states must represent those inferences—or somehow track them—in order to be able to produce adequate predictions of the others' behaviors.

Clearly, this is different from what the basic mind-reader M, who merely attributes pure PPs, does. Firstly, M can predict what another creature C will do merely by attributing to C *one* mental state (instead of multiple ones) which rigidly combines a perceptual content with a prescriptive one. Secondly, the behavioral prediction will directly follow from that single attribution without requiring from M any capacity to track or represent any inferential processes in C. Even a creature completely incapable of representing the inferential processes of others could be a mind-reader in this minimal sense. At least in this regard, it seems that attributing pure PPs is psychologically (or cognitively) simpler than full-blown mindreading.

Nevertheless, at this point, one might wonder whether the attribution of complex PPs is really different, in a relevant sense, from traditional full-blown mindreading. I will try to clarify why such worries arise for each of the attributions of complex PPs that I have distinguished, and how they can be removed.

Firstly, I would like to examine the case of the mind-reader who attributes PPPs. As we have seen, this kind of mind-reader attributes a descriptive/potentially directive mental state to another creature C which indicates, at the same time, what it is out there in the world and what C should do about it, if it happens to be appropriately motivated. To accomplish such a task, this mind-reader will have to track two different kinds of mental states: (a) the PPP and (b) the motivational state that “activates” the imperative face of the PPP. Presumably, this would allow mind-reader M to attribute to C, who faces an object O, a PPP that simultaneously indicates the presence of O and how to act towards it. But, it will also allow M to distinguish two kinds of situations: (i) one in which C, being properly motivated, will act as the prescriptive side of the PPP commands; from (ii) a different situation in which, lacking the necessary motivation, C will not act as the PPP prescribes. As a consequence, our mind-reader will be able to predict that, in situation (i), C will act in the way prescribed by the PPP, while, in situation (ii), C will not act in such a way. Thus, imagine that M is trying to anticipate the behavior of a rival C who is in front of a piece of food. M will be able to attribute to C the PPP: ‘food there / (if motivated) grab it and eat it!’ Now, M will be also capable of detecting whether C is motivated—i.e., hungry—(situation (i)), or not (situation (ii)), and to predict that in (i) C will grab the food and eat it, while in (ii) it will leave the food untouched.

Yet, at this point one may ask: why shouldn't we characterize M's capacity to discriminate situation (i) and (ii) in terms of M attributing to C in both situations the same descriptive mental state ('food there') but two different

desires (the desire to eat food in (i) and the desire not to eat food in (ii))? I think that this is a viable alternative and that a mind-reader who attributes beliefs and desires should be capable of distinguishing situation (i) and (ii), as well as being capable of making adequate predictions in both cases. But what I would like to defend is that the abilities displayed by such a mind-reader are substantially different from the ones displayed by the mind-reader M, who merely attributes PPPs. Secondly, I will argue that what M is doing is psychologically less demanding from what a mind-reader which attributes detached beliefs and desires would need to do in order to achieve similar results. Finally, I think that such a gain in simplicity is, at least in principle, a good reason to favor this way of characterizing the socio-cognitive abilities of creatures like non-human primates and young human infants that, for all we know, are cognitively less sophisticated than normal human adults.

Let us begin by clarifying where the difference lies between attributing detached beliefs and desires, on the one hand, and attributing PPPs, on the other. The first thing to point out is that, unlike detached mental states, a PPP inexorably constrains how the interpreted creature C behaves when facing an object O. In contrast, a creature which is, for example, in a purely descriptive perceptual state can give multiple different responses depending on how the aforementioned perception combines with her or his other mental states.

If we extrapolate this to mindreading, the following consequence follows. As stated above, traditional mind-readers will have to attribute at least two different kinds of mental states—i.e., beliefs and desires—with two different contents clearly specified, in order to predict how C is going to act. Even more, in order to make adequate predictions, they will have to attribute to C different desires, with different contents in situations (i) and (ii). On the contrary, a mind-reader M, which attributes PPPs, can make the correct behavioral predictions in both situations by attributing to C one and the same PPP, and then detecting if C is in a motivational state that can “activate” the prescriptive side of the PPP, causing C to act as the PPP commands (or not). The interesting thing is that M does not have to represent the intentional content of such motivational state, since the content that suffices to explain what C will do is already rigidly contained in the directive face of the PPP. Thus, M does not need to attribute, like a full-blown mind-reader, two different desires, each with its own content, in situation (i) and (ii). M can predict C’s behavior merely by attributing one isolated intentional mental state—the PPP—plus a “contentless” motivational state that merely indicates whether C’s is “on” or “off” to act as the PPP commands.

Imagine, for example, that M attributes the PPP ‘food there, (if motivated) grab it!’ to C. Then, M may still need to find out whether C is motivated to act as the PPP commands, but once that is done, all the information



that M needs to predict about how C will act—grabbing the food—is already contained in the PPP. M does not need to attribute to C an additional desire to grab the food in question.

Even more, as suggested above, M may be “tracking” C’s motivational states merely by detecting some behavioral and/or gestural signs reliably correlated with those states. Thus, for example, instead of perceiving C’s hunger, M may perceive its promptness to turn to the food, the way in which it leans its body, etc., as signs that indicate that C will try to fetch it.<sup>9</sup>

But why, one may insist, should we consider these differences between attributing PPPs and full-blown mindreading to be relevant from a psychological point of view? At least for two reasons. Once again, we can begin by pointing out that attributing PPPs is not only a different, but also a psychologically simpler task. Why is this so? Mainly because, unlike full-blown mind-readers, the mind-reader M which attributes PPPs to others will not need to represent any kind of inferential link between different decoupled mental states (e.g., beliefs and desires) with their respective contents. M will only have to posit a non-inferential (causal) relationship between a motivational state, devoid of any content, and the PPP attributed to C, in which the former “activates” the directive content of the latter. In other words: all M needs to do is to: (i) attribute a PPP representation to C and (ii) detect if that PPP is “on” or “off” or, in other terms, detect whether, at that moment, the PPP will or will not guide C’s behavior. This is all that ‘tracking C’s motivational states’ amounts to. But, reasoning about the causal link between a motivational state devoid of content and a representational PPP is not the same as being somehow able to represent the reasoning of other creatures, or the inferential links between the contents of their different representational mental states.

Furthermore, it can be argued that representing and understanding how different mental contents of C are inferentially linked in order to guide C’s behavior, as the full-blown mind-reader has to do, is a psychologically taxing

<sup>9</sup> It may be objected that if M tracks C’s motivational states by detecting some of C’s observable gestures or behaviors, then M may end up being a behavior-reader which only uses behavioral rules to predict C’s behavior. However, this is not strictly true, since in order to anticipate how C is going to act M is not relying directly on a rule linking some signs of motivation in C with C’s behavior. Rather, M is following a “hybrid strategy” which combines the behavioral signs of C’s motivational state with the attribution of one genuine representational mental state: a PPP. Furthermore, this is not a gratuitous stipulation, since we have a way of providing empirical evidence that this is the case. As stated in footnote 4, the present proposal requires that a mind-reading chimpanzee should be capable of attributing PPs with an illusory descriptive face. Thus it will always be possible to design an experimental protocol, following the general lines proposed by Lurz and Krachun (as suggested in footnote 5), to empirically test whether M indeed attributes PPPs instead of being a mere behavior-reader.

task for several reasons. It requires not only representing and “holding in mind” different descriptive and directive mental states, but also representing how they “fit” together, working out which are the semantic and/or logical links between their respective contents, and how they may jointly determine a particular action. At least *prima facie*, this seems to be tough cognitive work.<sup>10</sup>

Let us turn next to the attributions of pushmi-multiple-pullyu (PMPs) representations. In this case, our mind-reader M is capable of attributing to C the perception of different affordances in an object O. More specifically, M will be capable of attributing to C, in different contexts, different PMPs, each of them composed by a complex descriptive face in which the same object is represented as affording one of many different possibilities of action and, consequently, is associated with one of many potential imperative contents. Now, one may wonder: how is this different from being capable of attributing to C different descriptive mental states about O’s affordances and different directive mental states?

Even though it may seem harder to draw a clear distinction in this case, I think that there is still a relevant difference between the mind-reader M which attributes PMPs representations and a full-blown mind-reader. Firstly, even though M will be able to attribute to C the perception of a range of different affordances on the same object O, and a range of different goals towards O, the representation of each affordance of O will be rigidly connected to one and only one corresponding goal towards O. This is so, because the prescriptive content attributed to C will issue a command to act precisely as C perceives that O affords. Thus, if C has a PMP with the content ‘eatable food there!’, the prescriptive content will be: ‘go there and eat the food!’ And this should allow M to immediately predict how C is going to act: as the prescriptive face of the PMP commands.

Clearly, then, these are not detached descriptive and directive representations that M may combine at will. Rather, they should be thought as different pre-fixed pairs of two-faced representations about the same object O that M can attribute to C. Thus, what M needs to do is to work out which pair of ‘perception of O’s affordance / goal towards O’ should be attributed to C, given C’s context and motivations. But, once that is done, M will be attributing both contents simultaneously and this is what will allow M to predict directly, without further cognitive steps, how C is going to act. Now, all this is

<sup>10</sup> Following a similar line of reasoning, Butterfill and Apperly also acknowledge that one of the reasons why full-blown mindreading is a cognitively demanding task, which consumes working memory and other scarce cognitive resources, is that it involves attributing propositional attitudes which interact with each other in “uncodifiable complex ways.” Cf. (Butterfill and Apperly 2013, 610).

clearly different, and arguably simpler, than what a full-blown mind-reader will have to do. Unlike the latter, M will not need to track the inferential links between the contents of different detached mental states in order to predict C's behavior.

The contrast is even more acute if we accept that full-blown mind-readers must possess a full-range of mental concepts and must be capable of representing virtually any type of mental content, as well as all sorts of relations between mental states, environment and behavior (Hutto et al. 2011; Apperly and Butterfill 2009). All this makes full-blown mindreading a sophisticated and flexible ability that can be used in a general way, in many different contexts. But, it also dramatically increases the cognitive complexity of having to select, from a very wide and rich collection, those specific mental states that can be appropriately combined in inferences to lead each agent to act as she or he does in a particular situation. In contrast, the attributor of PMPs is only capable of attributing to other creatures the detection of a limited and already fixed combination of two-faced representations about what is out there in the world, what possibilities of actions it affords, and how one should act towards it.

Furthermore, even if we turn our attention to other minimal varieties of mindreading on offer, a crucial difference with the present proposal remains. We may, for example, think of models of mindreading that impose limitations on the kinds of mental attitudes that the mind-reader is capable of attributing. One clear case are those proposals according to which minimal mind-readers—like non-human primates or human infants—are only capable of attributing some “basic” mental states, like perceptions and goals, to others. Or we may think about proposals according to which basic mind-readers are capable of tracking only some mental contents of other creatures and only some of its inferential links, abandoning the more demanding varieties of inferential holism.<sup>11</sup> In any case, the complexity of the task which mind-readers face is reduced, because the variety and amount of mental states, and of inferential links, that must be represented is severely restricted. Nevertheless, as long as what these mind-readers attribute are detached mental states, they will still need to select which set of mental states to attribute to another creature C, and to track the inferential links between them in order to predict C's behaviors. This is a demand that a pushmi-pullyu mind-reader will not need to meet. And this is the very specific sense

<sup>11</sup> According to the most radical varieties of holism, in order to attribute a mental state, like a belief, to an agent, it is necessary to situate it in a maximally rich and complex web of other mental states logically related to it. There are compelling arguments in favor of abandoning such a position, especially when we are interested in animal minds (Glock 2000; Finkelstein 2007).

in which the attribution of PPs to others purports to be even simpler than other minimalist rivals.

### 3.2 Mindreading and attribution of pushmi-pullyus in human infants

Another interesting issue to consider is whether the different varieties of pushmi-pullyu mindreading presented so far could also be extended to human infants.<sup>12</sup> It is frequently claimed that human infants have mindreading abilities that are more basic and primitive than those of older children and adults, and many alternative ways of understanding those early capacities have been proposed. But, are there reasons to think that these simpler capacities could also consist in attributing pushmi-pullyus to others? Is this at least a viable hypothesis? Although this issue deserves a much more careful and detailed examination, let me offer a few tentative comments on why I think the varieties of mindreading which I have presented may be fruitfully extended to human infants.<sup>13</sup>

Let us focus, first, on some other experiments on infant goal attribution which may seem hard to fit in the present account. According to Michael and Christensen (2016), there is a large body of research providing evidence that infants attribute goals to others by flexibly drawing on different kinds of contextual information, including information about the agent's previous activities, preferences and other psychological states. In one of these studies, Woodward (1998; 2005) found out that infants take into account the prior actions of agents when attributing goals to them. In her study, 5 to 12-month-old infants were habituated to an event in which an agent, facing an array of two toys, A and B, systematically reached for toy A. During the proper test, the positions of the toys were switched and infants saw two alternated events: on some occasions, the agent preserved the original path, but

<sup>12</sup> In a recent paper, Michael (2015) draws on Millikan's notion of "unicepts" to defend that infants maybe be attributing beliefs to others by as early as 6 months. My proposal, instead, is that what infants are attributing to others are not beliefs but pushmi-pullyu representations. I am, then, making use of a different notion from Millikan's work, and applying it to explore a different way of understanding infant mindreading. Anyway, although I will not be able to do that here, it would be interesting to explore the possibility that infants (and chimpanzees) may be using unicepts to attribute different kinds of PPs.

<sup>13</sup> Such an extension would bring the additional benefit of allowing us to view the mindreading abilities of non-human primates as continuous—at least to a certain extent—with mindreading in humans. It would be possible to argue, for example, that our closest evolutionary relatives evolved some basic mindreading abilities and mechanisms which are similar to those deployed by human infants (and may be also similar to the ones that human adults exhibit in their more immediate and non-reflexive interactions with others). This does not imply denying, of course, that humans may also have additional mechanisms, operating from early infancy onwards, which are not present in other species (Barrett 2015).

ended up reaching toy B while, in others, the path was changed but the agent kept reaching for toy A. Now, infants looked longer when the agent reached for toy B. This suggests that they interpreted that the agent had the goal of reaching a specific object, rather than a specific location. But it also suggests, according to Michael and Christensen, that in attributing this specific goal, infants were taking into account the previous behavior of the agent during the habituation phase. More specifically, it can be argued that infants saw the selection of toy A over B in the habituation phase as revealing the preference of the agent for toy A over toy B, and that they expected this preference to guide the agent's action in the test situation.

Later studies suggest, furthermore, that infants may be sensible to other psychological states of agents, besides their preferences, when they are attributing goals to them. Luo and Baillargeon (2007), for example, ran a test with a habituation phase that included two conditions. In the "visible object" condition, the agent could see that there was another object B, but systematically chose object A. In the "hidden object" condition, there was a second object present but the agent could not see it. In the proper test, the agent selected one of the two objects. When 12.5-month-old infants had experienced the "visible object" condition, they expected the agent to keep choosing the same toy which she or he had repeatedly chosen in the habituation phase, but they did not show such an expectation when they had previously experienced the "hidden object" condition.

According to one way of interpreting these tests' results, then, infants would be integrating information about different psychological states of agents—like their preferences, their perceptual states and their goals—in order to predict how they are going to act. In other words, infants would be attributing detached mental states, and tracking their holistic inferential links, instead of merely attributing PPs. Now, I do not think this is the only plausible interpretation of the evidence. We could think, instead, that infants do attribute PPs, while also being sensitive to the previous actions of the agents (and not to their mental states). Thus, during the habituation phase of Woodward's experiments, infants could be attributing to the agent two PPPs with the contents 'there's an object A (if motivated) / grab it' and 'there's an object B (if motivated) / grab it.' Now, what our infants learn from experience is that the agent always chooses object A. Based on that observation, they come to expect that, when the agent has two PPs about two objects A and B—that is, when the agent has, at the same time, two perceptual descriptive contents about A and B and two potential goals (to grab one or the

other)—the PPP that gets “activated” (the one that the agent will actually act on) is the PPP about object A.<sup>14</sup>

We can also explain Luo and Baillargeon’s results in terms of the attribution of PPPs. According to this alternative interpretation, in the “visible object” condition of the habituation phase, when objects A and B are both visible to the agent, infants would attribute two PPPs to her or him, one about each object, and will come to learn by experience that this agent always acts according to what the PPP about object A indicates. In contrast, in the “hidden object” condition, where there is a second object B present but the agent cannot see it, the infant will only attribute one PPP to the agent (about object A). During the proper test, infants who had experienced the visible object condition will predict that the agent will reach for object A, because they would have previously learned that, when the agent has two contrasting PPPs, one about object A and the other about object B, the first one gets activated. On the contrary, infants who experienced the “hidden object” condition will not have attributed two PPPs to the agent in the habituation phase. During the test, they will attribute two PPPs for the first time (one for each object), but they will not have any previous experience about which of them is the one which usually guides the actions of the agent.

Turning to descriptive mental states, there is evidence that, from quite early on, infants have some understanding of the perceptual states of others; for example, they can establish what objects they can and cannot see (Sodian et al. 2007). Now, in these experiments, infants are interpreted as attributing perceptual states about some objects based on their expectations on how agents are going to act (or not) towards them. Thus, it seems *prima facie* plausible to offer the following alternative account for such data: what

<sup>14</sup> This interpretation can be questioned on the following grounds. Phillips, Wellman and Spelke (2002) ran an experiment to test whether infants recognized that a person who previously looked positively at an object would grasp that object rather than a different one. During the habituation phase, infants saw a woman, first looking and then picking up one of two stuffed animals (toy A). Then, infants were shown two types of tests. During one of them, the woman looked at toy B and then picked it up (consistent actions). During the other, she looked at toy A but picked up toy B (inconsistent actions). At 12 months (but not at 8 months) infants looked longer when the actions of looking and grasping were inconsistent than when they were consistent. Now, this may be taken as indicating that 8-month-olds do not attribute a PPP, with a descriptive and a prescriptive content, but only isolated goals disconnected from the states of visual attention of the agent. Another alternative, though, is to claim that these young infants may still be attributing PPPs, but on the basis of different cues than the ones used by older infants. Thus, initially, infants may attribute PPPs to agents about all those objects that they can see and that the agents can reach. Later on, 12-month-old infants will attribute a PPP about an object O to an agent A only when the latter has a direct line of gaze towards O, is attending to O, etc. (see my interpretation of Luo and Baillargeons’ test above). But the only thing that has changed here are the conditions of attribution of the PPP, not its descriptive/directive content.

infants are attributing to others in these experiments are pushmi-pullyu representations which indicate that there is an object nearby and, at the same time, prescribe a way of acting towards it.

Even more impressive is a range of experiments which has been interpreted as indicating that 12 to 18-month-old infants have some understanding of the false beliefs of others and take them into account when predicting their actions (Onishi and Baillargeon 2005; Buttelmann et al. 2009; Southgate et al. 2007). Now, while these studies suggest, at least, that infants are sensible to information which goes beyond what is perceived here and now, PPs are representations about presently perceived objects or states of affairs. Consequently, it does not seem appropriate to try to explain the results of the former in terms of the latter. Nevertheless, I think that such a task is still feasible if we are willing to make some further adjustments to our original proposal.

We can begin by distinguishing a new variety of pushmi-pullyu representations, that we may call PP\*s. The directive content of a PP\* is about an object *O* having certain properties  $P_1 \dots P_n$  iff: (i) agent *A*, the owner of the PP\*, has previously perceived, at time  $T_1$ , *O* having  $P_1 \dots P_n$ ; and (ii) *A* has not perceived *O* having properties incompatible with  $P_1 \dots P_n$  since  $T_1$ . In a nutshell: PP\*s are like traditional PPs but they represent how agents take the world as being in the present, based on states of affairs that they have perceived in the past.

Having introduced PP\*s, we can now characterize another variety of mindreading which will consist in attributing PP\*s to others. A mindreading infant *M* will be capable of attributing to an agent *A* a PP\* about an object *O* having properties  $P_1 \dots P_n$  iff: (i) at time  $T_1$ , *M* attributed to *A* a PP with a descriptive content: ‘object *O* is  $P_1 \dots P_n$ ’ and a prescriptive content: ‘Do *A* to / with object *O*!’, (ii) since time  $T_1$ , *M* has not detected that *A* encountered once again *O* manifesting different properties from the ones perceived at  $T_1$ . Putting it bluntly, *M* will be attributing to *A* a representation, which was acquired in the past, about how things are now (and about what *A* should do about them).

Such a notion should suffice, I conjecture, to explain the main evidence available on false belief attribution in human infants. Let me give one brief example. In Onishi and Baillargeon (2005) classic experiments, infants see that an agent *A* places an object *O* in one location—a dark box—but *A* does not see that, later on, the object switches its location to a light box. If infants expect *A* to search for *O* on the prior location (presumably because they have attributed a false belief to her or him) they should look longer when *A* searches for the object on the light box. Which is, indeed, what happened with 15-month-old infants.

Now, imagine that, instead of attributing false beliefs to agent A, infants detected that A perceived the object being put in the dark box at time  $T_1$  and attributed to A a PPP representation like: ‘Object O is in the dark box / (if motivated) go there and grab it.’<sup>15</sup> After that, the infants did not detect A perceiving object O at a different location. As a consequence, they attribute to A a PPP\* with a content ‘Object O is in the dark box / (if motivated) go there and grab it.’ Clearly then, they will expect the agent to go to the dark box and not to the light one and they will look longer if what happens is the latter.<sup>16</sup>

Of course, it is still an open question whether the present proposal would be able to adequately explain the main data that we have on the complex capacities of human infants in the domain of social cognition.<sup>17</sup> The brief considerations (over a few studies) that I have sketched here only attempt to establish that this is a promising alternative which deserves further attention.

Finally, there is another complex issue which I would like to discuss, at least in a preliminary way. As is well-known, there are already several alternative accounts which credit human infants and non-human animals with rudimentary mindreading capacities.<sup>18</sup> Hence, it seems legitimate to

<sup>15</sup> I am treating here the location of the object as the critical property of object O which appears in the content of the PP.

<sup>16</sup> On the contrary, the “true-belief” condition, in which A stays in the room and sees object O switching locations from the dark box to the light one, can be accounted for in the following way: at time  $T_1$ , the infant attributed to A the PPP: ‘object O is in the dark box / (if motivated) go there and grab it.’ Soon afterward, the infant perceived how A saw the object switching locations. As a consequence, the infant attributed to A a new PPP with the opposite content ‘Object O is in the light box / (if motivated) go to the dark box and grab it,’ which will give place, after A leaves the room and comes back, to the attribution of a new PPP\* with the same content of the last PPP. This is what explains why, in this condition, the infant looks longer to A when she/he searches for the object in the dark box.

<sup>17</sup> Am I requiring, like I did in the case of chimpanzees, that the infants in all these experiments should be capable of attributing illusory PPs to others? My main focus in this section has not been to give an answer to skeptics who claim that infants are mere behavior-readers. Rather, I have been interested in a more modest task: showing to those already convinced that we have good evidence on infant mindreading, that endowing infants with the capacity to attribute PPs to others can account for some of those studies. For this reason, I would like to remain provisionally neutral on whether we should also ask infant mind-readers to be able to attribute illusory PPs. Anyway, those who consider that a good answer to the skeptics about infant mindreading is needed can easily add the requirement that to qualify as mind-readers, infants should be capable of attributing illusory PPs to other’s, and design new protocols in order to gather relevant evidence on when infants can actually accomplish such a task (as suggested by Lurz and Krachun 2011).

<sup>18</sup> Thus, for example, some researchers claim that infants and animals are only capable of attributing to others some basic mental states (like intentions, desires, perceptions, etc.), while they are not capable of attributing more sophisticated ones (e.g., beliefs) (Call and



ask: are there good reasons to prefer the present proposal to its rivals or, once again, to consider it at least a worthwhile alternative to develop? Although this topic would require further development and discussion, I will present here one preliminary reason which favors the present account over some of its rivals.

To the best of my knowledge, all the alternative varieties of simple mind-reading on offer involve the attribution of detached mental states. In addition, at least some of them also explicitly require that simple mind-readers show some appreciation of the holistic inferential links between the different detached mental states of the others, as well as some understanding of how their behavior will vary as a result of the ways in which their mental states get inferentially combined.<sup>19</sup> Now, as Zawidzki (2011) argues, many current experiments on infant mindreading do not strictly test for such an appreciation of holism (not even for a moderate holism). They are not designed to assess, for example, whether infants are capable of tracking different ways in which the same mental states can cause different behaviors depending on which other mental states the agent has and how they are inferentially combined. The same goes, I think, for current studies on the mindreading abilities of non-human primates which, as far as I can see, do not test the capacity of animals to attribute many inferentially interrelated mental states to others, but only their abilities to attribute isolated mental states (having an internal goal, seeing an object or not seeing it, etc.).

But, then, if experiments on young infants' and non-human animals' mindreading abilities do not test for holism and extant evidence can be interpreted as suggesting that these creatures are only correlating some particular mental states with specific behaviors, why should we think that what these basic mind-readers are attributing are detached mental states with holistically and inferentially mediated causal influence on behavior? It seems more parsimonious to claim, instead, that they are attributing some kind of PPs to others, which allow them to predict the agents' behaviors without having to

Tomasello 2008; Wellman 2014). Other well-known alternative consists in differentiating two different mindreading systems (Apperly and Butterfill 2009). According to this last proposal, human infants and some non-human animals are capable of using the more primitive system, which allows them to have an implicit and fast—but also limited—understanding of beliefs, whereas from preschool onwards, children become capable of understanding beliefs in an explicit, deliberate, flexible and expansive way.

<sup>19</sup> Wellman (2014, 88–90, 179–184), for example, defends a progressive understanding of the interplay of several mental states which begins in infancy. See also Call and Tomasello (2008, 191). On the contrary, the “minimal theory of mind” defended by Apperly and Butterfill does not endorse such a requirement. The “registrations” that infants and animals would attribute to others, according to the authors, must have their effects on actions independently of each other and of any other mental state (see Apperly and Butterfill 2009, 963).

go through the laborious and cognitively demanding process of tracking the inferential links between their different mental states. This gain in simplicity seems to be, at least *prima facie*, a good reason to put some effort into the development of the present account.

To summarize: in the last decades, proposals that seek to clarify both the level and distinctive features of the mindreading abilities of non-human animals, without crediting them with full-blown mindreading, have flourished. This paper can be read as an effort to contribute to this debate by extending and complementing some ideas, originally presented in Lurz and Krachun (2011), in order to develop a new way of understanding the varieties of mindreading that some non-human primates, like chimpanzees, may have. To accomplish such a task, I have distinguished three different types of basic mindreading, each mainly consisting in the capacity to attribute some variety of pushmi-pullyu representations to others, and I have contrasted them with full-blown mindreading. Furthermore, I have suggested that pushmi-pullyu mindreading might be fruitfully extended to infant mindreading, and I have given one preliminary reason to favor it over some of its rivals. Much more should be done, both conceptually and empirically, to develop the present proposal and to evaluate its merits and limitations. Here, I have only tried to take some initial steps in that direction.

### Acknowledgments

I would like to thank Robert Lurz, Laura Duhau, Pamela Barone, Daniel Kalpokas, the reviewers for *Studia Philosophica Estonica* and the members of the GPyC for their very careful and insightful comments to earlier drafts of this paper.

### Bibliography

- Andrews, K. (2015). *The Animal Mind: An Introduction to the Philosophy of Animal Cognition*, Routledge, New York.
- Apperly, I. and Butterfill, S. (2009). Do humans have two systems to track beliefs and belief-like states?, *Psychological Review* **116**: 953–970.
- Barrett, C. (2015). *The Shape of Thought*, Oxford University Press, Oxford.
- Bermúdez, J. (2003). *Thinking Without Words*, Oxford University Press, Oxford.

- Buttelmann, D., Carpenter, M. and Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in active helping paradigm, *Cognition* **112**: 337–342.
- Butterfill, S. and Apperly, I. (2013). How to construct a minimal theory of mind, *Mind and Language* **28**: 606–637.
- Call, J. (2007). Past and present challenges in theory of mind research in nonhuman primates, *Progress in Brain Research* **164**: 341–353.
- Call, J., Hare, B., Carpenter, M. and Tomasello, M. (2004). ‘Unwilling’ versus ‘unable’: Chimpanzees’ understanding of human intentional action, *Developmental Science* **7**: 488–498.
- Call, J. and Jensen, K. (2006). Chimpanzees may recognize motives and goals, but may not reckon in them, *Novartis Foundation Symposium* **278**: 56–70.
- Call, J. and Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later, *Trends in Cognitive Science* **12**: 187–192.
- Duhau, L. (2006). *Thinking Conceptually*, PhD thesis, King’s College London.
- Finkelstein, D. (2007). Holism and animal minds, in A. Crary (ed.), *Wittgenstein and the Moral Life: Essays in Honor of Cora Diamond*, The MIT Press, Cambridge, MA, pp. 251–278.
- Flavell, J. (1992). Perspectives on perspective-taking, in H. Beilin and P. Puffall (eds), *Piaget’s Theory: Prospects and Possibilities*, Erlbaum, N.J. Hillsdale, pp. 107–139.
- Glock, H.-J. (2000). Animals, thoughts and concepts, *Synthese* **123**: 35–54.
- Goodall, J. (1968). The behavior of free-living chimpanzees in the Gombe Stream Reserve, *Animal Behavior Monographs* **1**: 161–311.
- Hare, B., Call, J. and Tomasello, M. (2001). Do chimpanzees know what conspecifics know?, *Animal Behavior* **61**: 139–151.
- Hare, B., Call, J. and Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding, *Cognition* **101**: 495–514.
- Hare, B. and Tomasello, C. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks, *Animal Behavior* **68**: 571–581.
- Hirata, S., Myowa, M. and Matsuzawa, T. (1998). Use of leaves as cushions to sit on wet ground by wild chimpanzees, *American Journal of Primatology* **44**: 215–220.

- Hutto, D., Herschbach, M. and Southgate, V. (2011). Social cognition: Mindreading and alternatives. Editorial to the special issue, *Review of Philosophy and Psychology* **2**: 375–395.
- Kaminski, J., Call, J. and Tomasello, M. (2008). Chimpanzees know what others know but not what they believe, *Cognition* **109**: 224–234.
- Krachun, C., Call, J. and Tomasello, M. (2009). Can chimpanzees (*Pan troglodytes*) discriminate appearance from reality?, *Cognition* **112**: 435–450.
- Luo, Y. and Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions?, *Cognition* **105**: 489–512.
- Lurz, R. (2011). *Mindreading Animals: The Debate over What Animals Know about Other Minds*, The MIT Press, Cambridge, MA.
- Lurz, R., Kanet, S. and Krachun, C. (2014). Animal mindreading: In defense of optimistic agnosticism, *Mind and Language* **29**: 428–454.
- Lurz, R. and Krachun, C. (2011). How could we know whether nonhuman primates understand other's internal goals and intentions? Solving Povinelli's problem, *Review of Philosophy and Psychology* **2**: 449–481.
- McGrew, W. (2004). *The Cultured Chimpanzee: Reflections on Cultural Primatology*, Cambridge University Press, Cambridge.
- Michael, J. (2015). Putting unicepts to work: A teleseomantic perspective on the infant mindreading puzzle, *Synthese* **forthcoming**.
- Michael, J. and Christensen, W. (2016). Flexible goal attribution in early mindreading, *Psychological Review* **123**: 219–227.
- Millikan, R. (1996). Pushmi-pullyu representations, in J. Tomberlin (ed.), *Philosophical Perspectives*, Vol. 9, Ridgeview Publishing, Atascadero, CA, pp. 185–200.
- Millikan, R. (2004a). On reading signs: Some differences between us and the others, in D. K. Oller and U. Griebel. (eds), *The Evolution of Communication Systems: A Comparative Approach*, The MIT Press, Cambridge, MA, pp. 15–30.
- Millikan, R. (2004b). *Varieties of Meaning*, The MIT Press, Cambridge, MA.
- Moll, H. and Meltzoff, A. (2011). How does it look? Level 2 perspective-taking at 36 months of age, *Child Development* **82**: 661–673.
- Onishi, K. and Baillargeon, R. (2005). Do 15-month-old understand false beliefs?, *Science* **308**: 255–258.

- Penn, D. and 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. Series B, Biological Sciences* **362**: 731–744.
- Phillips, A., Welman, H. and Spelke, E. (2002). Infants' ability to connect gaze and emotional expression to intentional action, *Cognition* **85**: 53–78.
- Povinelli, D. and Vonk, J. (2003). Chimpanzee minds: Suspiciously human?, *Trends in Cognitive Sciences* **7**: 157–160.
- Povinelli, D. and Vonk, J. (2004). We don't need a microscope to explore the chimpanzee's mind, *Mind and Language* **19**: 1–28.
- Schmelz, M., Call, J. and Tomasello, M. (2013). Chimpanzees predict that a competitor's preference will match their own, *Biological Letters* **9**: 1–5.
- Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*, Cambridge University Press, Cambridge.
- Sodian, B., Thoermer, C. and Metz, U. (2007). Now I see but you don't: 14-month-old can represent another person's visual perspective, *Developmental Science* **10**: 199–204.
- Southgate, V., Senju, A. and Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds, *Psychological Science* **18**: 587–592.
- Tomasello, M. and Call, J. (2006). Do chimpanzees know what others see—or only what they are looking at?, in S. Hurley and M. Nudds (eds), *Rational Animals?*, Oxford University Press, Oxford, pp. 371–384.
- Tomasello, M., Call, J. and Hare, B. (2003). Chimpanzees understand psychological states: The question is which ones and to what extent, *Trends in Cognitive Sciences* **7**: 153–156.
- Wellman, H. (2014). *Making Minds: How Theory of Mind Develops*, Oxford University Press, Oxford.
- Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach, *Cognition* **69**: 1–34.
- Woodward, A. (2005). The infant origins of intentional understanding, *Advances in Child Development and Behavior* **33**: 229–262.
- Zawidzki, T. (2011). How to interpret infant socio-cognitive competence, *Review of Philosophy and Psychology* **2**: 483–497.