

Inferring Other Minds: Failure of the Argument by Analogy

Daniel J. Povinelli
University of Louisiana at Lafayette

Steve Giambrone
University of Louisiana at Lafayette

Next to the ridicule of denying an evident truth, is that of taking much pains to defend it; and no truth appears to me more evident, than that beasts are endow'd with thought and reason as well as man. The arguments are in this case so obvious, that they never escape the most stupid and ignorant.

When . . . we see other creatures, in millions of instances, perform like actions, and direct them to like ends, all our principles of reason and probability carry us with an invincible force to believe the existence of a like cause. 'Tis needless in my opinion to illustrate this argument by the enumeration of particulars. The smallest attention will supply us with more than are requisite. The resemblance betwixt the actions of animals and those of men is so entire in this respect, the very first action of the first animal we shall please to pitch on, will afford us an incontestable argument for the present doctrine.

This doctrine is as useful as it is obvious. . . . 'Tis from the resemblance of the external actions of animals to

those we ourselves perform that we judge their internal [actions] likewise to resemble ours. . . .

—David Hume, 1739–40, *Treatise*, bk. I, pt. 3, sect. 16

I. INTRODUCTION

Darwin had a problem—or at least he thought he had a problem. He knew that the harshest reaction to his theory of evolution would not concern his claims about the evolution of animals. Although the devoutly religious would find any claim for evolution objectionable, the most severe resistance would center around what his theory suggested about the non-Divine origins of humans. Indeed, Darwin’s meticulous documentation of the similarities and differences among living and extinct species might easily have been brushed aside as merely interesting, if the implications for human origins were not so apparent. After all, by the time *The Origin of Species* was published in 1859, Europeans had already been forced to accept the remarkable morphological and behavioral similarities between monkeys, apes, and humans. The average Victorian, then, had already reached the conclusion that what really separated humans from other animals were not their bodies, but their minds.

Herein lay Darwin’s problem. On the one hand, he knew that if he dwelled on the human case he might merely inflame general anti-evolutionist sentiments. On the other hand, if he let the mind escape evolutionary scrutiny by allowing for its Divine origin (as did the co-discoverer of the principle of natural selection, Alfred Russell Wallace), this might raise a cloud of suspicion around the entire theory. As he later put it: “If no organic being excepting man had possessed any mental power, or if his powers had been of a wholly different nature from those of lower animals, then we should never have been able to convince ourselves that our high faculties had been gradually developed.”¹ Although Darwin perceived this problem early on, in *The Origin* he attempted to sidestep it by making only a few, concluding comments about humans: “Much light,” he vaguely promised, “will be thrown on the origin of man and his history.”² However, even at the time, Darwin was already privately preparing the evidence that humans, too, were a species descended. As he later confessed in *The Descent of Man*, “During many years I collected notes on the origin or descent of man, without any intention of publishing on the subject . . . as I thought that I should thus only add to the prejudices against my views.”³

But by the late 1860s, Darwin felt that the time had come to outline the evidence that every aspect of humans, including the most seemingly Divine

aspects of our minds, had been produced through the action of natural selection. In Chapters 3 and 4 of *The Descent of Man*, Darwin laid out the case for believing that the difference between the minds of humans and other animals was “certainly one of degree and not kind.”⁴ “My object . . .” Darwin wrote, “is to shew [*sic*] that there is no fundamental difference between man and the higher mammals in their mental faculties. . . . With respect to animals very low on the scale, I shall give some additional facts . . . shewing [*sic*] that their mental powers are much higher than might have been expected.”⁵ In the ensuing pages, he offered an impressive (and often amusing) array of anecdotes to illustrate the remarkable intellectual abilities of animals. Monkeys “revenge[d] themselves,”⁶ dogs played “practical jokes,”⁷ and in general animals were “constantly seen to pause, deliberate, and resolve.”⁸

Darwin’s method for comparing the mental abilities was admittedly loose: “As no classification of the mental powers has been universally accepted. . . . I will select those facts which have struck me most, with the hope that they may produce some effect on the reader.”⁹ But underlying his approach was an unquestioned assumption that where humans and other animals exhibited similar behaviors, so too they shared similar mental faculties. And, naturally, these were precisely the mental faculties that were known to accompany the human actions. Darwin’s approach was soon taken up by John George Romanes. Romanes’s goal was to formalize Darwin’s proposal that mental evolution followed the same general principles as organic evolution, and he did so in two major treatises, *Animal Intelligence* (1882) and *Mental Evolution in Animals* (1883).

Romanes’s ambition was no less than to establish a new scientific discipline—comparative psychology. Drawing an analogy to comparative anatomy, he argued that “just as anatomists aim at a scientific comparison of the bodily structures of organisms, so [comparative psychology] aims at a similar comparison of their mental states.”¹⁰ But anatomists had access to dead bodies that could be pinned and dissected; what comparable substance would comparative psychologists dissect? Understandably, Romanes turned to the only source of material that was available—anecdotal reports of the behavior of animals. Although he apologized for having no other alternative, Romanes relied on these accounts to launch the new science. But, unlike Darwin, Romanes spelled out the logical case for using the spontaneous behavior of animals to infer their mental states: “Starting from what I know of the operations of my own individual mind, and the activities which in my own organism they prompt, I proceed by analogy to infer from the observable activities of other organisms what are the mental operations that underlie them.”¹¹ Thus, comparative psychology was born of an argument by analogy. But Romanes’s method was really just a restatement of Hume’s

doctrine from a century earlier: where humans and animals share similar behavior, so too must they share similar minds.¹² In fact, Hume believed that there could not be “the least suspicion of mistake” in his doctrine, noting that he had provided “rather invincible proof” of its validity.¹³

In the twentieth century, the argument by analogy has come to be better associated with Bertrand Russell’s use of it to counter solipsistic arguments against the existence of other human minds, as opposed to its original formulation for establishing the presence of animal minds.¹⁴ In one sense, this is justified, because Russell’s version of the argument is extended and carefully stated:

From subjective observation [i.e., introspection], I know that A, which is a thought or feeling, causes B, which is a bodily act. . . . I know also that, whenever B is an act of my own body, A is its cause. I now observe an act of the kind B in a body not my own, and I am having no thought or feeling of the kind A. But I still believe on the basis of self-observation, that only A can cause B; I therefore infer that there was an A which caused B, though it was not an A that I could observe. On this ground I infer that other people’s bodies are associated with minds, which resemble mine in proportion as their bodily behavior resembles my own.¹⁵

A moment’s reflection on Russell’s solution to the problem of other human minds reveals that it is merely a special case of the broader solution offered by Hume, Darwin, and Romanes.

In this essay we challenge Hume’s seemingly “invincible” proof by critically examining the basis for the argument by analogy. Unlike advocates of this argument, we claim that the similarity of two species’ spontaneous behaviors is silent with respect to the similarity of at least one kind of mental representation—namely, the representation of mental states. Although our indictment is a general one, we focus on the question of whether the argument by analogy can help us to decide whether other species reason about mental states; that is, whether they have a “theory of mind.”¹⁶ We begin by reviewing recent empirical research on this question, which suggests that the striking behavioral similarities between humans and even their nearest living relatives, chimpanzees, are often accompanied by equally striking psychological dissimilarities (at least with respect to theory of mind). Next, we offer an evolutionary model for understanding how our proposition (one that Hume believed could only appeal to “the most stupid and ignorant”) might, in fact, be true. Finally, we offer a detailed critique of the argument by analogy, and in the process provide theoretical support for our evolutionary model.

II. WHAT YOUNG CHIMPANZEES KNOW ABOUT THE MIND: THE CASE OF SEEING

We begin our critical consideration of the argument by analogy by examining some of our laboratory's research on chimpanzees' understanding of mental states. In particular, we consider what chimpanzees understand about mental states, or as Premack and Woodruff put it, whether they have a theory of mind.¹⁷ Humans clearly exhibit a pervasive and automatic tendency to conceptualize the behavior of the self and others in terms of such mentalistic constructs as desires, thoughts, intentions, knowledge, beliefs, and emotions. But are we alone in this capacity?

Although Premack and Woodruff¹⁸ presented some evidence that chimpanzees might reason about internal mental states, more recent research suggests a very different answer to this question.¹⁹ Furthermore, it has become increasingly clear that "theory of mind" indexes a wide range of psychological territory that is far too complex to be addressed by any single experiment. As a partial remedy to this approach, a number of years ago we established a set of laboratories whose purpose was to simultaneously compare chimpanzees and young children's understanding of behavior. As part of this effort, we established a colony of seven chimpanzees who were dedicated to a research program designed to assess the nature of their social understanding. For the past nine years we have investigated their understanding of a wide range of mental states, including desires, beliefs, and intentions. However, our most intensive research has focused on determining what (if anything) chimpanzees understand about the mentalistic dimensions of visual perception, or seeing.

A. DEVELOPING AN UNDERSTANDING OF SEEING AS ATTENTION

There is general consensus among developmental psychologists that by about three to four years of age human children have developed an understanding of visual perception as a mentalistic event.²⁰ This research suggests the presence of at least two major developments in young children's understanding of seeing. By two to three years of age, children appear to understand that gaze refers to (or is "about") objects or events in the external world, and by four years of age they additionally appear to realize that visual perception creates internal mental representations in the minds of both self and others. Despite some consensus about these transitions, whether even younger children and infants' interest in the eyes and gaze direction of others demonstrates still earlier understanding of the psychological dimensions of seeing remains controversial.²¹

One phenomenon that has been offered as among the earliest instances of the infant's understanding of "seeing" is gaze-following. The development of gaze-following (or joint visual attention) allows human infants to orient themselves to the same object or event at which someone else is looking.²² Although there is debate about the timing of various facets of gaze-following in human infants, there is general consensus that it emerges in some form by about six months and that by about eighteen months infants are able to (a) follow an adult's gaze into space outside of their own visual field, and (b) reliably track gaze in response to joint movements of the head and eyes, or the movements of the eyes alone.²³ A number of researchers agree that the sophistication of gaze-following in eighteen-month-olds may qualify as evidence that they appreciate the aboutness or referential aspect of the gaze of others.²⁴ Other researchers have shown that eighteen-month-olds appear to use gaze orientation to associate a nonsense word with a novel object and to discover the reference of an affective outburst.²⁵

The ability to extract and use information about the direction of the gaze of another individual is a focal point for theoretical accounts of how and when infants develop an understanding of the mental experiences of others and, in particular, the mental state of attention. For example, some theorists interpret gaze-following as evidence of the infant's awareness of shared attention.²⁶ On this view, the underlying development of an ability to represent the shared attention of self and others allows for the gaze-following response to emerge, along with other behaviors such as proto-declarative pointing. Other theorists interpret gaze-following as reflecting "hard-wired" mechanisms that provide the foundation for the later development of an understanding of shared reference or meaning.²⁷ In any event, as we have seen, by two to three years of age, there is strong evidence that children have come to understand seeing in explicitly mentalistic terms.

B. GAZE-FOLLOWING IN CHIMPANZEES

Several years ago, our laboratory launched a series of studies designed to determine whether chimpanzees develop an understanding of the attentional aspect of seeing. To begin, we sought to determine if chimpanzees follow gaze, and if so, what degree of similarity to human gaze-following it would have.²⁸ The results of these studies revealed that chimpanzees may share many if not all of the aspects of gaze-following exhibited by eighteen-month-old human infants (see fig. 1). Based on the results of these studies, the following conclusions can be drawn about the specificity of gaze-following in chimpanzees: (1) chimpanzees extract specific information about the direction of gaze from others; (2) chimpanzees display the gaze-following response whether it is instantiated by movements of the head and eyes in concert, or simply the eyes alone; (3) chimpanzees will use another's

gaze to visually search into space outside their immediate visual field in response to eye plus head/upper torso movement, eye plus head movement, or just eye movement alone; (4) chimpanzees do not need to witness the shift in another's gaze direction in order to follow it into space outside their immediate visual field; and finally (5) chimpanzees possess at least a tacit understanding of how another's gaze is interrupted by solid, opaque surfaces. Furthermore, several recent reports have confirmed our speculation that gaze-following may be widespread among primate species.²⁹

There are two broadly different ways of interpreting the level of social understanding associated with chimpanzees' gaze-following abilities. First, chimpanzees and other nonhuman primate species (and even human infants) may understand "gaze" not as a projection of attention, but as a directional cue (i.e., a vector away from the eyes and face). It is possible that the ancestors of the modern primates evolved an ability to use the head/eye orientation

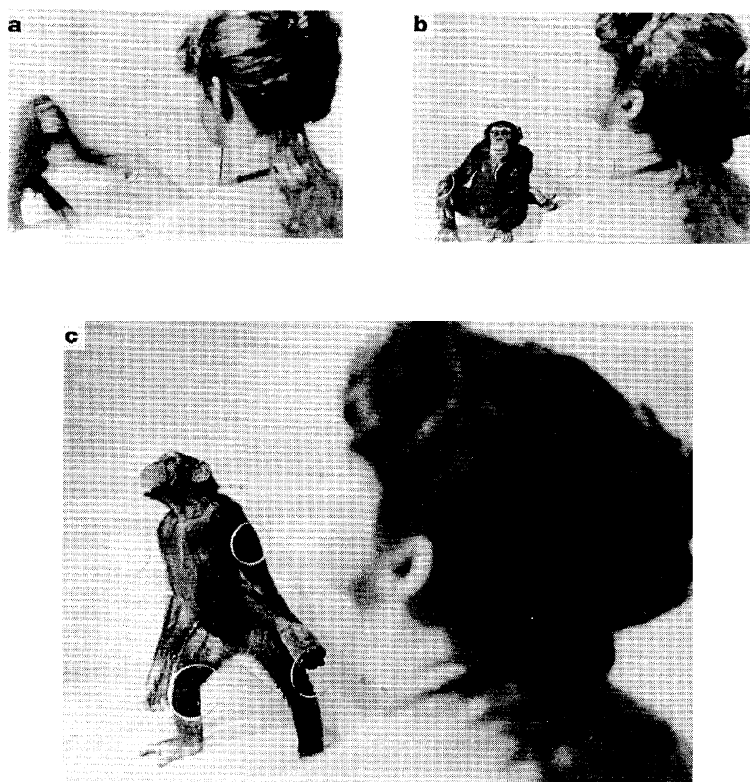


FIG. 1. Gaze-following in a six-year-old chimpanzee in response to movement of eyes + head.

of others to direct their own visual systems along a particular trajectory. Once their visual system encountered something novel, the orienting reflex would ensure that two chimpanzees, for example, would end up attending to the same object or event, without either ape representing the other's internal attentional state. This kind of gaze-following system may have evolved as a commensal relation in which useful information about predators³⁰ or social exchanges³¹ could be gathered at little or no cost to either individual involved.³² A second account is that apes follow gaze because they appreciate its connection to internal attentional states. We will refer to these accounts in the remainder of this essay as the "low-level" and "high-level" accounts, respectively. To summarize, the high-level model stipulates that chimpanzees form concepts about internal mental states (such as attention) and use these concepts to interpret the behavior of others. In contrast, the low-level model supposes that (at best) chimpanzees cogitate about behavioral propensities, not mental states.

If you have ever experienced a chimpanzee following your gaze, you know that it is nearly impossible to resist thinking that the chimpanzee is trying to figure out what you were looking at (that is, it is difficult to resist the high-level explanation of their behavior). However, a careful consideration of the problem reveals that nothing excludes the possibility that they are simply looking where you are looking, without entertaining ideas about your internal attentional state (that is, that the low-level model is correct). In order to intelligently choose between these very different accounts of gaze-following, we need to flesh them out a bit more clearly, and determine if they generate different predictions about how apes (and human infants) might respond in more revealing circumstances.

C. DO CHIMPANZEES UNDERSTAND THE ATTENTIONAL ASPECT OF SEEING?

In an effort to distinguish among these possibilities, over the past several years our laboratory has designed and conducted dozens of experiments involving our seven chimpanzees. In the interest of space, however, we will focus on what we believe are a particularly revealing series of studies concerning whether chimpanzees understand the psychological distinction between someone who can see them and someone who cannot. To address this question, we began by considering the chimpanzee's natural begging gesture—a gesture that this species uses in a number of different communicative contexts, including soliciting allies, requesting food, or reconciliation with others after hostile encounters.³³ Perhaps the most common example of captive chimpanzees' use of this gesture is when they request food from human caretakers. To us, this appeared to be an ideal natural context in which to investigate whether chimpanzees' use of their gesture is mediated by a notion of "seeing."

To begin, we trained our apes to deploy their begging gesture in a standardized routine. The apes entered their test unit in which they were separated from human experimenters by a Plexiglas partition. They then learned to gesture through a hole directly in front of a single, familiar experimenter who was either standing or sitting to their left or right. On each trial that they gestured through the hole directly in front of the experimenter, this person praised them and handed them a food reward. The apes quickly learned to gesture through the correct hole toward the experimenter (see fig. 2).

This training set the stage for examining the animals' reactions to two experimenters, one who could see them (and therefore respond to their gestures), and another who could not. Several conditions which instantiated this

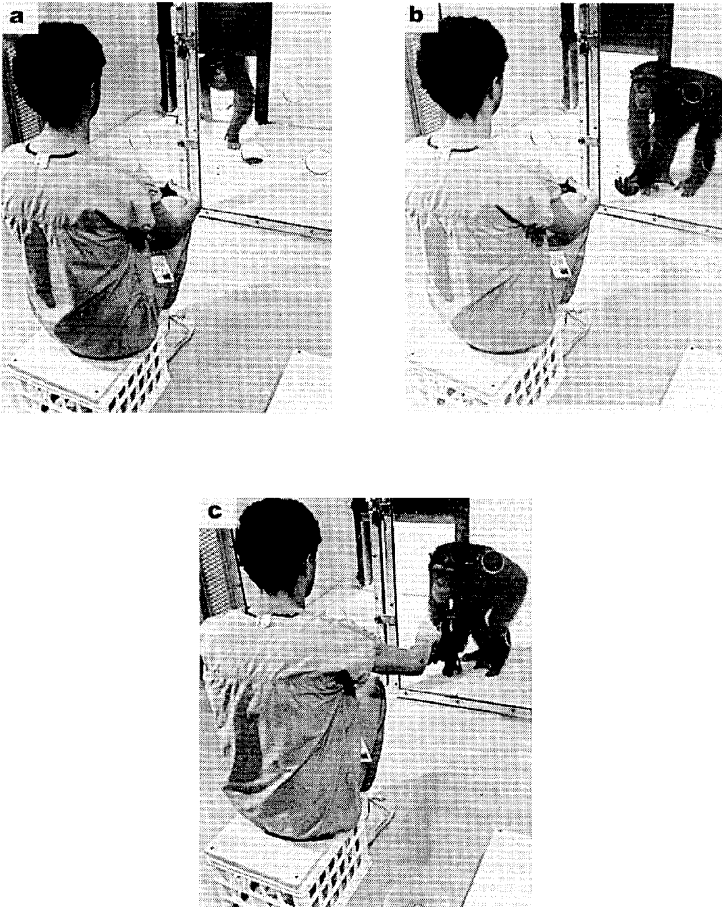


FIG. 2. Chimpanzee entering test unit and using species-typical begging gesture to request food from a familiar experimenter.

notion were created (see fig. 3). Although it is true that we designed these scenarios, we modeled them after the animals' natural play behavior in which they used objects in their enclosure to obstruct their vision. For example, they placed large plastic buckets over their heads, and then moved around their compound until they bumped into something. On a number of occasions we witnessed them lifting the buckets to peek, only to then replace it over their heads and continue to move about blindly. They engaged in the same activities using plastic bowls, burlap sacks, pieces of cardboard, and even their hands. These spontaneous behaviors, then, inspired the conditions of seeing and not seeing depicted in fig. 3: blindfolds, buckets, hands-over-the-eyes, and back-versus-front. The high-level model predicted that the subjects would gesture to the person who could see them, whereas the low-level model generated the counterintuitive prediction that the apes would not prefer either person.

The animals' first reaction to these conditions was to pause. But then, having apparently noticed the novelty of the conditions, they were then just as likely to gesture to the person who could *not* see them as to the person who could. In other words, the chimpanzees displayed no preference for gesturing toward the experimenter who could see them. In contrast, on the trials where only a single person was present, the apes gestured through the hole directly in front of him or her on virtually every trial. Thus, despite

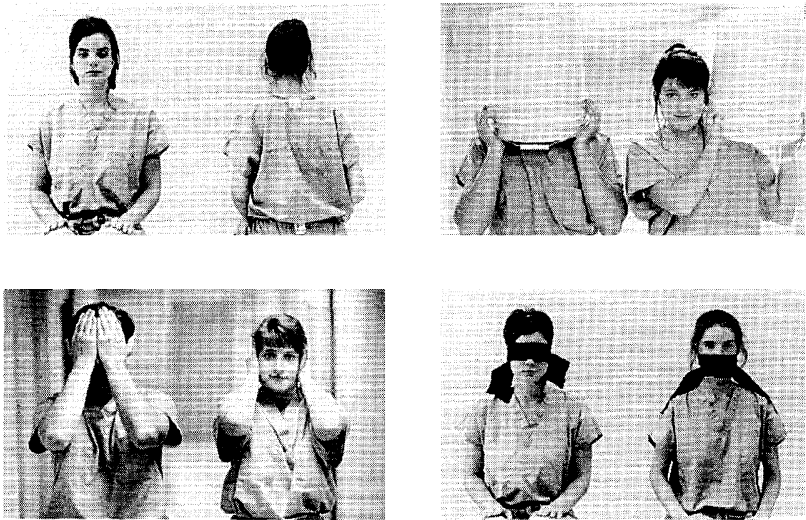


FIG. 3. Conditions used to test chimpanzees for their understanding of seeing/not seeing: (a) blindfolds, (b) buckets, (c) hands-over-eyes, (d) back-versus-front.

their general interest and motivation in the test, when it came to the seeing/not seeing conditions, the animals appeared oblivious to the psychological distinction between the two experimenters.

In contrast to the blindfolds, buckets, and hands-over-the-eyes trials, on back-versus-front trials the animals were correct from trial 1 forward. Why did the apes perform well on this condition but not the others? One possibility was that the back/front condition was the easiest situation in which to recognize the difference between seeing and not seeing. In this way, the high-level model could be sustained. But the low-level model could explain these results as well. Perhaps in reaction to the back/front trials, the apes were simply doing what they had (inadvertently) been taught to do in the training phase—enter the test lab, look for someone who happened to be facing forward, and then gesture in front of him or her. Instead of reasoning about who could see them, perhaps the apes were following a simple procedural rule, e.g., “Gesture in front of a person facing forward.”

We conducted several additional experiments to distinguish between these two revised versions of the low- and high-level models. We reasoned that if the high-level model were correct, then the apes ought to perform well on other conditions that were equally natural. For example, what if the apes encountered two persons facing away from them, but one of them was looking back over the shoulder toward the ape? Recall that the low-level account could explain the apes’ performance on the back/front condition by arguing that they were just following a rule about gesturing to the frontal aspect of a person. But in this looking-over-the-shoulder condition, we eliminated the frontal posture altogether; the only thing distinguishing the two experimenters was that one of their faces was visible. Because no frontal stimulus was present, the low-level account predicted that the subjects would perform randomly on the looking-over-the-shoulder trials—which is exactly what they did.

Several additional experiments using new conditions revealed that with enough experience, the animals began to reliably gesture to the person who could see them. For example, in one condition, an experimenter covered her face with a screen, whereas the other held the screen over her shoulder in such a way that she could still clearly see the subjects. With repeated experience on this test, the apes’ performances gradually improved, until they were beginning to perform at levels exceeding chance. Furthermore, follow-up studies revealed the interesting (although not completely unexpected) fact that our apes’ correct responding had generalized from the screens’ condition to several of the original conditions as well. One possible interpretation of this finding was that the subjects had learned another procedural rule, e.g., “Gesture in front of the person whose face is visible.” However, given that they were now performing correctly on several different conditions, it

seemed possible that they might finally have figured out that we were asking them about *seeing*.

So additional experiments were conducted to distinguish between these possibilities. In one of these experiments, we administered the original set of conditions to the apes (blindfolds, buckets, etc.). The high-level model predicted that the apes would perform excellently on all of them. In contrast, the low-level account predicted excellent performance in all of the conditions except one—blindfolds (where blindfolds covered the eyes of one person versus the mouth of the other). The reason for the prediction was that in this condition an equal amount of the face of each person was visible (see fig. 3). Although it seems obvious to us that only one of these persons can see, if the apes were relying on an arbitrary procedural rule about the presence or absence of a face, they would be left to choose the person whose eyes were covered as often as the person whose mouth was covered—and this is exactly what they did.

Other tests confronted the chimpanzees with two experimenters whose eyes and faces were both clearly visible. However, only one of them was looking in the direction of the ape. The other appeared distracted, with his or her head and gaze directed above and behind the chimpanzee. The high-level model predicted that the subjects would gesture to the experimenter who was looking toward them, whereas the low-level model (because the eyes and faces of both experimenters were visible) predicted that they would not prefer to gesture to one over the other. In response to this condition, the apes followed the distracted experimenter's gaze into the rear corner of the ceiling, but then, as the low-level model had predicted, they were just as likely to gesture to the distracted person as toward the person who was looking in their direction. Thus, although they clearly processed the information about the distracted experimenter's direction of gaze, they did not appear to appreciate its attentional significance.

In a final set of tests, the apes were confronted with a number of conditions that attempted to discern which features of the experimenters were controlling their choices (e.g., fronts, faces, eyes). The results revealed that the apes were relying on a hierarchical rule structure of the following type: (1) "Gesture to a person whose *frontal aspect* is visible; if no frontal aspect is visible (or both fronts are visible), gesture to the person whose *face* is visible; if no face is visible (or both faces are visible), gesture to the person whose *eyes* are visible." For example, in an eyes open/closed condition, the apes initially had no preference for the person whose eyes were open, but after a number of trials their performance improved. However, even here, additional control tests revealed that when the eyes and face were pitted against one another, the face rule was more important (see fig. 4)! In short, through a combination of innate sensitivities to the face, eyes, and overall

posture of others and trial-and-error learning, the subjects learned a simple set of procedural rules. The notion of “seeing,” then, did not appear to be a concept that the apes used in their decision making.

Finally, the chimpanzee’s performances were compared with that of two-, three-, and four-year-old children who were trained to gesture to familiar adult experimenters to request brightly colored stickers. They were then tested on several of the conditions we had used with the apes (screens, hands-over-the-eyes, and back-versus-front). Unlike the apes, however, even the youngest children were correct in most or all of the conditions from their very first trial forward.³⁴

D. GENUINE VERSUS “AS-IF” UNDERSTANDING

Through their experiences on these tests our apes had learned relationships about their own gestures, on the one hand, and the experimenters’ postures, faces, and eyes, on the other. In fact, in some cases they had learned these relationships so well that they could generalize this understanding to new conditions within four trials. This raised the question of how we were to distinguish between the following, very different explanations of their behavior. One possibility was that the apes had not originally understood the concept of visual attention (or just did not apply it in this context), but as a consequence of the feedback they received in our tests, constructed such a concept (or, again, learned to apply it in the context). But a second possibility was that they had simply learned the procedural rules described earlier; indeed, had learned them so well that they were able to behave exactly as if they were reasoning about visual attention.

At this point, one might suppose that the animals had learned something of fairly enduring significance. In order to test this assumption, we conducted two additional sets of tests using these general procedures, first when the subjects were seven years old, and again when the animals were eight to nine years old. At the first longitudinal follow-up, the apes were retested using several of the original conditions³⁵ (e.g., eyes open/closed, screens, back/front). However, despite what they had learned a year earlier, the animals responded at chance levels on all the conditions except, as before, back/front. Indeed, it was only after close to fifty trials of the screens condition that the animals’ performance reached levels significantly above chance. This distinction is even more sobering when one considers that our apes were not idle during this intervening year. On the contrary, they participated in numerous studies that explored their understanding of attention in other ways. One reading of these results is that far from serving to assist them, at some level these other studies may actually have interfered with their performance on this task. Likewise, at eight to nine years of age the apes

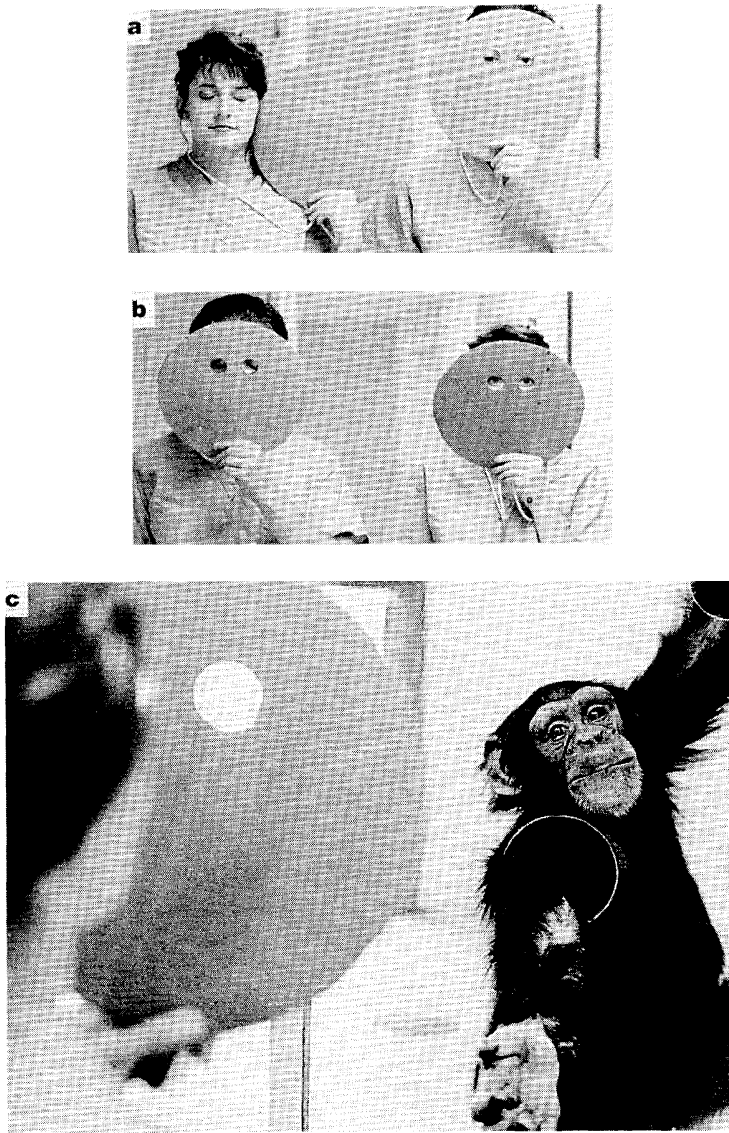


FIG. 4. Two of the conditions used to distinguish rules used by chimpanzees on seeing/not seeing tests: (a) face vs. eyes, (b) eyes vs. no eyes, and (c) Mindy gesturing to the correct experimenter. Subjects preferred the experimenters whose face was visible, over those who could see them (see Povinelli and Eddy, "What Young Chimpanzees Know about Seeing," *Monographs of the Society for Research in Child Development* 61 [2, ser. no. 247], Experiment 14, for details).

initially responded randomly on a number of the original conditions, although within about sixteen trials or so, most of the apes *did* relearn how to solve many of these problems.³⁶ Indeed, some of our chimpanzees learned to solve them all.

This performance raised the question of whether they had finally constructed an understanding of seeing. In order to examine this, we mixed together two conditions that several of our apes appeared to understand well: looking-over-the-shoulder and eyes open/closed. By mixing these conditions together—that is, taking the *correct* posture from the looking-over-the-shoulder condition and combining it with the *incorrect* posture from the eyes open/closed condition (see fig. 5)—we realized that we could once again subject the high- and low-level models to a critical test. After all, if the subjects had developed a genuine understanding of “seeing,” then they ought to perform excellently in this new, mixed condition—after all, it was just another case of seeing/not seeing. On the other hand, if they were still relying on a set of hierarchical rules related to the fronts, faces, and eyes of the experimenters, then not only would they *not* prefer the (correct) person looking over her shoulder, they ought to prefer the (incorrect) person facing forward with eyes closed. This is because the low-level model explicitly stipulated that the forward-facing posture was more important than whether the eyes were open or closed. Remarkably, the apes displayed a statistically significant preference for the incorrect experimenter.³⁷



FIG. 5. Final condition used in Year 3 consisting of eyes closed (incorrect), and looking-over-the-shoulder (correct) options. The subjects preferred the *incorrect* option (see text for details).

Although the low-level model had consistently done a better job of predicting the chimpanzees' behavior on these seeing/not seeing tests, the sophistication of their gaze-following abilities was, nonetheless, impressive—and at a qualitative level difficult to ignore. One possible solution to this tension was to assume that there were fundamental flaws in our seeing/not seeing tests. For example, perhaps the apes had trouble reasoning about the different visual perspectives of two persons at once. In a general effort to evaluate such critiques, the chimpanzees were tested to determine if they would show better evidence of understanding the attentional aspect of visual perception in situations more directly involving their gaze-following abilities.³⁸ In order to do so, these same chimpanzees, as well as three-year-old children, were trained to search under two opaque cups for a hidden reward. In the critical test trials, we kept them ignorant as to the reward's location, but instead let them witness the experimenter gazing either at the correct cup (*at-target*) or above the correct cup at the ceiling of the room (*above-target*) (see fig. 6). If the high-level model of gaze-following were correct, the subjects ought to understand the referential significance of the gaze of the experimenter and therefore choose the correct cup on the *at-target* trials. In contrast, they ought to choose *randomly* between the two cups on the *above-target* trials. This is because organisms with a theory of attention (for example, human children) should interpret the posture of looking at the ceiling as indicating that the person is psychologically (attentionally) disconnected from the cups—conveying no information about the location of the reward.

The results revealed that the three-year-old children selected the cup at which the experimenter was looking on the *at-target* trials, but chose randomly between the two cups on the *above-target* trials. In direct contrast, however, the chimpanzees did not discriminate between the *at-target* and *above-target* trials. Rather, they entered the test unit, moved to the side of the apparatus in front of the experimenter's face, and then chose the cup closest to them. But did the apes even notice the direction of the experimenter's gaze on the *above-target* trials? The results indicated that the subjects followed the experimenter's gaze on 71 percent of the *above-target* trials. Thus, they clearly noticed and followed the experimenter's gaze. However, unlike three-year-old children, the chimpanzees behaved according to the predictions of the low-level model.

Although we have merely scratched the surface of our emerging knowledge of how chimpanzees and humans differ psychologically, the data we have reviewed is sufficient to allow us to illustrate several critical flaws in the argument by analogy. Before we begin, however, we note that rather than accepting these data, one might instead choose to assume either (1) that there are errors in the nature of the experiments or (2) that there are errors

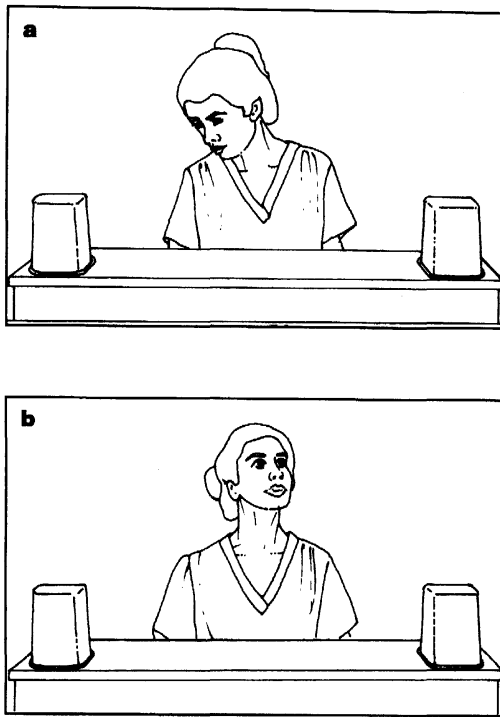


FIG. 6. Conditions used to determine if chimpanzees and three-year-old human children understand the referential aspect of gaze: (a) at-target, (b) above target.

in our interpretation of them. Indeed, there are a number of researchers who still hold out for a chimpanzee theory of mind.³⁹ However, there is a growing body of evidence to the contrary—evidence which indicates that in addition to not understanding the attentional significance of gaze, chimpanzees may also not understand the referential significance of pointing or gazing, the intentions underlying behavior, or the shared intentions underlying joint or cooperative actions.⁴⁰ Quite simply, chimpanzees may not have a theory of mind.

How, then, do we reconcile the apparent absence of theory of mind in chimpanzees with the remarkable sophistication of the spontaneous social maneuvering of chimpanzees and other primates?⁴¹ As we shall see, this reconciliation must begin by abandoning the argument by analogy, and instead exploring the possibility that identical behaviors may be generated and/or attended by different psychological representations. We now turn toward an evolutionary framework within which it is possible to suppose that this is precisely the case.

III. REINTERPRETING BEHAVIOR

In this section, we present our model of the evolution of social complexity and theory of mind. The standard view of the evolution of social understanding is that the evolution of sociality in primates may have driven the emergence of certain kinds of social intelligence.⁴² More recently, this idea has been applied in the context of the emergence of psychological capacities related to theory of mind.⁴³ Indeed, the remarkable complexity of the spontaneous social behavior of chimpanzees (behavior Frans de Waal found fit to label as “Machiavellian”)⁴⁴ gave rise to the idea that theory of mind evolved to cope with problems encountered in group living, and hence the further idea that it might be present in many primate species.⁴⁵ One problem with such a view, of course, is that although sociality is widespread among the primates, the experimental evidence suggests that theory of mind is not.⁴⁶ Indeed, even those who hold out for theory of mind in chimpanzees, doubt the existence of such capacities in other, equally social primates.⁴⁷

For the past several years, we have been exploring an alternative explanation of the evolution of theory of mind. We have advanced the idea that most (if not all) of the sophisticated behavioral patterns shared by humans, chimpanzees, and other primates, evolved long before humans, theory of mind, or even introspection itself.⁴⁸ In short, we have proposed that a novel psychological system for generating and sustaining higher-order representations,⁴⁹ including the representation of other minds,⁵⁰ may have evolved in the human lineage without radically altering our basic behavioral patterns. However, once this psychological system evolved, humans may have found themselves in the position of being able to represent ancient behavioral patterns in explicitly psychological terms, and of using these new representations to modulate an existing behavioral repertoire in order to cope with the newly uncovered mental world (in addition to the directly observable aspects of the social and physical world with which their ancestors had been coping for millions of years). If this view is correct, the most crucial differences between humans and apes are defined by cognitive—not behavioral—innovations. We refer to this as the “reinterpretation” hypothesis.

In short, the reinterpretation hypothesis proposes that the evolution of sophisticated patterns of social behaviors during the course of primate and mammalian evolution was not accompanied by the representation of mental states. Rather, it envisions that the human lineage specialized by evolving a novel representational system that made such representations possible. However, this new system did not replace older ones. Rather, we envision that it was integrated into the earlier psychological systems, so that both now operate in concert. Thus, unlike the social intelligence hypothesis, our proposal does not assume that behavioral complexity and theory of mind

evolved concurrently. Indeed, the reinterpretation hypothesis proposes that the behavioral complexity characteristic of the social primates is mediated by psychological processes which differ little from those present in the earliest mammals.⁵¹ Our model draws a distinction between the emergence of behavioral complexity that is generated through fairly low-level psychological mechanisms, on the one hand, and an understanding of such behaviors in terms of mental states, on the other.

An important part of the reinterpretation hypothesis is that this new psychological system did not suddenly endow the human lineage with myriad new basic behaviors, nor did it interfere with already existing and very useful ones. Rather, this new psychology allowed for the representation of ancient behaviors in explicitly psychological terms. Such representations, in turn, may have given humans the ability to organize existing behaviors in novel and increasingly efficient or productive ways. Nonetheless, because we suppose that these new behavioral patterns were constructed from a set of existing *basic* behavioral elements, it would be difficult to point to any particular behavior and claim that it was not present, or not in principle possible, prior to the evolution of the new system. Indeed, we believe that part of the resistance to the idea that different kinds of representation sustain similar behavior, stems from the fact that many theorists have assumed that the evolution of theory of mind should have resulted in the emergence of numerous novel basic behaviors. In this way, it was expected that we ought to be able to identify a fairly straightforward relationship between the emergence of theory of mind and a class of new social behaviors.⁵² However, as we shall see, the overt behavioral manifestation of theory of mind may be far more subtle than previously suspected.

Some may balk at the view we have just presented, arguing either (1) that the level of social complexity exhibited by primates demands an understanding of mental states, or (2) that evolution of such a novel psychological system must surely have given rise to scores of novel behaviors. Indeed, given our claim that the evolution of the representation of mental states was not accompanied by the emergence of a fundamentally new set of basic behaviors, some might wonder if we are adopting an epiphenomenalist position (in which mental states attend, but play no causal role, in behavior). However, given that we believe that humans (1) possess representations of mental states, and that (2) these representations are part of a causal network responsible for generating many of our behaviors, this cannot be true.

In order to clarify our position, as well as to show how it is internally consistent, we now illustrate the operation of a real system in which the addition of new representational capacities neither changes the basic behavioral repertoire of the system, nor *directly* causes any new behaviors of the system. In detailing this example, we clarify the major points of the re-interpretation hypothesis, including how our position is not a form of epiphenomenalism.

IV. EPIPHENOMENALISM AND THE REINTERPRETATION HYPOTHESIS

Consider a man driving an automobile across a highway.⁵³ Underneath the metal skin of his vehicle is a network of mechanical and electro-mechanical operations that can be described as causing the automobile's motion. From inside the automobile, of course, the driver has no way of directly observing these operations. However, an instrument panel sits before his eyes, keeping track of (representing) various aspects of the automobile, including its speed, the temperature of its engine block, the rpm of its driveshaft, etc. Furthermore, let us imagine that the vehicle is traveling 120 km/hr, and let us focus on the instrument panel's representation of the vehicle's speed. Now, although the speedometer's reading of "120" always co-varies with the reality of the automobile's moving at the speed of 120 km/hr, the current speedometer's representation of this speed is not the cause of the vehicle's current speed.

At this point, one might retort: "Ah, so you are epiphenomenalists after all! Just as you imply the representation of the automobile's speed plays no causal role in generating its motion, so, too, you imply that the mental states that attend and represent our behavior (or the behavior of others) play no causal role in its generation!" However, such a charge could only be sustained if we believed that there were no causal interaction between the representation of the automobile's speed, on the one hand, and the automobile's actual speed, on the other. Clearly, however, the representation of the speedometer can causally interact with the motion of the automobile through feedback loops, without being the cause of the current motion of the vehicle, nor indeed being the immediate cause of any subsequent motion of the vehicle. Of course, we must be careful about the causal relationship between the representation and the speed of the vehicle. First, although the representation is not a cause of the speed that it is representing, it is sometimes an indirect cause of the future speed of the vehicle. For instance, the reading on the speedometer is sometimes used by the driver to modulate the automobile's motion in relation to factors such as his anticipated time of arrival at his destination, his concern for his safety, and local traffic laws. Thus, the driver processes the implications of the representation and adjusts the pressure of his foot upon the accelerator accordingly. Indeed, having done so, he may then check the representation again, continuing this process until the representation matches the representation he desires. In summary, the representation of the automobile's speed that is generated by the speedometer causally interacts with other mechanisms which, in turn, directly generate the motion of the automobile.

Now imagine that our original driver sees another vehicle approaching

and wonders if this second automobile, like his own, contains a speedometer. Furthermore, let us suppose that, based on the fact that the other automobile is behaving the way his does, he concludes that the answer is “yes.” Does this conclusion follow? It may or may not be true, but the behavioral properties of the second vehicle by themselves do not provide good evidence for the conclusion. Even though we have established that speedometers can causally interact with an automobile’s speed, their presence is not necessary for an automobile to move at any particular speed. And so, there would be no good means for the first driver to use the observed spontaneous behavior of the second automobile to determine whether it possessed a speedometer.

Further, imagine following two different automobiles around a city for a full day, one of which has a speedometer, the other of which does not. Both automobiles will exhibit the same set of definable basic behaviors (starting, stopping, turning right or left, accelerating, decelerating, etc.), and thus we doubt that there would be any principled manner of using their overt behavior in *this* circumstance to determine which automobile contains a representation of its speed and which does not. Indeed, and perhaps even more revealing, if we return to our original automobile, and disabled its speedometer, there need not be any qualitative change in its behavior.⁵⁴

At this point, some might concede that the argument by behavioral analogy for the existence of speedometers is indeed weak, but note that this was not the best argument for the inference in the first place. In other words, despite the weakness of the argument by behavioral analogy, there is nonetheless good reason to suppose that the second automobile has a speedometer because it is the same general kind of thing as the first (i.e., that there are other relevant similarities). Thus, except under some rather unusual circumstances (e.g., someone having removed the speedometer, the speedometer not having been installed in the factory, etc.), the general similarity of the two automobiles provides good reason to suppose that, most of the time, the inference to the presence of the speedometer will be correct.

Although this seems like a reasonable argument, let us examine it a bit further to determine if it is helpful for the problem at hand (i.e., inferring second-order mental states in animals). First, we might ask how similar the appearance between the two automobiles need be in order to be confident of the inference. Certainly the same year and make lends itself to the inference. Even the same year will do—indeed, even an automobile seventy years younger may do. But what about automobiles nearer the origin of the invention of the speedometer, or a mixture of modern automobiles and ones that were invented prior to the addition of the speedometer? Here we see that our inference becomes very risky indeed. What we witness from behind the instrument panel of a 1992 Saturn may tell us little (a priori) about the likely

state of affairs inside a 1908 Model-T. Given this, one might wish to shift back to the argument by behavioral analogy, and claim that differences in the behavioral properties of the 1992 Saturn versus the 1908 Model-T can settle the issue. But, of course, we have already shown the weakness of this argument.

There are, we believe, some behaviors that could be used to settle the issue, but, significantly, the behaviors would not (as we have seen) be of a spontaneous, ordinary sort. So what behaviors could be used? We can imagine numerous cases, but take this one. Imagine that we utter the following sounds to both drivers: “Accelerate to 60 km/hr!” The Saturn driver will do so easily by consulting his representational device. But the Model-T driver will flounder. However, note that the driver of the Model-T could learn to perform the feat correctly on command—even if he had no concept of speed at all. We simply let him behave at random (which will naturally include his varying the speed of the vehicle) and reinforce the target behavior until he comes to understand the connection between the sounds “60 km/hour” and a certain combination of kinesthetic and visual experiences.⁵⁵ We propose that our experimental comparisons of chimpanzees and children do exactly the same sort of thing. And, as we have illustrated above, we can get our chimpanzees to act precisely “as-if” they were using second-order representations.

At this point, one may wonder why the speedometer was put into automobiles in the first place. After all, we have claimed that the behavioral properties of automobiles were not obviously affected by its addition. However, note that although the automobile’s *basic* set of behaviors was not enlarged by the addition of the speedometer, many of them became easier to perform (e.g., maintaining a constant speed, estimating times of arrival), while still others became practical for the first time. Furthermore, we suppose that the addition of the speedometer increased the sophistication of the *organization* of the behavior of automobiles. Indeed, we can even imagine that some new constraints were established (e.g., speed limits). Thus, the addition of the speedometer can be seen as a very useful (and, as traffic laws were expanded and enforced, indispensable) part of automobiles. But note that if someone were to doggedly insist on being shown some behavior which the Saturn (with its speedometer) can perform, but the Model-T (which we allege to lack a speedometer) cannot, we would have a very difficult time doing so if we only had recourse to its spontaneous behaviors. This is specifically the demand of those who are particularly convinced by the argument by analogy.⁵⁶ This demand, however, is based on the incorrect assumption that the emergence of novel representational systems must lead to observable behaviors that are not only novel, but also otherwise impossible. In the context of the evolution of second-order intentional states, the reinterpretation hypothesis rectifies this misunderstanding.

We recognize that our automotive illustration has its limitations. But it at least establishes two important points. First, it shows how the emergence of new, functionally useful representational systems may generate new representations that causally interact with behaviors without having been responsible for the emergence of those behaviors to begin with, nor being the *direct* cause of subsequent occurrences of those behaviors. Second, and more importantly, it shows how the addition of these new representations need not be associated with an expansion in the existing set of *basic* behaviors, nor with an expansion in the set of *possible* behaviors.

V. THE ARGUMENT BY ANALOGY: A FORMAL CRITIQUE

So far, our critique of the argument by analogy for the existence of other minds has been indirect. In this section, we first adapt Russell's argument as an argument by analogy for the existence of theory of mind in chimpanzees (which we take to be a fair version of the argument that Hume, Darwin, and Romanes would have made in this case), and then examine it formally.

- P1 I (and other humans) exhibit bodily behavior of type B (i.e., those normally thought to be caused by second-order intentional states).
- P2 Chimpanzees (and other species) exhibit bodily behaviors of type B.
- P3 My own bodily behaviors (and those of other humans) of type B are usually⁵⁷ caused by⁵⁸ my (and other humans') second-order mental states of type A.
- C Therefore, bodily behaviors of type B exhibited by chimpanzees are caused by their second-order mental states of type A; and so, a fortiori, chimpanzees have second-order mental states of type A.

As should be apparent from Sections II and III, we challenge the inductive cogency of this argument.

A. QUESTIONING THE MAIN PREMISE (P3)

First, note that the argument pivots on our confidence in P3, which is supposed to rest upon the testimony of introspection. As Russell noted,⁵⁹ there are actually two claims in this premise, which when adapted to the case of second-order intentional states become the following: (1) we claim to know that second-order mental states of type A cause behavior of type B by "subjective observation," and (2) we claim to know that behavior of type B is only (or at least generally) caused by second-order mental states of type A (again by "self-observation").

With respect to Russell's first claim, all of us are familiar with instances of acting on deliberations about our desires and what behaviors might be effective in securing their objects. For immediate purposes, we accept that at least some of these cases (and perhaps others as well) are instances of higher-order mental states causing our behaviors.

The second claim, however, is dubious. It is not at all clear that "self-observation" tells us that behaviors which are *sometimes* caused by our higher-order mental states are *usually* caused by them. Gaze-following in humans is, no doubt, sometimes caused by the belief that the other person is looking at something, coupled with a desire to know what that something is. But is it usually? Suppose, for example, we are in the middle of a conversation with Joe. During the conversation, he happens to glance behind us, and we follow his gaze without a break in the conversation. Now suppose that we are later asked why we did this. We would, no doubt, reply that we saw our friend looking at something and wanted to know what it was.

There are several ways of thinking about the nature of the faculty (let us call it "introspection") that generates this report—indeed, that generates any such self-reports. On some accounts, introspection is seen as another perceptual faculty; in others, introspection is a non-perceptual, but nonetheless cognitive faculty; in still others, "introspection" is merely a form of internal explanation (story-telling) which is more or less simultaneous with the supposed events of the story; such "story-telling" may either be sheer confabulation or relatively solid theoretical speculation.

Our view of such explanations is somewhat different. We contend that many self-reports about the causes of our behavior are not generated by self-observation at all, but rather by after-the-fact explanations produced according to already-held general beliefs about why we behave the way we do. Thus, we contend that self-reports such as the one in the gaze-following example from above, are not derived from propositional memories of what we had earlier learned by self-observation (i.e., "direct" introspective perception), nor from a current, vivid recollection of the behavior and the mental states that purportedly attended it. Indeed, we think that second- (or higher-) order mental states do not accompany the kinds of behavior in question as often as is commonly supposed. Thus, one's confidence in P3 must be lowered. Similarly, one's confidence in the conclusion needs to be lowered.

Some may react to this by granting that the original premise (P3) was overstated, but that now the conclusion simply needs to be understood as implying that the chimpanzee behaviors in question are caused by second-order intentional states *as much or as little as in humans*.⁶⁰ However, this misses the point. Once one takes seriously the fact that type B behaviors need not be directly caused by second-order intentional states, the impor-

tance of the reinterpretation hypothesis becomes apparent. Rather than dismissing this observation as trivial, the reinterpretation hypothesis elevates the significance of the falsity of P3. It becomes much easier to understand how a set of basic behaviors were in full operation prior to the emergence of the capacity to sustain second-order intentional states. Thus, the fact that many type B behaviors occur in humans without second-order intentional states causing them can be seen as a direct consequence of the fact that human evolution featured the integration of a new psychological system into a more ancient one. P3 may not simply be false, it may be false because of the way in which second-order intentional states evolved.

B. QUESTIONING THE INFERENCE

As we have just seen, the falsity of P3 can be understood in the context of the reinterpretation hypothesis. But as a logical point, even if we were to grant P3, the inference for second-order mental states from humans to chimpanzees would still not be a strong one. The strength of arguments by analogy depends upon the extent of the similarities, as well as the relevance of the similarities. Likewise, of course, such arguments are weakened by relevant dissimilarities.

The experimental results that we detailed above provide compelling evidence that humans and chimpanzees behave in strikingly different ways in controlled laboratory tests (that is, that they are relevantly dissimilar). First, their initial behavior in experimental contexts designed to test for the presence of certain reasoning processes reveals that they initially behave very differently from even young children. Second, the chimpanzees' slow rate of acquisition and the limited nature of retention provide further evidence of this dissimilarity. In addition to the results of these laboratory tests, it is important to note that there are striking dissimilarities between humans and chimpanzees in their natural social behaviors and material culture⁶¹—dissimilarities that have been glossed over due to an insistence on evolutionary continuity which began (as we have seen) with Darwin.⁶² Finally, of course, there are obvious dissimilarities in brain size and structure. For example, the human lineage has expanded the overall size of the brain three-fold since its divergence from the common ancestor of the great ape/human clade.⁶³ Chimpanzees, in contrast, may not have expanded their brain size whatsoever. In summary, there are significant, highly relevant dissimilarities between humans and chimpanzees which preclude the argument by analogy's being inductively strong.

Researchers have long known about some of these dissimilarities between humans and chimpanzees (and, indeed, other animals as well) in social behavior, material culture, and brain size, but have assumed that they were trumped by overwhelming similarities. But we suspect that the real

reason for downplaying the dissimilarities has been somewhat different. Rather than relying exclusively on the argument by analogy, many theorists have also (at least implicitly) relied on an argument to the best explanation.⁶⁴ These theorists (and we include our past selves in this camp) might well maintain that, although it is persuasive to see animals behaving in the same way as we do, the inference to psychological similarity is made all the more compelling *because there is no better explanation* for their behavior—especially given the assumption that the best explanation would have to be consistent with the idea of intellectual continuity in the evolution of mental functioning.⁶⁵

Thus, a modern defender of the argument by analogy might reason as follows: “Granted, Russell called it an argument by ‘analogy,’ but as he himself pointed out, he used the label ‘analogy’ rather loosely. But regardless of his formalism, the most persuasive version of the argument is this: Given that humans exhibit bodily act B, which is caused by mental state A, and given that animals exhibit bodily act B, the best hypothesis available is that they also experience mental state A. After all, how else could they be generating such behaviors?” We believe that the reinterpretation hypothesis provides exactly such a better explanation (currently the best explanation), because it accounts not only for the relevant similarities between humans and chimpanzees, but also the relevant dissimilarities—including those new ones uncovered in our laboratory. And, consistent with this hypothesis, models appealing to lower-level mechanisms have consistently done a better job of predicting our chimpanzees’ behavior in crucial experimental situations.

Finally, one might be inclined to concede that there is no simple causal connection between psychological states and the behaviors they accompany, but note that we have not eliminated the possibility that second-order psychological states merely accompany the behaviors in question. We concede that although this is true in the strict sense, a suitably modified version of the argument by analogy (see n. 58) is even weaker than the original, since the relevance of the behavioral similarity to the presumed presence of the mental states consisted in the latter’s supposed causal role in generating the behaviors. If one claims that there is no causal relationship, the similarity of human and chimpanzee behaviors is about as relevant for the conclusion that they have similar mental states as is the fact that both species have feet.⁶⁶

VI. THE END OF THE ARGUMENT BY ANALOGY?

Arguments for mental states in others typically take the form of either an argument by analogy or an argument to the best explanation. We have

shown that neither of these can support a claim for higher-order mental states in chimpanzees. In doing so, we have argued that the best available scientific evidence is strongly consistent with what we conjecture, namely, that chimpanzees do not have second-order mental states at all. If true, and if this conclusion can be generalized to other great apes as well, then an evolutionary analysis would strongly imply that no other species (currently known) except humans possess such states.

But what about other mental states and processes? Until now, we have intentionally restricted our analysis to the case of second-order intentional states. However, it is worth considering whether our analysis has implications for whether animals possess other mental states and processes. One particularly controversial issue concerns the question of consciousness.⁶⁷ If chimpanzees truly do not have second-order intentional states, then they may not be conscious entities at all. After all, without second-order intentional states there may be no reflexive aspect to chimpanzee cognition. Although this view of consciousness is certainly not universally accepted,⁶⁸ the traditional philosophical analysis holds that there is no consciousness without *self*-consciousness.⁶⁹ Another, less-speculative area concerns chimpanzees' understanding of folk physics. Investigations in our laboratory and elsewhere are beginning to reveal striking differences between humans and chimpanzees in this area. For example, although widely misunderstood, Wolfgang Köhler's landmark studies⁷⁰ revealed sobering differences between humans and chimpanzees in their understanding of even the basic aspects of the mechanics of the world. More recent research in our laboratory has confirmed and extended these findings,⁷¹ and although it is too early to be certain, we believe that this may be symptomatic of an inability to reason about unobservable phenomena. One strong interpretation of such data is that chimpanzees may have no genuine notion of "cause" at all. In any event, our general point is that the argument by analogy has outlived its usefulness in this context. Indeed, rather than relying on any very general argument about the sorts of minds and kinds of knowledge that other animals may possess, researchers ought to be testing for the presence or absence of specific capacities using a rigorous experimental approach that embraces the method of multiple working hypotheses.⁷²

Finally, we recognize that, as is the case with humans, some chimpanzees may display abilities that others do not (e.g., self-recognition in mirrors).⁷³ Thus, although there is no compelling evidence to support this view, it is certainly possible that some chimpanzees possess second-order intentional states, whereas most do not. An extreme view of this notion is that exposing chimpanzees to the company and culture of humans (i.e., "enculturating" them), can lead them to construct second-order intentional states.⁷⁴ While we have vigorously pursued the method of strong inference

in investigating our chimpanzees' understandings, we recognize that they represent only one type of rearing condition. Although our apes have been raised in extensive contact with humans, there are certainly other apes which have received far more intensive exposure to the social and material culture of humans.⁷⁵ However, claims that such apes possess psychological abilities not possessed by other apes are simply not well substantiated.⁷⁶ But the fact remains that even in humans very little is known about how mental states develop in the first place. Thus, the question of whether the chimpanzee's normal ontogenetic pathways can be altered substantially by exposing them to human culture should be addressed in rigorous, well-controlled studies. Though such studies will be difficult and costly, the alternatives are simply not scientifically acceptable.

NOTES

The research described in this article was supported by NSF Young Investigator Award SBR-8458111 to D. J. Povinelli, and would not have been possible without the talents of numerous staff and students who have assisted in the training and testing of the chimpanzees over the past five years, including Anthony Rideaux, James Reaux, Donna Bierschwale, Laura Theall, and Timothy Eddy. We thank Gerald Massey for drawing our attention to the cited passages by David Hume. Photographs are by Donna T. Bierschwale and Corey G. Porché.

1. Charles Darwin, *The Descent of Man* (1871; reprint, New York: Modern Library, 1982), 445.
2. Charles Darwin, *The Origin of Species* (1859; reprint of 6th ed., New York: Modern Library, 1982), 373.
3. Darwin, *The Descent*, 389.
4. *Ibid.*, 494.
5. *Ibid.*, 446.
6. *Ibid.*, 449.
7. *Ibid.*, 450.
8. *Ibid.*, 453. Indeed, less than halfway through chap. 3, Darwin felt that his essential task was complete: "It has, I think, now been shewn [*sic*] that man and the higher animals, especially the Primates . . . [a]ll have the same senses, intuitions, and sensations,—similar passions, affections, and emotions, even the more complex ones, such as jealousy, suspicion, emulation, gratitude, and magnanimity; they practice deceit and are revengeful; they are sometimes susceptible to ridicule, and even have a sense of humour; they feel wonder and curiosity; they possess the same faculties of imitation, attention, deliberation, choice, memory, imagination, the association of ideas, and reason, though in very different degrees. The individuals of the same species graduate in intellect from absolute imbecility to high excellence. They are also liable to insanity, though far less often than in the case of man" (Darwin, *The Descent*, 456–57).
9. *Ibid.*, 446.
10. George John Romanes, *Mental Evolution in Animals* (New York: Appleton, 1883), 5.
11. George John Romanes, *Animal Intelligence* (London: Kegan Paul, 1882), 1–2.
12. David Hume, *A Treatise of Human Nature. 2 Vols.*, ed. L. A. Selby-Bigge (1739–40; reprint, Oxford: Clarendon, 1978), 176–79.

13. *Ibid.*, 178.
14. Bertrand Russell, *Human Knowledge: Its Scope and Limits* (London: Unwin Hyman, 1948).
15. *Ibid.*, 489. Russell states the underlying inferential principle as the postulate of analogy: "If, whenever we can observe whether A and B are present or absent, we find that every case of B has an A as a causal antecedent, then it is probable that most B's have A's as causal antecedents, even in cases where observation does not enable us to know whether A is present or not. . . . This postulate, if accepted, justifies the inference to other minds" (*ibid.*, 486).
16. D. Premack and G. Woodruff, "Does the Chimpanzee Have a Theory of Mind?" *Behavioral and Brain Sciences* 1 (1978): 515–26. Premack and Woodruff use the term "theory of mind" to refer to the ability to conceptualize behavior in terms of unobservable mental states. "A system of inferences of this kind," they noted, "is properly regarded as a theory because such states are not directly observable, and the system can be used to make predictions about the behavior of others" (515). In current philosophical discussions, having a "theory of mind" corresponds to being a second- (or higher-) order intentional system. See, e.g., D. Dennett, "Intentional Systems in Cognitive Ethology: The 'Panglossian Paradigm' Defended," *Behavioral and Brain Sciences* 6 (1983): 343–55.
17. Premack and Woodruff, "Does the Chimpanzee."
18. *Ibid.*
19. For discussions of this issue, see D. J. Povinelli, "Reconstructing the Evolution of Mind," *American Psychologist* 48 (1993): 493–509; M. Tomasello, A. C. Kruger, and H. H. Ratner, "Cultural Learning," *Behavioral and Brain Sciences* 16 (1993): 495–552; A. Whiten, "When Does Smart Behaviour-Reading Become Mind-Reading?" in *Theories of Theories of Mind*, ed. P. Carruthers and P. Smith (Cambridge: Cambridge University Press, 1996), 277–92; D. J. Povinelli and T. J. Eddy, "What Young Chimpanzees Know about Seeing," *Monographs of the Society for Research in Child Development* 61 (2, ser. no. 247) (1996); M. Tomasello and J. Call, *Primate Cognition* (Oxford: Oxford University Press, 1997); D. J. Povinelli and C. G. Prince, "When Self Met Other," in *Self-awareness: Its Nature and Development*, ed. M. Ferrari and R. J. Sternberg (New York: Guilford Press, forthcoming).
20. J. D. Lempers, E. R. Flavell, and J. H. Flavell, "The Development in Very Young Children of Tacit Knowledge Concerning Visual Perception," *Genetic Psychology Monographs* 95 (1977): 3–53; J. H. Flavell, B. A. Everett, K. Croft, and E. R. Flavell, "Young Children's Knowledge about Visual Perception: Further Evidence for the Level 1–Level 2 Distinction," *Developmental Psychology* 17 (1981): 99–103; J. H. Flavell, E. R. Flavell, F. L. Green, and S. A. Wilcox, "Young Children's Knowledge about Visual Perception: Effect of Observer's Distance from Target on Perceptual Clarity of Target," *Developmental Psychology* 16 (1980): 10–12; J. H. Flavell, S. G. Shipstead, and K. Croft, "What Young Children Think You See When Their Eyes Are Closed," *Cognition* 8 (1978): 369–87; D. K. O'Neill and A. Gopnik, "Young Children's Ability to Identify the Sources of Their Beliefs," *Developmental Psychology* 27 (1991): 390–97; A. Gopnik and P. Graf, "Knowing How You Know: Young Children's Ability to Identify and Remember the Sources of Their Beliefs," *Child Development* 59 (1988): 1366–71; H. Wimmer, G. J., Hogrefe, and J. Perner, "Children's Understanding of Informational Access as a Source of Knowledge," *Child Development* 59 (1988): 386–96; D. J. Povinelli and S. deBlois, "Young Children's (*Homo sapiens*) Understanding of Knowledge Formation in Themselves and Others," *Journal of Comparative Psychology* 106 (1992): 228–38; T. K. Ruffman and D. R. Olson, "Children's Ascriptions of Knowledge to Others," *Developmental Psychology* 25 (1989): 601–6.
21. C. Trevarthen and P. Hubley, "Secondary Intersubjectivity: Confidence, Confiders, and Acts of Meaning in the First Year," in *Before Speech: Beginning of Interpersonal Communication*, ed. A. Lock (London: Academic Press, 1978), 183–229; I. Bretherton

- and E. Bates, "The Emergence of Intentional Communication," in *New Directions for Child Development: Vol. 4*, ed. I. Uzgiris (San Francisco, Calif.: Jossey-Bass, 1979); "The Emergence of Intentional communication as a problem-solving strategy in the gorilla," in *Language and Intelligence in Monkeys and Apes: Comparative Developmental Perspectives*, ed. S. T. Parker and K. R. Gibson (Cambridge: Cambridge University Press, 1994), 333–55; J. C. Gomez, "Visual Behavior as a Window for Reading the Minds of Others," *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading*, ed. A. Whiten (Oxford: Blackwell, 1991), 330–43; W. Phillips, S. Baron-Cohen, and M. Rutter, "The Role of Eye Contact in Goal Detection: Evidence from Normal Infants and Children with Autism or Mental Handicap," *Development and Psychopathology* 4 (1992): 375–83; S. Baron-Cohen, "How to Build a Baby That Can Read Minds: Cognitive Mechanisms in Mindreading," *Current Psychology of Cognition* 13 (1994): 513–52.
22. M. Scaife and J. Bruner, "The Capacity for Joint Visual Attention in the Infant," *Nature* 253 (1975): 265–66.
 23. Ibid. See also G. Butterworth and E. Cochran, "Towards a Mechanism of Joint Visual Attention in Human Infancy," *International Journal of Behavioral Development* 3 (1980): 253–72; G. Butterworth and N. Jarrett, "What Minds Have in Common Is Space: Spatial Mechanisms Serving Joint Visual Attention in Infancy," *British Journal of Developmental Psychology* 9 (1991): 55–72. For a discussion of additional limitations of eighteen-month-olds, see V. Corkum and C. Moore, "Development of Joint Visual Attention in Infants," in *Joint Attention: Its Origins and Role in Development*, ed. C. Moore and P. Dunham (Hillsdale, N.J.: Lawrence Erlbaum 1994, 61–83.
 24. E.g., Corkum and Moore, "Development of Joint Visual Attention."
 25. D. A. Baldwin and L. J. Moses, "Early Understanding of Referential Intent and Attentional Focus: Evidence from Language and Emotion," in *Children's Early Understanding of Mind*, ed. C. Lewis and P. Mitchell (Hillsdale, N.J.: Erlbaum, 1994), 133–56.
 26. E.g., Baron-Cohen, "How to Build a Baby"; S. Baron-Cohen and J. Swettenham, "The Relationship between SAM and TOMM: The 'Lock and Key' versus the 'Tadpole and Frog' Hypotheses," in *Theories of Theories of Mind*, ed. Carruthers and Smith.
 27. Butterworth and Cochran, "Towards a Mechanism"; Butterworth and Jarrett, "What Minds Have in Common"; R. P. Hobson, *Autism and the Development of Mind* (Hillsdale: Erlbaum, 1993); Corkum and Moore, "Development of Joint Visual Attention."
 28. See Experiment 12 in Povinelli and Eddy, "What Young Chimpanzees Know," 85–95; D. J. Povinelli and T. J. Eddy, "Chimpanzees: Joint Visual Attention," *Psychological Science* 7 (1996): 129–35; D. J. Povinelli and T. J. Eddy, "Specificity of Gaze-Following in Young Chimpanzees," *British Journal of Developmental Psychology* 15 (1997): 213–22; D. J. Povinelli, D. T. Bierschwale, and C. G. Čech, "Comprehension of Seeing as a Referential Act in Young Children, but Not Juvenile Chimpanzees," *British Journal of Developmental Psychology* 17 (1999): 37–60.
 29. N. J. Emery, E. N. Lorincz, D. I. Perret, M. W. Oram, and C. I. Baker, "Gaze Following and Joint Attention in Rhesus Monkeys (*Macaca mulatta*).," *Journal of Comparative Psychology* 111 (1997): 286–93; M. Tomasello, J. Call, and B. Hare, "Five Primate Species Follow the Visual Gaze of Conspecifics," *Animal Behavior* 55 (1998): 1063–69.
 30. For example, see C. P. van Schaik, M. A. van Noordwijk, B. Warsono, and E. Sutriyono, "Party Size and Early Detection of Predators in Sumatran Forest Primates," *Primates* 24 (1983): 211–21.
 31. For example, see M. R. A. Chance, "Attention Structure as the Basis of Primate Rank Orders," *Man* 2 (1967): 503–18.
 32. See Povinelli and Eddy, "What Young Chimpanzees Know"; Povinelli and Eddy, "Chimpanzees: Gaze-Following"; Likewise, Butterworth and Jarrett, "What Minds Have in Common," and Corkum and Moore, "Development of Joint Visual Attention," have

- argued that the earliest manifestation of gaze-following in human infants may be controlled by similar low-level psychological phenomena, but that these early mechanisms draw infants into the kinds of social interactions that lead them to construct the notion of "attention."
33. See J. D. Bygott, "Agonistic Behavior, Dominance, and Social Structure in Wild Chimpanzees of the Gombe National Park," in *The Great Apes*, ed. D. A. Hamburg and E. R. McCown (Menlo Park, Calif.: Benjamin/Cummings, 1979), 405–28; F. de Waal, *Chimpanzee Politics: Power and Sex among Apes* (New York: Harper & Row, 1982).
 34. Povinelli and Eddy, "What Young Chimpanzees Know," Experiment 15.
 35. See D. J. Povinelli, "Growing Up Ape," *Monographs of the Society for Research in Child Development* 61 (2, ser. no. 247) (1996).
 36. J. E. Reaux, L. A. Theall, and D. J. Povinelli, "A Longitudinal Investigation of Chimpanzees' Understanding of Visual Perception" *Child Development* 70 (1999): 275–90.
 37. *Ibid.*, see Experiment 4.
 38. See Povinelli, Bierschwale, and Čech, "Comprehension of Seeing."
 39. See contributions to A. E. Russon, K. A. Bard, and S. T. Parker, eds., *Reaching into Thought* (Cambridge: Cambridge University Press, 1996).
 40. Povinelli and Eddy, "What Young Chimpanzees Know; Povinelli, Bierschwale, and Čech, "Comprehension of Seeing"; D. J. Povinelli, H. K. Perilloux, J. E. Reaux, and D. T. Bierschwale, "Young and Juvenile Chimpanzees' Reactions to Intentional versus Accidental and Inadvertent Actions," *Behavioral Processes* 42 (1998): 205–18; D. J. Povinelli, J. E. Reaux, D. T. Bierschwale, A. D. Allain, and B. B. Simon, "Exploitation of Pointing as a Referential Gesture in Young Children, but Not Adolescent Chimpanzees," *Cognitive Development* 12 (1997): 327–65; Reaux, Theall, and Povinelli, *A Longitudinal Investigation*; D. K. O'Neill and D. P. Povinelli, *Chimpanzees, Children, and the Coordination of First and Third Person Intentional Relations* (1997). MS submitted for publication; M. Tomasello, J. Call, and A. Gluckman, "Comprehension of Novel Communicative Signs by Apes and Human Children," *Child Development* 68 (1997): 991–1222; J. Call and M. Tomasello, "The Production and Comprehension of Referential Pointing by Orangutans (*Pongo pygmaeus*)," *Journal of Comparative Psychology* 108 (1994): 307–17. Recently, Michael Tomasello and his colleagues have argued that "enculturated" chimpanzees may develop at least some aspects of theory of mind (especially those related to joint attention) as the results of the experiences with humans. See M. Tomasello, "The Power of Culture: Evidence from Apes," *Human Development* 38 (1995): 46–52; M. Tomasello, "Chimpanzee Social Cognition," *Monographs of the Society for Research in Child Development* 61 (161–73, ser. no. 247) (1996). Although this remains an open question, in our view the evidence in support of such claims is at present very weak. For a critique, see Povinelli, "Growing Up Ape."
 41. E. W. Menzel Jr. "A Group of Young Chimpanzees in a One-Acre Field," in *Behavior of Non-human Primates: Modern Research Trends*, ed. A. Schrier and F. Stollnitz (New York: Academic Press, 1974), 83–153; J. Goodall, *The Chimpanzees of Gombe: Patterns of Behavior* (Cambridge, Mass.: Belknap, Harvard University Press, 1986); de Waal, *Chimpanzee Politics*; F. de Waal, "Deception in the Natural Communication of Chimpanzees," in *Deception: Perspectives on Human and Nonhuman Deceit*, ed. R. W. Mitchell and N. S. Thompson (Albany, N.Y.: SUNY Press, 1986), 221–44; A. Whiten and R. W. Byrne, "Tactical Deception in Primates," *Behavioral and Brain Sciences* 11 (1988): 233–44.
 42. See A. Jolly, "Lemur Social Intelligence and Primate Intelligence," *Science* 153 (1966): 501–6; N. K. Humphrey, "The Social Function of Intellect," in *Growing Points in Ethology*, ed. P. P. G. Bateson and R. A. Hinde (Cambridge: Cambridge University Press, 1976), 303–17.
 43. See G. G. Gallup Jr. "Self-awareness and the Emergence of Mind in Primates," *American Journal of Primatology* 2 (1982): 237–48; G. G. Gallup Jr. "Toward a Comparative

- Psychology of Mind," in *Animal Cognition and Behavior*, ed. R. E. Mellgren (New York: North-Holland, 1983), 473–510; G. G. Gallup Jr. "Do Minds Exist in Species Other Than Our Own?" *Neuroscience and Biobehavioral Reviews* 9 (1985): 631–41.
44. De Waal, *Chimpanzee Politics*.
 45. For example, Whiten and Byrne, "Tactical Deception"; R. W. Byrne, *The Thinking Ape* (Oxford: Oxford University Press, 1995); R. W. Byrne and A. Whiten, *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans* (New York: Oxford University Press, 1988); A. Whiten, ed., *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading* (Cambridge: Basil Blackwell, 1991).
 46. See discussion above; D. J. Povinelli, K. A. Parks, and M. A. Novak, "Do Rhesus Monkeys (*Macaca mulatta*) Attribute Knowledge and Ignorance to Others?" *Journal of Comparative Psychology* 105 (1991): 318–25; D. J. Povinelli, K. A. Parks, and M. A. Novak, "Role Reversal in Rhesus Monkeys, but No Evidence of Empathy," *Animal Behaviour* 44 (1992): 269–81; D. L. Cheney and R. M. Seyfarth, *How Monkeys See the World: Inside the Mind of Another Species* (Chicago and London: University of Chicago Press, 1990); H. Kummer, G. Anzenberger, and C. K. Hemelrijk, "Hiding and Perspective Taking in Long-tailed Macaques (*Macaca fascicularis*)," *Journal of Comparative Psychology* 110 (1996): 97–102; P. S. Silverman, "Can a Pigtail Macaque Learn to Manipulate a Thief?" in *Deception*, ed. Mitchell and Thompson, 151–67.
 47. For example, Capuchin monkeys (*Cebus apella*), fail to understand a cooperative task; see R. Chalmeau, E. Visalberghi, and A. Gallo, *Animal Behavior* 54 (1997): 1215–25; Byrne, *The Thinking Ape*; Whiten, "When Does Smart Behaviour-Reading Become Mind-Reading."
 48. Povinelli, "Growing Up Ape"; Povinelli and Prince, "When Self Met Other."
 49. There are numerous proposals for the nature of the psychological system that could generate and sustain such higher-order representations that are compatible with our view: J. Perner, *Understanding the Representational Mind* (Cambridge, Mass.: MIT Press, 1991); A. Karmiloff-Smith, *Beyond Modularity: A Developmental Perspective on Cognitive Science* (Cambridge, Mass.: MIT Press, 1992); D. Olson and R. Campbell, "Constructing Representations," in *Systems of Representation in Children: Development and Use*, ed. C. Pratt and A. F. Garton (New York: John Wiley & Sons, 1993), 11–26; A. Gopnik and A. Meltzoff, *Words, Thoughts, and Theories* (Cambridge, Mass.: MIT Press, 1997). However, there are interesting parallels between Karmiloff-Smith's developmental account of "representational redescription" and our evolutionary reinterpretation proposal (see Povinelli and Prince, "When Self Met Other").
 50. Modularity theorists may object to our characterization of the emergence of theory of mind. Rather than being a domain-general representational system that manifests itself in different ways (including the representation of other minds), modularity theorists maintain the notion of more domain-specific capacities that evolved to serve specific functions. See J. Fodor, *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge, Mass.: Bradford Books/MIT Press, 1987); A. M. Leslie, "ToMM, ToBY, and Agency: Core Architecture and Domain Specificity," in *Mapping the Mind: Domain Specificity in Cognition and Culture*, ed. L. A. Hirschfeld and S. A. Gelman (New York: Cambridge University Press, 1994), 119–48); S. Baron-Cohen, *Mindblindness: An Essay on Autism and Theory of Mind* (Cambridge, Mass.: MIT Press, 1995). Although our views are somewhat different, our argument is not affected by this debate.
 51. See D. J. Povinelli and J. G. H. Cant, "Arboreal Clambering and the Evolution of Self-conception," *Quarterly Review of Biology* 70 (1995): 393–421; Tomasello and Call, *Primate Cognition*. We are not claiming that there is a complete heterogeneity of psychological abilities among primates (or even mammals in general). Individual lineages may evolve cognitive specializations in various domains to cope with unique challenges posed by their environments. For an excellent case study in microtine rodents, see

- S. J. C. Gaulin, "Evolution of Sex Differences in Spatial Ability," *Yearbook of Physical Anthropology* 35 (1992): 125–51. However, we argue that the social complexity of the primates—as expressed through such behaviors as gaze-following, deception, selective retaliation, reconciliation, and appeasement—may have evolved long before the emergence of the ability to represent other minds.
52. Gallup, "Self-awareness and the Emergence of Mind."
 53. For purposes of simplicity, we speak of "the automobile" and "the driver" separately, when what is really at issue is the automobile/driver system.
 54. We recognize that in real biological systems, a new representational system may be so "installed" that it would be difficult, if not impossible, to disable only it. However, regardless of one's views on the degree of modularity of these psychological systems in question (see n. 50), our general point stands.
 55. Note that in our example, as soon as the feedback is removed, the driver's new ability may quickly deteriorate. Indeed, we sometimes witness such deterioration in the long-term retention of our chimpanzee subjects. The occurrence of such deterioration would seem to depend upon the stability of the mechanisms that they relied upon to judge the target behavior correctly. In our example, the deterioration may be rapid precisely because in the absence of regular feedback, the relevant kinesthetic/visual senses may not be extremely reliable.
 56. An excellent case in point is pointing. Several carefully controlled, programmatic experimental studies have demonstrated that chimpanzees do not mean or comprehend their begging/reaching gestures in a protodeclarative manner. See Povinelli and Eddy, "What Young Chimpanzees Know"; Reaux, Theall, and Povinelli, *A Longitudinal Investigation*; Tomasello, Call & Gluckman, "Comprehension of Novel Communicative Signs"; Povinelli, Reaux, Bierschwale, Allain, and Simon, "Exploitation of Pointing." Nonetheless, several researchers have insisted that they do, relying on little more than the mere production of the gesture in concert with gaze alternation: D. A. Leavens, W. D. Hopkins, and K. A. Bard, "Indexical and Referential Pointing in Chimpanzees (*Pan troglodytes*)," *Journal of Comparative Psychology* 110 (1996): 346–53; M. A. Krause and R. S. Fouts, "Chimpanzee (*Pan troglodytes*) Pointing: Hand Shapes, Accuracy, and the Role of Eye Gaze," *Journal of Comparative Psychology* 111 (1997): 330–36.
 57. Russell claims "only" here (*Human Knowledge*, 449), but we assume that most persons would find such a claim to be very implausible. We weaken this claim to "usually" in order to avoid attacking a straw man.
 58. We assume here an ordinary, robust notion of causation which includes real connection between cause and effect. We later (sect. V.B) deal with a version of the argument by analogy which embraces a Humean notion of causation as merely constant (temporally ordered) conjunction.
 59. Russell, *Human Knowledge*, 489.
 60. Along the same lines, one might recast P3 as claiming that behaviors of type B owe their initial evolutionary and/or ontogenetic appearance in humans to second-order mental states of type A (adjusting the conclusion appropriately). If true, it could be argued that all we have done is to show that after the behaviors become established they can become automatized and occur in the absence of second-order intentional states. However, this changes the entire nature of the argument by analogy. Initially, we were to think that the premises were secure—indeed, that the difficult premise (P3) was given by the testimony of reliable introspection. In the recast argument, P3 becomes a theoretical hypothesis. More importantly, it becomes a theoretical hypothesis that is certainly no better supported than a competing hypothesis—namely, the reinterpretation hypothesis.
 61. Tomasello, Kruger and Ratner, "Cultural Learning"; D. J. Povinelli and L. R. Godfrey, "The Chimpanzee's Mind: How Noble in Reason? How Absent of Ethics?" in *Evolutionary Ethics*, ed. M. Nitecki & D. Nitecki (Albany, N.Y.: SUNY Press, 1993), 227–324.

62. See Povinelli, "Reconstructing the Evolution of Mind"; T. M. Preuss, "The Argument from Animals to Humans in Cognitive Neuroscience, in *The Cognitive Neurosciences*, ed. M. S. Gazzaniga (Cambridge, Mass.: MIT Press, 1995), 1227–41.
63. Review by T. M. Preuss and J. H. Kaas, "Human Brain Evolution," in *Fundamental Neuroscience*, ed. F. E. Bloom, S. C. Landis, J. L. Robert, L. R. Squire, and M. J. Zigmond (San Diego, Calif.: Academic Press, 1999), 1283–1311.
64. Indeed, it may be the case that all arguments by analogy can (or at least should) be reconstructed as arguments to the best explanation.
65. In a different context, Bertrand Russell illustrates both the argument to the best explanation and the idea that any such explanation will preserve intellectual continuity: "[I]t is a highly probable inference that there is nowhere a very wide mental gap. It is, of course, possible that there may be, at certain stages in evolution, elements which are entirely new from the standpoint of analysis, though in their nascent form they have little influence on the behavior and no very marked correlatives in structure. But the hypothesis of continuity in mental development is clearly preferable if no psychological facts make it impossible. We shall find, if I am not mistaken, that there are no facts which refute the hypothesis of mental continuity, and that, on the other hand, this hypothesis affords a useful test of suggested theories as to the nature of mind" (Russell, *Analysis of Mind* [London: George Allen Unwin, 1921], 41). For discussions of how the notion of psychological continuity became integrated into the behavioral and brain sciences, see Povinelli, "Reconstructing the Evolution of Mind"; Preuss, "The Argument from Animals to Humans in Cognitive Neuroscience."
66. A traditional epiphenomenalist or a contemporary supervenience theorist might argue that specific second-order mental states, though causally inert, accompany specific behaviors in humans because both the mental states and the behaviors are "brought about" by a common (presumably neurophysiological) state. They might then conclude that the similarity of a particular chimpanzee behavior to a particular human behavior would be relevant to whether chimpanzees have second-order mental states. Our response is twofold. First, the claim of causal inertness is very speculative, and hence there is no reason to reject the experimental evidence that higher-order mental states do, at least some of the time, play a causative role in human behavior. It then becomes very difficult to explain why they do not play the same role in chimpanzee behavior. Second, this new argument by analogy suffers equally from our main critique of the original argument. We have shown that chimpanzees exhibit behavior (including learning rates and retention) that is relevantly dissimilar to human behavior. So, even the epiphenomenalist/supervenience theorist would be forced to grant that the behaviors are caused by different neural states. And, given that these different neural states may generate different behaviors, they may likewise generate different mental states.
67. For example, Gallup, "Self-awareness and the Emergence of Mind in Primates"; L. Davis, "Self-consciousness in Chimps and Pigeons," *Philosophical Psychology* 2 (1989): 249–59; R. W. Mitchell, "Mental Models of Mirror Self-recognition: Two Theories," *New Ideas in Psychology* 11 (1993): 295–325.
68. See N. Nelkin, *Consciousness and the Origins of Thought* (Cambridge: Cambridge University Press, 1996).
69. For a recent defense of this view, see R. Genarro, *Consciousness and Self-consciousness: A Higher-Order Thought Theory of Consciousness* (Philadelphia: Benjamins North America, 1996).
70. W. Köhler, *The Mentality of Apes*, 2d ed. (New York: Vintage, 1927).
71. D. J. Povinelli, *Folk Physics for Apes* (Oxford: Oxford University Press, 2000).
72. T. C. Chamberlin, "The Method of Multiple Working Hypotheses," *Journal of Geology* 5 (1897): 837–48; see also J. R. Platt, "Strong Inference," *Science* 146 (1964): 347–53.
73. K. B. Swartz and S. Evans, "Not All Chimpanzees Show Self-recognition," *Primates* 32 (1991): 483–96; D. J. Povinelli, A. B. Rulf, K. R. Landau, and D. T. Bierschwale, "Self-

recognition in Chimpanzees (*Pan troglodytes*): Distribution, Ontogeny, and Patterns of Emergence,” *Journal of Comparative Psychology* 107 (1993): 347–72.

74. For example, Tomasello, “The Power of Culture.”

75. For example, W. N. Kellogg, and L. A. Kellogg, *The Ape and the Child: A Study of Environmental Influence on Early Behavior* (New York: McGraw-Hill, 1933); C. Hayes, *The Ape in Our House* (New York: Harper and Brothers, 1951); F. Patterson and E. Linden, *The Education of Koko* (New York: Holt, Rinehart, and Winston, 1981); E. S. Savage-Rumbaugh and R. Lewin, *Kanzi: On the Brink of Humanity* (New York: John Wiley, 1995).

76. Povinelli, “Growing Up Ape.”