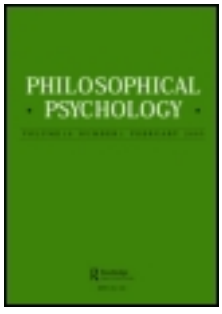

Philosophical Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cphp20

If chimpanzees are mindreaders, could behavioral science tell? Toward a solution of the logical problem

Robert Lurz

Version of record first published: 15 Jun 2009

To cite this article: Robert Lurz (2009): If chimpanzees are mindreaders, could behavioral science tell? Toward a solution of the logical problem, Philosophical Psychology, 22:3, 305-328

To link to this article: http://dx.doi.org/10.1080/09515080902970673

Philosophical Psychology
Vol. 22, No. 3, June 2009, 305–328

If chimpanzees are mindreaders, could behavioral science tell? Toward a solution of the logical problem

Robert Lurz

There is a persistent methodological problem in primate mindreading research, dubbed the 'logical problem,' over how to determine experimentally whether chimpanzees are mindreaders or just clever behavior-readers of a certain sort. The problem has persisted long enough that some researchers have concluded that it is intractable. The logical problem, I argue, is tractable but only with experimental protocols that are fundamentally different from those that have been currently used or suggested. In the first section, I describe what the logical problem is (and is not) and how it can, in principle, be solved. In the second section, I illustrate how a well-known experimental protocol by Hare et al. (2000) fails to solve the logical problem. In the third section, I do the same for a protocol by Heyes (1998). (I do the same in the appendix for a recently suggested protocol by Penn and Povinelli (2007).) In the fourth section, I describe a novel experimental protocol for visual perspective-taking and argue that it succeeds to discriminate between the relevant mindreading and behavior-reading hypotheses. In addition, this new experimental protocol employs procedures that are realistic enough to suppose that chimpanzees might very well succeed in passing them.

Keywords: Appearance-reality Distinction; Behavior-reading; Chimpanzees; Mindreading; Optical-illusion

1. Introduction

The *mindreading hypothesis* in primatology is the hypothesis that chimpanzees have mental concepts, such as *see* and *belief*, that they apply to themselves and others

Robert Lurz is Associate Professor at Brooklyn College, CUNY.
Correspondence to: Robert Lurz, Brooklyn College, Philosophy, 2100 Bedford Avenue, Brooklyn, NY 11210. Email: rlurz@brooklyn.cuny.edu

ISSN 0951-5089 (print)/ISSN 1465-394X (online)/09/030305-24 © 2009 Taylor & Francis
DOI: 10.1080/09515080902970673

(principally, conspecifics) for the purpose of predicting others' behaviors. The hypothesis takes a minimally functionalist view of mental concepts as they are applied to others (Penn & Povinelli, 2007; Whiten, 1996). On this view, mental concepts when applied to others describe subjects as being in a non-observable state that causally mediates between observable features of their environment and their observable behavior. As a result, mental concepts, on this view, are always applied to others on the basis of some observational fact or judgment about the other subject's behavior or environment. The *behavior-reading hypothesis* in primatology, on the other hand, is the hypothesis that chimpanzees lack mental concepts altogether and anticipate the behaviors of other animals on the basis of what they know or believe (from experience, inference, or innately) about the contingencies existing between such behaviors and the observable environment.

Since mental concepts are applied to others on the basis of some observational fact or judgment about the other subject's behavior or environment, every mindreading hypothesis in primatology has a *complementary behavior-reading hypothesis*, a hypothesis in which the test subject is taken to use the very same observational fact to anticipate another animal's behavior that, on the mind-reading hypothesis, the test subject uses as the observational basis for applying the mental concept in question. The methodological problem that has plagued primate mindreading research is that with all standard experimental protocols, any positive result that can plausibly be explained or predicted by a mindreading hypothesis can plausibly be explained and predicted by its complementary behavior-reading hypothesis. It is this problem that Povinelli and Vonk (2004) and Hurley and Nudds (2006) label the 'logical problem'. The question that constitutes the logical problem, then, is whether an experimental protocol can be designed whose positive results can be plausibly explained and predicted by a mindreading hypothesis but not by its complementary behavior-reading hypothesis.[1]

Since the logical problem is sometimes confused with other methodological issues in primate mindreading research, it is important to state what it is not. The logical problem is *not* that current experimental protocols fail to control for *any* plausible behavior-reading hypothesis, for they obviously do. Some behavior-reading hypotheses have been controlled for and eliminated by current experiments (see, for example, Bräuer, Call, & Tomasello, 2007; Hare, Call, Agnetta, & Tomasello, 2000; Hare, Call, & Tomasello, 2001). Nor is the logical problem that no current experimental protocol can control for *every* plausible behavior-reading hypothesis, for no experimental protocol can do that. Rather, the logical problem is that no current experimental protocol can control for a plausible behavior-reading hypothesis that is *complementary* to the mindreading hypothesis under consideration.

The reason that this problem has been described as a *logical* problem is that some researchers have thought that it is, in principle, experimentally insoluble, and that, as a result, primate mindreading research is doomed from the start. Although rarely

stated explicitly the general argument for the problem's insolubility can be expressed in the following steps:

> *Step one*. Consider a mindreading hypothesis (MRH) for a chimpanzee A. According to MRH, A anticipates that another chimpanzee B will perform behavior *r* on the grounds that B is in some mental states *m*, and that B's being in *m* is likely to lead him to do *r*.
>
> *Step two*. Since mental concepts applied to others are based on observational facts about the subject's behavior or environment, A must apply the mental concept *m* to B on the grounds of some observational fact, *s*, about B's behavior or environment.
>
> *Step three*. Now generate the complementary behavior-reading hypothesis (CBRH) by replacing A's mental-state attributions of *m* to B with A's observational grounds for this attribution – that is, A's judgment that *s* obtains. According to CBRH, A anticipates that B will do *r* simply on the grounds that *s* obtains, and that when *s* obtains, B is likely to do *r*.
>
> *Step four*. If it is plausible to suppose that prior to his anticipation of B's behavior, A has experienced *s*-type conditions followed by *r*-type behaviors in other animals or in himself, then it is plausible (perhaps, even more so) to suppose (as CBRH does) that A anticipates B doing *r* simply on the grounds that *s* obtains, as opposed to supposing (as MRH does) that A anticipates B's behavior on the grounds that *s*'s obtaining is evidence of B being in mental state *m*.
>
> *Step five*. No experimental protocol is able to eliminate the plausible supposition stated in step four.
>
> *Conclusion*. No experimental protocol can distinguish a mindreading hypothesis (MRH) from its complementary behavior-reading hypothesis (CBRH).

The conclusion follows, of course, only if step five is true. And although it may be true that no current experimental protocol to date has eliminated the supposition made in step five, it does not follow that no experimental protocol can. In fact we can see, in general terms, what an experimental protocol would have to do in order to eliminate such a supposition. What is needed is an experimental protocol in which a chimpanzee, A, is tested on whether he will anticipate another chimpanzee, B, doing *r* on the basis of some observable fact *s* such that from the set up of the experiment it is plausible to suppose, on the assumption that A is a mindreader that possesses the mental concept *m*, that:

(i) A has reason to believe that the observable fact *s* is evidence of mental state *m* in B;
(ii) A has reason to believe that mental state *m* in B will lead to B doing *r*; but
(iii) prior to his anticipation of B's behavior, A has *no* independent reason to believe—independent, that is, from what can be inferred from (i) and (ii)—that *s*-type conditions lead to *r*-type behavior in other animals or in himself.

In such an experimental situation, a complementary behavior-reader is unlikely to have the right sort of belief to make the prediction that B will do *r* when *s* obtains, since he has no reason to believe that the observable condition *s* is likely to lead to B doing *r*. A mindreader, on the other hand, has the potential to acquire the right sort

of belief, for he need only combine his beliefs identified in (i) and (ii) to arrive at the belief that B is likely to do *r* when *s* obtains.[2] So we have here a general framework for how, in principle, to solve the logical problem. Before I show how to realize this framework in an actual experimental protocol, it will be instructive to see how a well-known experiment fails to solve the logical problem as a result of its failure to implement the above framework.

2. Hare et al.'s (2000) Visual-Perspective Taking Experiment

In Hare et al. (2000), a subordinate and a dominant chimpanzee are housed in separate rooms on either side of a middle room where they compete for food. In the preliminary phase of the experiment, the doors of the chimpanzees' rooms are opened fully, and while the dominant is in the center of his own room and the subordinate is in the center of his own room (or inside the dominant's room), a single piece of food is placed either inside the subordinate's room or inside the dominant's room. In this stage of the experiment, the subordinate chimpanzee learns quickly that he can retrieve food that is both visibly and physically inaccessible to the dominant chimpanzee, and that the dominant chimpanzee will always retrieve food that is visible and accessible to him.

In the *occluder test*, two pieces of food are placed in the center of the middle room about 2 meters apart; one piece is placed out in the open, visible to both chimpanzees, and the other is placed on the subordinate's side of an opaque barrier, visible only to the subordinate. The doors of the chimpanzees' rooms are then raised enough so that they both can see the middle room, the respective pieces of food, the opaque barrier, and each other. The subordinate's door is opened first, allowing him a slight head start into the middle room, after which the dominant's door is opened, allowing him to enter the middle room as well. In the *transparent barrier test*, the opaque barrier in the occluder test is switched with a clear transparent barrier. One food piece is placed on the subordinate's side of the transparent barrier while the other piece is placed out in the open. The test then proceeds like the occluder test.

On the mindreading hypothesis, if the subordinate chimpanzee has the mental concept *see*, then he might use it to predict the dominant's behavior so as to increase his own chances of retrieving food. In general, the hypothesis would go something like this. Based upon his experiences with the dominant chimpanzee in the preliminary test phase, the subordinate mindreading chimpanzee could learn:

(a) The dominant chimpanzee will attempt to retrieve all and only food from the location where he *sees* it (or *saw* it) placed.

And based upon his own experiences with transparent and opaque barriers, or upon his experiences of other chimpanzees' behaviors in response to opaque and transparent barriers, the subordinate chimpanzee could learn:

(b) The dominant chimpanzee does not *see* food behind opaque barriers, but he does *see* food behind clear, transparent barriers.

Finally, we would expect that the subordinate mindreading chimpanzee would use his knowledge of (a) and (b) in the occluder test to predict that the dominant chimpanzee will *not* attempt to retrieve the food behind the barrier and, therefore, use this information to retrieve the food for himself. And we would expect that the subordinate chimpanzee would also use his knowledge of (a) and (b) in the transparent barrier test to predict that the dominant chimpanzee *will* attempt to retrieve the food behind the barrier (as well as the food in the open), and that he should not, therefore, try to retrieve either piece of food for himself. And this is precisely what Hare and colleagues discovered: the subordinate chimpanzee showed a decided preference toward retrieving food hidden behind the opaque barrier in the occluder test but showed no tendency toward retrieving food behind the transparent barrier in the transparent test. From these results and other control studies, Hare and colleagues concluded that chimpanzees know what other chimpanzees can and cannot *see*.

The problem with this experimental protocol, as argued by Penn and Povinelli (2007) and Povinelli and Vonk (2004), is that it could just as well be passed by a complementary behavior-reading chimpanzee. Recall that the concept *see* is a mental concept, and mental concepts, when applied to others, are applied on an observable basis. A principle part of the observable basis for thinking that another subject sees an object is that there is a certain observable relation between the object and the subject's open eyes, a relation that Heyes (1994, 1998) calls 'eye-object line', Penn and Povinelli (2007) call 'uninterrupted visual access', and I, following others (Okamoto-Barth, Call, & Tomasello, 2007), shall call 'direct line of sight'. Whatever this observable relation is called, the idea is that in order to judge that another subject sees an object, one must base one's judgment on the observational belief or fact that there is 'an unobstructed, notional, straight line between [the subject's open] eyes' and the object seen (Heyes, 1998, p. 113). Of course the sorts of things that obstruct this notional line of sight and prevent it from being 'direct' or 'uninterrupted' are opaque barriers of certain sizes. So to judge that a subject, S, has direct line of sight with an object, O, is to judge that there is no opaque barrier (of a certain size) on the straight line between S's open eyes and O.

Opaque barriers, of course, are barriers that prevent one from *seeing* items behind them. But this does not mean that judgments about direct line of sight must rest upon judgments about seeing or that the concept *direct-line-of-sight* must include the concept *see*. A subject who makes judgments about direct line of sight needs some way of conceptualizing opaque barriers, of course, but it need not conceptualize them *as* barriers that prevent one from *seeing* items behind them. In fact, a subject's concept of opaque barriers may be primitive in that it cannot be defined or analyzed (by us) by means of other concepts but at best can be functionally identified by means of its inferential role in the mind of the subject. So, for example, a subject's, S's, primitive concept of opaque barriers can be functionally identified (roughly) as the concept C^* such that if S sees (or seems to see) an object O behind a barrier Y of a certain size, then, *ceteris paribus*, S believes that Y is not a C*; and if S sees (or seems to see) Y but does not see (or seem to see) O but believes (based upon the contents

of his working memory of the environment) that O is behind Y, then, *ceteris paribus*, S believes that Y is a C*. Such a subject, then, could come to learn that certain types of barriers, such as windows, glass bottles, and mesh nets, are not C*-type barriers, while other types of barriers, such as wooden doors, painted walls, and rocks of certain size, are.[3] And so, for such a subject to judge that another subject S* has direct line of sight with an object O is to judge that there is no C*-type barrier (of a certain size) on the straight line between S*'s open eyes and object O. A complementary behavior-reader, then, can possess the concept *direct-line-of-sight*.

Furthermore, a subject with the concept *C** of opaque barriers can reasonably judge from the *contents* of his own visual experience along with the *contents* of his working memory about the environment whether he has or lacks direct line of sight with an object. For visual experiences do not simply represent intrinsic features of objects, such as color and shape, but certain relational features as well (Harman, 1990). A subject's visual experience of a tree, for example, does not just represent an object of a certain color and shape; it also represents the object as being in a certain direction and distance from here, where 'here' is roughly the location where the subject's eyes are. What the subject sees, when he sees the tree, is, among other things, an object of a certain shape, size and color that is a certain distance and direction on a straight line from here. Hence if the subject sees the tree, then based on the *content* of his own visual experience—and not on the higher-order fact or judgment *that he is having such an experience*—he has reasonable grounds for believing that the tree is in his direct line of sight, that there is no C*-type barrier of a certain size intersecting the straight line between his eyes and the tree. If, on the other hand, an opaque barrier (Y) of a certain size were to intersect this straight line, then, based upon the content of his current visual experience (i.e., that he has direct line of sight with Y) and what he remembers and continues to believe about the location of the tree, as well as his belief that his location relative to the tree has not changed since the barrier appeared, the subject has reasonable grounds for believing that he no longer possesses direct line of sight with the tree—that there is now a C*-type barrier that intersects the straight line between the tree and his eyes. So a subject can reasonably judge in his own case whether he has or lacks direct line of sight with an object without having or employing the mental concept *see*, or any mental concepts. What is more, if a subject also has the concept *see* and can introspectively judge that he sees such-and-such object, then he will also be in a position to judge, on the basis of the content of his visual experience, that he has direct line of sight with the object. Hence, introspective judgments about seeing will always be confounded with perceptual judgments about direct line of sight.[4]

In Hare et al.'s (2000) experiment, for example, since the subordinate chimpanzee can see the floor (and other items) behind the transparent barrier that he is looking at, he can judge, if he has the primitive concept *C**, that this barrier is not a C*-type barrier and that, therefore, he has direct line of sight with the floor (and other items) behind the barrier. On the other hand, since the subordinate chimpanzee cannot see the floor (or any other item) behind the opaque barrier that he presumably believes (based on the contents of his working memory) is still there, he can judge, if he has

the concept *C**, that that barrier is a C*-type barrier and that, therefore, he does not have direct line of sight with the floor (or any other item) behind it.

If Hare and colleagues are correct and the subordinate chimpanzee is making judgments about what the dominant chimpanzee can and cannot see, then the chimpanzee must be basing these judgments on beliefs or facts about what objects the dominant chimpanzee has and does not have direct line of sight with. In particular, he must be basing them on the respective judgments (1) that the dominant has direct line of sight with the food behind this (the transparent) barrier and (2) that dominant does not have direct line of sight behind that (the opaque) barrier. Arguably the subordinate's reason for judgment (1) is some type of analogical inference along the lines: since I have direct line of sight with objects behind this barrier, it is likely that the dominant does as well. And arguably a similar analogical inference supports the subordinate's judgment (2): since I do not have direct line of sight with objects behind that barrier, it is unlikely that the dominant does either.

But now the question is why we could not predict and explain the subordinate chimpanzee's performance just as well in terms of his making judgments (1) and (2) above. Suppose, then, that the subordinate chimpanzee is a complementary behavior-reader and does *not* have the mental concept *see* but the observational counterpart concept *direct-line-of-sight*. Based upon his experiences in the preliminary test phase of the experiment, such a behavior-reading chimpanzee could learn:

> (a′) The dominant chimpanzee will attempt to retrieve all and only food in those locations with which he has (or has had) direct line of sight.

And based upon his own experiences with the transparent barrier and with the opaque barrier, as well as upon his experiences of the dominant's behavior toward food placed behind walls in the preliminary testing phase of the experiment, the complementary behavior-reading subordinate chimpanzee could come to learn:

> (b′) The dominant chimpanzee has direct line of sight with food behind this (transparent) barrier but does not have direct line of sight with food behind that (opaque) barrier.

The subordinate chimpanzee, then, could use his knowledge of (a′) and (b′) in the occluder test to predict that the dominant chimpanzee will *not* attempt to retrieve the food behind that barrier, as well as use it in the transparent barrier test to predict that dominant chimpanzees *will* attempt to retrieve food behind this barrier. And so, Hare et al.'s experimental protocol does not succeed in distinguishing a genuine mindreading chimpanzee from its complementary behavior-reading counterpart. Again, let me reiterate that the problem here is not that Hare et al.'s protocol cannot distinguish the mindreading hypothesis under consideration from any plausible behavior-reading hypothesis, for it obviously can. Rather, the problem is that it cannot distinguish the mindreading hypothesis from the complementary behavior-reading hypothesis. And that, again, is the logical problem to be solved.

3. Heyes' (1998) Experience-Projecting Protocol for Visual Perspective-Taking

In Heyes (1998), an experimental protocol is described in which the test animal is allowed the opportunity to learn something about its own mental state of *seeing* in relation to a novel observable stimulus and is then tested to determine whether it will use this knowledge to anticipate the behavior of a trainer when presented with the novel stimulus. In the pretraining phase of the experiment, the chimpanzee is trained to wear different colored goggles over its eyes. Red goggles are equipped with an opaque lens cover that prevents the chimpanzee from seeing anything in its environment; blue goggles are equipped with a clear transparent lens cover that allows the chimpanzee to see things in its environment. From a distance, the only discernable difference between the goggles is their respective colors.

In the training phase of the experiment, the chimpanzee watches while one of four containers is bated by an experimenter behind a screen while a second experimenter (the Knower) watches and a third experimenter (the Guesser) does not (e.g., he may have his back turned or be out of the room). The Knower and the Guesser (upon his return) then point to different containers. The chimpanzee is rewarded for choosing the container indicated by the Knower. The probe trial stage of the experiment is just like the training stage except that the Knower, while watching the baiting process, wears blue (transparent) goggles and the Guesser, who is present and whose head is directed toward the baiting process like the Knower's, wears red (opaque) goggles. Heyes reasoned that 'if chimpanzees have the concept ''see,'' then on probe trials one would expect them to choose the Knower, wearing transparent goggles, more often than the Guesser, wearing opaque goggles' (p. 113). On this mindreading hypothesis, the chimpanzee could learn through analogical reasoning about its own ability with seeing things in its environment while wearing the different color goggles that:

(a) Red goggles prevent one from *seeing* objects in the environment, while blue goggles do not.

That is, the chimpanzee could learn from his own case that red goggles, but not blue goggles, prevent him from seeing things in the environment, and from this he could reason by analogy that similar results are to be expected for other subjects wearing the respective goggles. From his experiences in the training phase of the experiment, the mindreading chimpanzee could learn:

(b) Food is under the container indicated by the trainer that *saw* the baiting process.

Finally from his knowledge of (a) and (b), the mindreading chimpanzee in the probe trial could predict:

(c) Food is under the container indicated by the trainer with the blue goggles who *saw* the baiting process.

The question, of course, is whether on the complementary behavior-reading hypothesis one would also expect the behavior-reading chimpanzee with the concept *direct-line-of-sight* to prefer the Knower with blue goggles over the Guesser with red

goggles in the probe trial. And the answer is yes. For the complementary behavior-reading chimpanzee could also learn through analogical reasoning about his own ability to have direct line of sight with objects in the environment while wearing the different colored goggles that:

> (a′) Red goggles prevent one from having direct line of sight with objects in the environment, while blue goggles do not.

That is, the chimpanzee could discover in his own case that blue goggles allow direct line of sight with objects in the environment, while red goggles place an opaque barrier between his open eyes and objects in his environment that, from the contents of his working memory, he knows to be there at the moment of his donning the red goggles. And from this, he could reason by analogy that similar results are likely to occur for other subjects wearing the respective goggles. From his experiences in the training phase of the experiment, the behavior-reading chimpanzee could learn:

> (b′) Food is located under the container indicated by the trainer who had direct line of sight with the baiting process.

Finally from its knowledge of (a′) and (b′), the behavior-reading chimpanzee in the probe trial could predict:

> (c′) Food is located under the container indicated by the trainer wearing blue goggles who had direct line of sight with the baiting process.

Hence, Heyes' experience-projecting protocol for visual-perspective taking seems to fare no better than Hare et al.'s (2000) at discriminating between genuine mindreading chimpanzees and their complementary behavior-reading counterparts.

In response to an earlier criticism along these lines by Lurz (2001), Heyes (2001) responded by challenging the assumption that in the pretraining phase of the experiment, the chimpanzee could learn in its own case that red goggles, but not blue goggles, prevent direct line of sight with objects in the environment. Her reasoning for saying this was:

> When I, or a chimpanzee, put on the [blue], translucent goggles, I see what is before me; I do not, as if from a combination of first and third person perspectives, see myself wearing the [blue] goggles, the object before my eyes, and myself seeing that object. (p. 1144)

But given what we said in section 2, Heyes' objection here can be seen to miss its mark. As shown there, to know (or to be justified in believing) that one has direct line of sight with an object does not rest upon seeing oneself seeing the object (whatever that might involve), and neither does it rest upon any introspective judgments about oneself *seeing* the object; rather it rests upon the content of one's own visual experience and working memory of the environment. What one sees when one sees an object is (in part) that the object bears a certain spatial relation to oneself—namely, a direct, unobstructed line. Heyes seems to assume that in seeing 'what is before me', one does not see the object as being in a particular direction and distance from oneself. But, again, visual experiences do not simply represent intrinsic features

of objects but certain relational features, such as direct line of sight, as well. When the chimpanzee dons the blue goggles, he has visual experiences of objects in the environment *as being directly before him* (*or his eyes*). While wearing the red goggles, however, the chimpanzee does not have visual experiences of objects in the environment but visual experiences of the back of the opaque lens of the goggles. What he sees (in part) is the opaque lens being directly before him (or his eyes). But based on his working memory of the objects in the environment, he has reason to believe the objects are where they were before his donning the red goggles and, hence, reason to believe that he now no longer has a direct line of sight with these objects that are still before him (or his eyes).

Moreover, as we saw in section 2, since introspective judgments about seeing are confounded with perceptual judgments about direct line of sight, the mindreading chimpanzee in Heyes' experiment would have access to the same direct line of sight information that the complementary behavior-reading chimpanzee has. This is precisely what Heyes sought to avoid in her experimental protocol and what prevents it from solving the logical problem. What is needed, as I hope is becoming clear, is an experimental protocol that can disconfound the mental concept (*m*) in question and its observable basis for ascription (*s*). As noted, I do not think that this can be done for *seeing* and its observable basis for ascription, but it can be done for *seeing*'*s* close cousin, the concept *seeing-as*, and its observable basis—or so I argue in the next section.

4. An Experimental Protocol that can Solve the Logical Problem

As in Hare et al. (2000), a subordinate and a dominant chimpanzee are housed in separate rooms on either side of a middle room where they compete for food. The chimpanzees' rooms are equipped with guillotine doors that can be raised enough to allow the animals to see the middle room and each other without allowing them to enter the room. The chimpanzees will compete for two types of food placed in the middle room. The first type is a highly desirous yellow banana; the second type is not even real food, but a plastic banana that is colored orange (or dark orange). From a distance, the only discernable difference between real and fake bananas is their respective color.

4.1. Preliminary Phase

In the preliminary phase of the experiment, the doors of the chimpanzees' rooms are opened fully, and while the dominant is in the center of his own room and the subordinate is in the center of his own room (or inside the dominant's room), a single piece of food (a yellow real banana or an orange (or dark orange) fake banana) is placed either inside the subordinate's room or inside the dominant's room. The chimpanzees are then allowed to compete for the placed food. In this stage of the experiment, the subordinate learns quickly that he can retrieve yellow bananas that are both visibly and physically inaccessible to the dominant (e.g., those that are

placed inside his own room while he is there) and that the dominant will retrieve all the yellow bananas that are visibly and physically accessible to him (e.g., those placed within the dominant's room while he is there) but will not attempt to retrieve the orange bananas (after discovering that they are fake.)

4.2. Pretraining Phase

In the pretraining phase of the experiment, we expose only the subordinate to two different types of transparent barriers: a clear transparent barrier that does not change the apparent color of objects behind it, and a red transparent barrier that causes yellow objects placed behind it to look orange and orange objects to look slightly darker orange. Through the placement of various yellow and orange objects behind the barriers, the subordinate is shown or allowed to discover (if a mind-reader) that he can see (and, hence, has direct line of sight with) objects placed behind either type of barrier, but that he sees yellow and orange objects behind the red barrier as orange and dark orange respectively, and yellow and orange objects behind the clear barrier as yellow and orange respectively. Moreover, the subordinate is allowed (or trained if necessary) to retrieve the objects he sees placed behind the barriers. This can be achieved by making the objects interesting for the chimpanzee and encouraging him to retrieve them. One thing that needs to be avoided in this phase of the experiment is the chimpanzee coming to have a greater inclination toward retrieving yellow objects behind clear barriers than behind red barriers. For if the subordinate chimpanzee were a complementary behavior-reader and came to have such a differential retrieval bias, he could, in the critical test (below), project this bias onto the dominant chimpanzee and thereby give the same prediction of the dominant's behavior as his mindreading counterpart is expected to do. Measures can be taken here, however, that would eliminate the possibility of such a differential retrieval bias. First, before the chimpanzee is shown a yellow object placed behind the red barrier, he is encouraged and rewarded for retrieving orange objects placed behind the red barrier and orange and (particularly) yellow objects placed behind the clear barrier. The chimpanzee's level of inclination can be measured by the speed and alacrity with which he retrieves the objects. Once his speed and alacrity for retrieving the various yellow and orange objects behind the clear barrier matches his speed and alacrity for retrieving the various orange objects behind the red barrier, he is then introduced to a yellow object behind the red barrier. To reinforce the likelihood that the chimpanzee will be *initially* inclined toward retrieving the object we make sure the object is extremely attractive to the animal—perhaps, by making it a new toy with salient features. We can then measure his initial speed and alacrity here to make sure that it matches the initial speed an alacrity with which he retrieved yellow objects behind the clear barrier. The chimpanzee is thereafter encouraged and rewarded for retrieving yellow objects placed behind the red barrier. The chimpanzee is promoted to the critical test stage of the experiment only when his overall speed and alacrity for retrieving yellow objects behind the red barrier matches his overall speed and alacrity for retrieving yellow objects behind the clear barrier. These measures (and perhaps

others, as well) could be used to secure that the chimpanzee is no less inclined, both initially and thereafter, toward retrieving yellow objects behind the red barrier than he is toward retrieving yellow objects behind the clear barrier. I see no reasons to think that a chimpanzee could not demonstrate clear signs of success on these measures. After the training phase, the chimpanzees are returned to their respective rooms and the following critical tests are run.

> *Red barrier test*. Two red transparent barriers are placed in the center of the middle room about 2 meters apart. A yellow banana and an orange (or dark orange) fake banana are placed on the subordinate's side of each barrier. The doors are then raised enough to allow each chimpanzee to see the barriers, the food, and each other; the doors are then raised further to allow the chimpanzees to enter the room; the subordinate is given a slight head start.
>
> *First alternative barrier test*. A clear transparent barrier replaces one of the red transparent barriers described in the above test. A yellow banana is placed on the subordinate's side of the red barrier, and an orange (or dark orange) fake banana is placed on the subordinate's side of the clear barrier. The test continues as above.
>
> S*econd alternative barrier test*. This test is just like the first alternative barrier test except that a yellow banana is placed on the subordinate's side of the clear barrier and a yellow banana is placed on the subordinate's side of the red barrier.

Now for the relevant predictions. If the subordinate chimpanzee is a mindreader and has the perceptual concepts *see* and *seeing-as*, then we would expect him to use these concepts to predict the dominant's behavior in the critical tests so as to optimize his own chances of retrieving and eating real bananas. If the subordinate chimpanzee is such a mindreader, then one would expect that, from his experiences in the preliminary stage of the experiment, he could learn:

> (a) Yellow bananas that the dominant does not *see* and (hence) does not *see as* yellow can be retrieved, since the dominant does not retrieve them; yellow bananas that the dominant *sees* and *sees as* yellow are not retrievable, since the dominant retrieves them.

And from his experiences in the pretraining phase of the experiment, the mindreading subordinate could learn through analogical reasoning from its own case:

> (b) Red barriers make one *see* yellow objects behind them *as* orange and orange objects *as* dark orange; clear barriers do not change the way the color of objects behind them *look*.

With his knowledge of (a) and (b), the mindreading subordinate chimpanzee could then predict and choose in the critical tests as follows:

> (c) The yellow banana behind the red barrier is retrievable, since the dominant, *seeing it as orange*, will not attempt to retrieve it.

Now let us consider this mindreading chimpanzee's complementary behavior-reading counterpart. To reach the complementary behavior-reading counterpart we need to replace the mindreading chimpanzee's judgments of seeing in (a) and (b) with their observational grounds. In (a), for example, the mindreading chimpanzee is

taken to judge that the dominant retrieves all and only those yellow bananas that he sees and sees as yellow. Since judgments of seeing are grounded on judgments of direct line of sight, the observational grounds for this judgment must be the judgment (a′) that the dominant retrieves all and only those yellow bananas with which he has direct line of sight. In (b), the mindreading chimpanzee judges that both red and clear barriers allow one to see objects behind them (although only red barriers change the apparent color of yellow objects). Again, since judgments of seeing are grounded on judgments of direct line of sight, the observational grounds for the mindreading chimpanzee's judgments of seeing here must be his judgment (b′) that red and clear barriers allow direct line of sight with objects behind them. With (a′) and (b′) we can now construct the behavior-reading hypothesis that is complementary to the above mindreading hypothesis. On the complementary behavior-reading hypothesis, the behavior-reading chimpanzee, from his experiences with the dominant chimpanzee in the preliminary stage of the experiment, would be expected to learn:

(a′) Yellow bananas with which the dominant does not have direct line of sight can be retrieved, since the dominant does not retrieve them; yellow bananas with which the dominant has direct line of sight are not retrievable, since the dominant retrieves them.

And from his experiences in the pretraining phase of the experiment, the subordinate chimpanzee would learn through analogical reasoning from his own case:

(b′) Red and clear barriers allow direct line of sight with objects behind them.

But from his knowledge of (a′) and (b′), the complementary behavior-reading chimpanzee would be expected to predict and choose quite differently in the critical tests from his mindreading counterpart. For based on (a′) and (b′), the complementary behavior-reading chimpanzee would be expected to predict and choose as follows:

(c′) The yellow banana behind the red barrier is not retrievable, since the dominant, having direct line of sight with it, will attempt to retrieve it.

Therefore, the complementary behavior-reading hypothesis predicts that the subordinate will behave differently in the critical tests from what the mindreading hypothesis predicts. Hence, the above protocol has the capacity to distinguish a mindreading hypothesis from its complementary behavior-reading hypothesis. The important element in the protocol that enables it to make this distinction is the fact that the observational grounds for the mindreading chimpanzee's judgment that the dominant sees the yellow banana behind the red barrier (i.e., that the dominant (i) has direct line of sight with the yellow banana behind the barrier) is *novel* relative to the predicted behavior of the dominant in the critical tests (i.e., the dominant's behavior of (ii) not retrieving the yellow banana behind the red barrier). Given the unlikelihood that the subordinate chimpanzee has ever experienced red transparent barriers prior to the pretraining phase, it is quite implausible to suppose that he has

ever experienced (i) associated with (ii) in other chimpanzees. So it is implausible to suppose that the subordinate would be able to predict that the dominant will (ii) not retrieve the yellow banana behind the red barrier on the basis of his having (i) direct line of sight with the yellow banana as a result of the subordinate having witnessed this contingency between (i) and (ii) in other chimpanzees.

Could the subordinate chimpanzee have predicted that the dominant would do (ii) on the basis of his having (i) as a result of the subordinate observing *in his own case* the association between (i) and (ii) (or by his observing *in his own case* the association between (i) and his being (iii) initially disinclined to retrieve yellow objects behind the red barrier)? This is not plausible, either. For recall that measures were taken in the pretraining phase precisely to prevent the subordinate from making such associations. Therefore, the subordinate chimpanzee cannot reasonably be interpreted as using a perceived contingency between (i) and (ii) (or between (i) and (iii)) in his own case or in the case of other chimpanzees as his grounds for predicting in the critical tests that the dominant will not attempt to retrieve the yellow banana behind the red barrier.

The above experimental protocol, then, can distinguish experimentally a mind-reading chimpanzee from its complementary behavior-reading counterpart. What the above experimental protocol cannot do, of course, is distinguish the above mindreading hypothesis from all conceivable behavior-reading hypotheses. But again, no experimental protocol can do that, and so this is not a mark against the present protocol *qua* a solution to the logical problem. There are behavior-reading hypotheses that yield the same prediction of the subordinate's behavior in the critical tests as the above mindreading hypothesis. But what is important to note about these alternative behavior-reading hypotheses is that they are, by definition, not the above complementary behavior-reading hypothesis and, hence, do not constitute the logical problem for the above mindreading hypothesis. Furthermore there is no general argument to suppose that they cannot in principle be distinguished from the mindreading hypothesis by running further control tests or experiments.

To illustrate, consider the following alternative behavior-reading hypothesis, ABRH, suggested by Juan Gómez (personal correspondence). According to ABRH, the subordinate behavior-reading chimpanzee, as a result of his experience in the pretraining phase, does not learn (b′), that red and clear barriers allow direct line of sight with objects behind them, as the complementary behavior-reading hypothesis supposes, but learns:

> (b″) that red barriers with yellow objects behind them prevent direct line of sight with the yellow object but allow instead direct line of sight with a numerically distinct but otherwise analogous orange object.

According to ABRH, the subordinate chimpanzee does not understand that the object that he has direct line of sight with while it is behind the red barrier (and looking orange) is the same object that he has direct line of sight with when he moves to the other side of the barrier and looks directly at the object (and now appears yellow). Rather, the chimpanzee thinks that there are two objects—one orange and

one yellow—and that he has direct line of sight with the former while facing the red barrier and direct line of sight with the latter while on the other side of the barrier. And so, it is argued, this behavior-reading chimpanzee, armed with this knowledge of red barriers, as well as with his background knowledge that the dominant chimpanzee retrieves yellow but not orange bananas with which he has direct line of sight, could then predict that that the dominant will *not* attempt to retrieve the yellow banana behind the red barrier, since he does not have direct line of sight with the yellow banana but with an orange banana. And this is the same prediction of the subordinate's behavior that the mindreading hypothesis makes.

There are two serious problems with ABRH, however. First, there is a real question of whether ABRH is coherent. According to the hypothesis, the subordinate chimpanzee in the critical test thinks that the dominant has direct line of sight with an orange banana. But direct line of sight, like seeing, is a factive relation: one cannot have direct line of sight with an object that does not exist in physical space at the time one bears this relation to it, any more so than one can see an object that does not exist in physical space at the time of seeing it. And this is no accident, for direct line of sight is part of the observable basis of judgments of seeing, as explained above. Therefore, the subordinate in ABRH must take the orange banana with which he judges the dominant to have direct line of sight to exist in a physical location at the time that the dominant has direct line of sight with it. But what physical location? If the dominant's direction of gaze is any indication, it is the location that the yellow banana occupies on the subordinate's side of the red barrier. But the subordinate does not (and cannot) see any orange banana in that location or in any other location in the room at this time. It is difficult to understand, then, why the subordinate would think that there is an orange banana existing in physical space at this time; and therefore, it is difficult to understand how the subordinate could think, as ABRH hypothesizes, that the dominant has direct line of sight with an orange banana. In fact, the relation that ABRH seems to credit the subordinate with attributing to the dominant and the (non-existent) orange banana is an *intentional* relation. But intentional relations between subjects and objects are mental relations, *par excellent* (see Chisholm, 1957), and no behavior-reading chimpanzee could possess a concept of such a relation.

Second, even if this coherence problem can be resolved, there is an experimental protocol that can distinguish ABRH from the mindreading hypothesis under consideration. If ABRH is true, then chimpanzees learn, as a result of their exposure to the red barrier in the pretraining phase, that when yellow objects are placed behind red transparent barriers, they do not have direct line of sight with the yellow object but with a numerically distinct but otherwise analogous orange object. By contrast, the mindreading hypothesis assumes that, as a result of their exposure to the red barrier in the pretraining phase, chimpanzees learn that when yellow objects are placed behind red transparent barriers, they continue to see (and, hence, have direct line of sight with) the yellow objects, although the objects look orange. Whether chimpanzees understand red transparent barriers as ABRH hypothesizes or as the mindreading hypothesis hypothesizes would appear to be an eminently

tractable empirical question.[5] Adapting an experimental protocol used by Sato, Kanazawa, and Fujita (1997) on chimpanzees for amodal completion tasks with an opaque barrier, an amodal completion task with a red transparent barrier can be designed that is capable of distinguishing the two hypotheses under consideration. After exposing the subordinate chimpanzee to the red barrier in the pretraining phase, we can run the following experiment.

4.3. Training Phase

In the training phase of the experiment, the chimpanzee learns a simple three-choice delayed match-to-sample task. The target stimuli are (i) an intact yellow block (30h × 5w × 5l cm), (ii) a disjoint rectangular block consisting of an orange block (15h × 5w × 5l cm) on which a yellow block (15h × 5w × 5l cm) is placed inline atop, and (iii) an intact orange block (15h × 5w × 5l cm) placed directly in front of a larger intact yellow block (30h × 5w × 5l cm) (see figure 1). The chimpanzee is shown through manual demonstration that stimuli (ii) and (iii) consist of two separate blocks. In the training phase, the target stimuli are always presented in front of a red transparent barrier (15h × 0.25w × 30l cm). The target stimuli are presented for about a second and then hidden behind a retractable screen. The chimpanzee is then presented with three cards on which each target stimulus is pictured. The chimpanzee is trained to pick the card that has the picture of the target stimulus that was presented earlier.

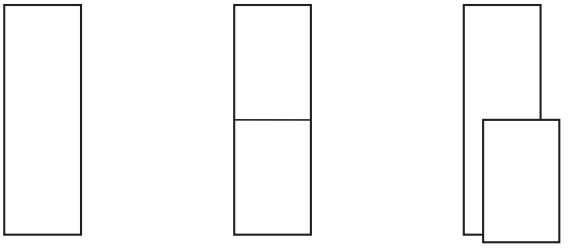

Figure 1 Target stimuli.

4.4. Testing Phase

Once the chimpanzee reaches criterion on the training task, he is administered the following critical test (as well as various control tests). In the critical test, the chimpanzee is presented with a red transparent barrier with an intact yellow block behind it. Given the dimension of the block and the barrier, the top half of the yellow block (15 cm) is above the barrier. For convenience of reference, let us label the top

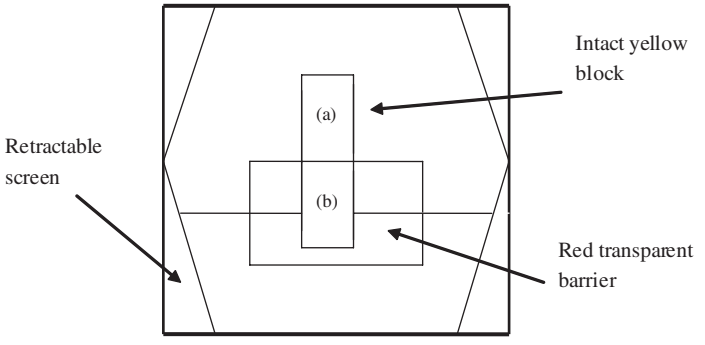

Figure 2 Red transparent barrier test.

part of the block that is above the barrier (a) and the bottom part that is behind the barrier (b) (see figure 2). The target stimulus is shown for about one second and then hidden behind the screen. The chimpanzee is then given the three cards from which to choose.

Now for the relevant predictions. The mindreading hypothesis would predict that the chimpanzee would treat (a) and (b) as parts of one yellow block and would likely choose the picture of the intact yellow block. ABRH, on the other hand, would predict that the chimpanzee would not treat (a) and (b) as parts of one intact block, for ABRH predicts that the chimpanzee will think that the orange block (b) with which it has direct line of sight is numerically distinct from the yellow block (a) with which it has direct line of sight. Therefore, ABRH would predict that the chimpanzee would choose one of the cards containing the disjoint blocks. ABRH, then, can be empirically distinguished from the mindreading hypothesis.

It is important to note at this point that the visual-perspective taking protocol presented here can be realized in different ways, and it may be that some of these ways are easier or more practicable to run, or are less cognitively demanding for the chimpanzees. So it is useful to show how the above protocol could be realized by using different materials. Staying within the same modality (i.e., vision), one could run the above protocol with clear transparent barriers that distort the sizes of objects placed behind them. In the preliminary stage of the experiment, subordinate and dominant chimpanzees would compete for two different size foods (small bananas and large bananas, for example), and the subordinate would learn that the dominant will always retrieve the larger of two bananas to which he has visual and physical access. In the pretraining phase of the experiment, the subordinate is introduced to three different kinds of clear transparent barriers that are distinguished by different colored trim: a non-distorting, black-trimmed barrier that does not change the apparent size of objects behind it; a minimizing, red-trimmed barrier that makes objects behind it look smaller than they are; and a magnifying, blue-trimmed barrier that makes objects behind it look larger than they are. Through the placement of various objects behind the barriers, the subordinate comes to learn, if he is

a mindreader, that he can see (and, thus, has direct line of sight with) objects behind all three barriers but that objects look larger than they are behind the blue-trimmed barrier and smaller than they are behind the red-trimmed barrier. As with the other experimental protocol, measures are taken in this phase of the experiment to guarantee that the subordinate does not learn that he is more likely to retrieve (or is initially more inclined to retrieve) objects placed behind the blue-trimmed (magnifying) barrier than he is to retrieving (or being initially inclined to retrieve) objects placed behind the red-trimmed (minimizing) barrier. One can eliminate the possibility of such associations being made by placing objects that are novel and salient to the chimpanzee behind the red-trimmed barrier and objects that are less interesting (but still interesting) to the chimpanzee behind the blue-trimmed barrier, as well as by introducing the chimpanzee to the blue-trimmed barrier after he has reached criterion for retrieving, without hesitation, objects behind the red-trimmed barrier. Again, I see no reason to think that chimpanzees could not show clear signs of passing these measures.

In the testing phase of the experiment, the subordinate and dominant chimpanzees compete for food placed behind the different barriers. The critical tests are those in which (a) a small banana is placed behind a magnifying barrier and a large banana is placed behind a minimizing barrier, (b) a small banana is placed behind a magnifying barrier and another small banana is placed behind a non-distorting barrier, and (c) a large banana is placed behind a minimizing barrier and another large banana is placed behind a non-distorting barrier. The mindreading hypothesis would predict that the subordinate chimpanzee would optimize its choices of bananas by anticipating that the dominant will retrieve the banana that it incorrectly *sees as* larger. So, for example, in test (a), the mindreading hypothesis would predict that the subordinate chimpanzee will attempt to retrieve the larger of the two bananas, since the dominant, incorrectly seeing the smaller banana as larger, will attempt to retrieve it. The complementary behavior-reading hypothesis, however, would predict that the subordinate chimpanzee will *not* attempt to retrieve the larger banana. For according to this hypothesis, the subordinate chimpanzee understands that the dominant has direct line of sight with both bananas behind the barriers, and he knows from past experience that whenever the dominant has direct line of sight with a large and small banana, he always retrieves the former over the latter.

There are, of course, various possible behavior-reading hypotheses that would yield the same prediction. But these alternative hypotheses can be handled in the same way that ABRH was handled above. First, the hypotheses are shown not to be complementary behavior-reading hypotheses and, therefore, not to constitute a logical problem for the mindreading hypothesis under consideration. Second, it can be shown how certain control experiments or further tests, analogous to those described above for ABRH, can distinguish these hypotheses from the mindreading hypothesis under consideration. For example, suppose that it is hypothesized that chimpanzees, after their exposure to the distorting barriers in the pretraining phase, learn that (e.g.) the blue-trimmed (magnifying) barrier prevents direct line of sight with a small object placed behind it but allows direct line of sight with a numerically

distinct but otherwise analogous large object. Again, an experimental protocol similar to the one described above can distinguish this hypothesis from the mindreading hypothesis that assumes that the chimpanzees would learn that the blue-trimmed barrier allows direct line of sight with the small object behind it but makes the object look large. In the critical test of the experiment, the chimpanzee is presented with a blue-trimmed (magnifying) barrier with an intact block behind it. One half of the block (a) is above the barrier; the other half (b) is behind the barrier and looking larger than (a) (see figure 3). The mindreading chimpanzee would treat (a) and (b) as parts of the same block; the behavior-reading chimpanzee under consideration would not. This difference would be manifested in their choices of different cards depicting an intact block and various disjoint blocks.

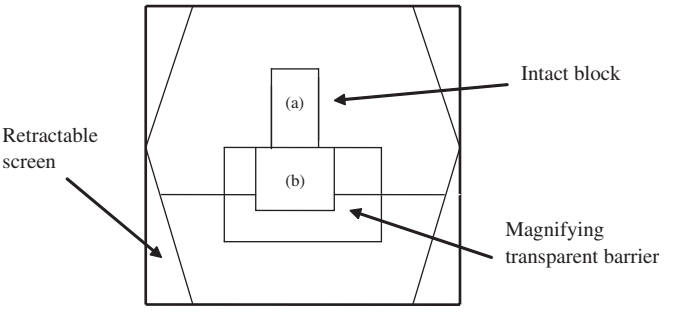

Figure 3 Magnifying barrier test.

And third, in some cases, the hypotheses can be shown to be incoherent, as in the case of a hypothesis that predicts (similar to ABRH) that in the critical tests the subordinate chimpanzee thinks that the red-trimmed (minimizing) barrier blocks the dominant's direct line of sight with the large banana behind it but allows him to have direct line of sight with a (non-existent!) small banana, and that the blue-trimmed (magnifying) barrier blocks the dominant's direct line of sight with the small banana behind it but allows him to have direct line of sight with a (non-existent!) large banana.

Finally, it needs mentioning that the experimental protocol being offered here is designed not only to test whether chimpanzees have the concept *seeing-as* but also the concept *see*, since the possession of the former concept arguably entails the possession of the latter. Admittedly, the concept *seeing-as* is a more complex concept than the concept *see*; and, therefore, it is possible that chimpanzees have the latter concept but not the former. And so it is possible for chimpanzees to fail the critical tests in the experimental protocol and still be mindreaders that possess the concept *see*. This may seem to be a problem for the experimental protocol, but I do not believe so. First, the explicit aim of the experimental protocol is to provide a *sufficient* condition for possessing the concepts *seeing-as* and (by entailment) the concept *see*,

not a necessary condition. So the fact that failure to pass the critical tests is consistent with chimpanzees possessing the concept *see* (as well as *seeing-as*) is not a mark against the protocol. Second, given the methodological problem at issue, I see no way of solving it successfully for the concept *see* without testing for the concept *seeing-as*, since the concept *seeing-as*, unlike the concept *see*, allows for the possibility of mistaken perception which, in turn, allows for the possibility of a prediction that cannot be made simply on the basis of an observable fact like of direct line of sight.[6] Of course, since the concept *see* is part of the concept *seeing-as*, if a chimpanzee passes the critical tests for the concept *seeing-as*, he passes it for the concept *see*, as well.

5. Conclusion

No experimental protocol is perfect, and the above experimental protocol does not aim to be. Its aim is to provide a general strategy for designing experiments that have a far better chance than any proposed or used of experimentally distinguishing mindreading hypotheses from their complementary behavior-reading hypotheses. It is the rationale behind the experimental protocol, not its particular materials or procedures, that is relevant. Without such a change in the underlying rationale, behavioral science shall never know whether chimpanzees are mindreaders.

Appendix 1. Why Penn and Povinelli's (2007) Experimental Protocol cannot Solve the Logical Problem

The most recent attempt to offer an experimental protocol aimed at solving the logical problem is by Penn and Povinelli (2007). In their protocol, a subordinate and a dominant chimpanzee are housed in separate rooms on either side of a middle room where they compete for food. The middle room is divided into 5 (or more) evenly spaced stalls. At the center of each stall is a bucket where food will be hidden. The chimpanzees' rooms are equipped with guillotine doors that can be opened enough to allow the animals to see the adjoining middle room, each other, and all the buckets in the stalls. In the first stage of the experiment, the subordinate and the dominant are exposed individually to a series of non-competitive trials. In each trial, a large food reward and a small food reward are placed under two separate buckets in full view of the chimpanzee who is watching from his room. The chimpanzee is then released into the middle room and is allowed to approach and retain the contents of only one bucket. The trial is repeated until the chimpanzee reliably approaches and retains the larger food reward. In the second stage of the experiment, a subordinate and a dominant chimpanzee watch from their respective rooms while two buckets are baited with the different amounts of food. The chimpanzees are then released into the middle room and allowed to compete for the hidden food. Only subordinate-dominant dyads in which the subordinate reliably retrieves the

smaller food reward and the dominant reliably retrieves the larger food reward participate in the final stage of the experiment. In the final stage, a number of experimental tests are run (in random order) in which the dominant's visual access to the middle room and baiting process are manipulated. The following four tests are, according to Penn and Povinelli, the critical tests:

> *Removed uninformed test*. The subordinate and the dominant watch while two buckets are baited with the different amounts of food. While the dominant's door is down, the contents of one of the buckets is removed and the bucket is replaced empty.
>
> *Replaced test*. The subordinate and the dominant watch while one of the buckets is baited with one of the two food rewards. While the dominant's door is down, the food reward is moved to a new bucket and the old bucket is baited with the other food reward.
>
> *Misinformed test*. The subordinate and the dominant watch while two buckets are baited with the different amounts of food. While the dominant's door is down, one of the buckets (which may not have been baited) is moved to where one of the baited buckets is, and that bucket and its contents are moved to a new location, and the bucket occupying that location is moved to the location occupied by the first bucket.
>
> *Swapped test*. The subordinate and the dominant watch while one of the buckets is baited with one of the two food rewards. While the dominant's door is down, the locations of the baited buckets is swapped.

Penn and Povinelli claim that 'there is no way for a [subordinate chimpanzee] to reliably pass these critical conditions without the ability to keep track of the counterfactual state of affairs from the dominant's cognitive perspective while simultaneously keeping track of the occurrent state of affairs from the subject's own perspective'—that is, without attributing false beliefs to the dominant chimpanzee regarding the location and amount of food (p. 740). More explicitly, the mindreading hypothesis that Penn and Povinelli have in mind here runs as follows. From his experiences in the first stage of the experiment, the mindreading subordinate chimpanzee could learn:

> (a) Choose the bucket with the largest food reward possible.

And from his experiences in the second stage of the experiment, the mindreading subordinate chimpanzee could learn:

> (b) The dominant chimpanzee will choose the bucket that he *saw* baited with the larger food reward and currently *believes* still is.

Finally with its knowledge of (a) and (b), the mindreading subordinate chimpanzee in the critical tests could predict the behavior of the dominant chimpanzee and optimize his own choices of food rewards. First, in the removed uninformed test, the mindreading subordinate chimpanzee would be expected to predict and choose as follows:

> ($c_{removed}$) If the smaller reward remains under the bucket after the removal process, then choose that bucket, since the dominant, *falsely believing* of the other bucket that it is still baited with larger food reward, will choose that bucket; however, if the

> larger reward remains under the bucket after the removal process, then do not choose that bucket, since the dominant, *correctly believing* that that bucket is still baited with the larger food reward, will choose it.

In the replaced test, the mindreading subordinate chimpanzee would be expected to predict and choose as follows:

> ($c_{replaced}$) If the moved reward is the larger food reward, then choose it, since the dominant, *falsely believing* of the originally baited bucket that it has the larger food reward, will choose the originally baited bucket; however, if the moved reward is smaller, then choose the originally baited bucket with the larger food reward, since the dominant, *falsely believing* that it is baited with the smaller reward, will not choose it.

In the misinformed test, the mindreading subordinate chimpanzee would be expected to predict and choose as follows:

> ($c_{misinformed}$) If the bucket with the larger reward was moved, then choose it, since the dominant, *falsely believing* of the bucket that is now in its place is baited with the larger reward, will choose that bucket; however, if the bucket with the larger reward was not moved, choose the other baited bucket, since the dominant, *correctly believing* that the bucket originally baited with the larger reward still is, will choose that bucket.

And finally in the swapped test, the mindreading subordinate chimpanzee would be expected to predict and choose as follows:

> ($c_{swapped}$) Choose the bucket with the larger reward, since the dominant, *falsely believing* of the bucket that is now in its place is baited with the larger reward, will choose that bucket.

However, given what was said in section 2 in response to Hare et al.'s (2000) experiment, it should be fairly obvious what is wrong with Penn and Povinelli's experimental protocol. Since the mindreading subordinate chimpanzee's false-belief attributions to the dominant are based on his past-tense seeing attributions to the dominant and these, in turn, are based upon his past-tense direct-line-of-sight attribution to the dominant, the complementary behavior-reading chimpanzee employing his concept *direct-line-of-sight* would be expected to behave in precisely the same way as Penn and Povinelli expect their mindreading subordinate chimpanzee to behave. For example, in the first stage of the experiment, the complementary behavior-reading chimpanzee could learn:

> (a′) Choose the bucket with the largest reward possible.

And from his experiences in the second stage of the experiment, the complementary behavior-reading chimpanzee could learn:

> (b′) The dominant chimpanzee will choose the bucket that he had direct line of sight with while it was being baited with the larger food reward.

And from his knowledge of (a′) and (b′), the complementary behavior-reading chimpanzee would be expected to make all the same predications and choices that

his mindreading counterpart is expected to make. In the removed uninformed test, for example, the complementary behavior-reading subordinate would be expected to predict and choose as follows:

> ($c'_{removed}$) If the smaller reward remains under the bucket after the removal process, then choose that bucket, since the dominant, having had direct line of sight with the other bucket while it was being baited with the larger reward, will choose that bucket; however, if the larger reward remains under the bucket after the removal process, then do not choose that bucket, since the dominant, having had direct line of sight with that bucket while it was being baited with the larger reward, will choose it.

As an exercise for the reader, I leave the task of describing the complementary behavior-reading chimpanzee's predictions of the dominant's behavior and his own subsequent choices of buckets in the remaining three critical tests. I hope it is clear that they will be the same, *mutatis mutandis*, as those of his mindreading counterpart's.

Acknowledgments

I wish to thank the anonymous referees of *Philosophical Psychology*, Dick Byrne, Simon Fitzpatrick, Laurie Santos, and (especially) Juan Gómez and Peter Carruthers for their extremely help comments and advice on earlier drafts of this paper.

Notes

[1] The logical problem, of course, is not restricted to mindreading studies on chimpanzees but extends to mindreading studies on all animals. However, this paper focuses on the former since chimpanzees have been the main experimental subjects in such studies, and (as we shall see) the experimental protocols that are relevant to the logical problem have been with chimpanzees.

[2] Penn and Povinelli (2007) have called for specifying the 'unique causal work' that mindreading provides to chimpanzees over and above complementary behavior-reading. The above experimental framework shows that mindreaders have the capacity to predict certain kinds of novel behaviors that complementary behavior-readers cannot. More precisely, call a type of behavior *r strongly novel* relative to some observable fact *s* for animal *A* iff *A* has no independent reason to believe that *s*-type conditions will lead to *r*-type behavior. What the above experimental framework identifies are types of situations in which a mindreading subject can predict others' behaviors that are *strongly novel* (relative to observable fact or cue *s*) on the basis of *s* but in which a complementary behavior-reader could not.

[3] Of course, chimpanzees are unlikely to encounter glass windows or mesh nets in the wild; nevertheless, it is quite likely that they will have encountered other types of transparent barriers (e.g., sparsely grouped trees or vegetation, sheets of rain, water, etc.) that they could categorize as being non-C* type. Hence, although the experimental protocol advanced in this paper employs a transparent barrier that does not occur in the chimpanzee's natural environment, this in no way entails that only captive chimpanzees who have had prior experience with such artificial barriers will be able to pass the critical tests, or that the experimental protocol itself fails to bear any resemblance to the sorts of mindreading/

behavior-reading problems involving transparent and opaque barriers that chimpanzees are likely to encounter in the wild.

[4] This point will bear on my critique of Heyes' (1998) protocol which, I argue, fails to disconfound the test subject's introspective judgments about seeing and his perceptual judgments about direct line of sight.

[5] How chimpanzees understand optical illusions of the sorts employed in the proposed experimental protocol (as well as in general) is a major project in its own right and ideally ought to be tested in advance of the proposed protocol. To my knowledge, very little (unfortunately) has been done to test chimpanzees' understanding of such illusions; however, Juan Gómez (personal communication) has expressed an interest in running some preliminary experiments pertaining to this question.

[6] A similar line of reasoning was once made by Dennett (1978) when he wrote: "a tactic for embarrassing behaviorists in the laboratory is to set up experiments that deceive the subject: If the deception succeeds, their behavior is predictable from their false beliefs [or false perceptions] about the environment, not from the actual environment . . .. So if we are to have good evidence that some system *S* is *not* a behaviorist—is a second-order intentional system—it will only be in those cases where behaviorism would not explain *S*'s success in manipulating [or predicting] another system's behavior" (p. 275).

References

Bräuer, J., Call, J., & Tomasello, M. (2007). Chimpanzees really know what others can see in a competitive situation. *Animal Cognition*, *10*, 439–448, doi 10.1007/s10071-007-0088.

Chisholm, R. (1957). *Perceiving*. Ithaca: Cornell University Press.

Dennett, D. (1978). Conditions for personhood. *Brainstorms*. Cambridge, MA: MIT Press.

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

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behavior*, *61*, 139–151.

Harman, G. (1990). The intrinsic qualities of experience. *Philosophical Perspectives*, *4*, 31–52.

Heyes, C. (1994). Cues, convergence and a curmudgeon: A reply to Povinelli. *Animal Behavior*, *48*, 242–44.

Heyes, C. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, *21*, 101–148.

Heyes, C. (2001). Theory of mind and other domain-specific hypotheses. *Behavioral and Brain Sciences*, *24*, 1143–44.

Hurley, S., & Nudds, M. (2006). *Rational animals?* Oxford: Oxford University Press.

Lurz, R. (2001). How to solve the distinguishability problem. *Brain and Behavioral Sciences*, *24*, 1142–3.

Okamoto-Barth, S., Call, J., & Tomasello, M. (2007). Great ape's understanding of other individuals' line of sight. *Psychological Science*, *18*, 462–468.

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 B*, *362*, 731–744.

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

Sato, A., Kanazawa, S., & Fujita, K. (1997). Perception of object unity in a chimpanzee (*Pan troglodytes*). *Japanese Psychological Research*, *39*, 191–199.

Whiten, A. (1996). When does smart behavior-reading become mind-reading? In P. Carruthers & P. Smith (Eds.), *Theories of theories of mind* (pp. 277–292). Cambridge: Cambridge University Press.