Social Understanding in Chimpanzees: New Evidence From a Longitudinal Approach

Daniel J. Povinelli
University of Southwestern Louisiana

Behavior, it would now seem, is not necessarily an open window into the mind. If it were, humanity's quest to end its intellectual and spiritual isolation from the rest of the natural world would have ended with Romanes' (1882) publication of Animal Intelligence over a hundred years ago. For there, Romanes took up Darwin's (1871/1882) claim that there was "no fundamental difference" (p. 445) in the mental functioning of humans and other species. Romanes (1882) proceeded to catalogue examples of animals behaving in ways that suggested not only that the animals had minds, but that they were also reasoning about minds in both themselves and in others. In lieu of being able to pin psychological structures of animals on the dissecting table, Romanes (1883) sought what he saw as the next-best alternative: collect and catalogue their behaviors. These behaviors, Romanes surmised, could serve as behavioral ambassadors of the hidden psychological structures of animals. After all, he reasoned, we know the mental activities that accompany the production of our own behavior, and so, through analogy we ought to be able to infer the mental activities of animals. Similarity in behavior, Romanes concluded, was evidence of similarity in mind.

Of course, Darwin and Romanes were really just putting into practice what Hume (1739–1740/1978) had proclaimed a century earlier: Where humans and animals display similar behavior, so too must they experience similar mental states. "'Tis from the resemblance of the external actions of animals to those we ourselves perform," Hume observed, "that we judge
their internal [actions] likewise to resemble ours . . ." (p. 176). In this chapter, I return to the problem that Hume\(^1\) claimed to have solved, namely, how to assess the mental states that accompany similar behaviors in different species. I examine several cases from our research program in which chimpanzees and human children exhibit strikingly similar behavior, and ask whether such behavioral similarity justifies an inference of comparable psychological similarity. Despite the confident urgings of Hume, Darwin, and Romanes, I provide a simple, theoretical framework for suspecting that it does not.

**CHIMPANZEE SOCIAL UNDERSTANDING**

Chimpanzees are remarkably social creatures. Even casual observations of their interactions with each other reveal a complexity and subtlety that startles the uninitiated. For example, consider the actions of one of our female chimpanzees, Brandy, who discovers a banana poking out of some straw in a secluded area of her compound. Just as she is about to seize the banana, in the distance she notices the alpha male of the group, Apollo, approaching her. Instead of reaching for the banana, Brandy turns her back on it. Apollo greets her with a chimpanzee hug, glances around, and then departs. Once he is out of sight, Brandy removes the banana from its hiding place and consumes it, quietly food-barking in pleasure. Menzel (1974) described the microgenesis of such behaviors many years ago during a landmark set of studies of young chimpanzees, and in the ensuing decades many primatologists have urged us to consider such episodes as evidence that chimpanzees see one another as more than just hairy collections of limbs, torsos, and facial expressions moving through space. Surely, they might insist, Brandy sees Apollo as a subjective being, alert with his own first-person experiences of attending, wanting, knowing, and believing. Indeed, her reactions would seem to provide prima facie evidence that she not only understands that Apollo possesses most or all of these mental states, but that she can manipulate his mental states through her actions. With Hume's two-century-old proclamation still echoing, we are encouraged to believe that it could not be otherwise. Similarity in behavior guarantees similarity in mind.

I, too, was once convinced by such observations that chimpanzees possess a mentalistic understanding of others very similar to our own. Indeed, the

---

\(^1\)Russell (1948) is perhaps better known for his exposition of the argument by analogy, although he developed it not to justify our belief in the minds of other species, but to justify our belief in the minds of other humans. The logical structure of his argument, however, parallels Hume's. Through introspection we can know that our own mental state of type A caused our own bodily act of type B, and therefore if we see another person engage in a bodily act B, we can reasonably infer the presence of mental state A.
description of chimpanzee social maneuvering in de Waal's (1982) *Chimpanzee Politics* was enough to convince even the skeptics that chimpanzees navigated their way through a complex social milieu by reasoning about the mental lives of those around them (see also de Waal, 1986; Goodall, 1986). Soon, reports surfaced of various species of monkeys acting in deceptive ways, and by 1988 Whiten and Byrne had published a report in the *Behavioral and Brain Sciences* that amounted to a veritable taxonomy of "tactical deception" in primates. Indeed, primates were not alone in this regard, and numerous other taxa ranging from avians to elephants have been shown to manipulate each other’s behavior in some sense or another (see Mitchell & Thompson, 1986).

The flood of species that were soon reported to engage in social deception left me, and others, questioning the utility of spontaneous behaviors (anecdotes) as evidence of animals' appreciation of mental states. But the nature of these critiques were often more polemical than theoretical. Some critiques sought to show how such behaviors could emerge through various forms of simple learning, others admonished primatologists for treating learning as something that only occurred in psychological laboratories, and still others appealed to parsimony to resolve the issue (see commentaries following Whiten & Byrne, 1988). Few, however, directly grappled with the more fundamental philosophical problem that Hume claimed to have solved (for an exception, see Thomas, 1988). Indeed, Hume's solution was a general one. He did not claim that certain kinds of particularly dramatic behavior (such as deception) could be taken as evidence for understanding the mental states of animals; rather, the argument by analogy held that the mental states that *attend* our behavior cause them, and that we can therefore use the presence of a particular behavior to infer the presence of a particular mental state (see also Romanes, 1882, pp. 1–2; Russell, 1948, pp. 482–486). Deception might tempt researchers more than other kinds of social interactions, but as far as Hume's argument was concerned, the hesitancy to extend such attributions into all aspects of animal behavior simply reflected a weakness at the knees.

There were at least two possible reactions to the mounting evidence of deception (and other forms of behavioral similarity) between humans and other primates. On the one hand, it was possible to follow Hume into believing that all of these species were reasoning about mental states (excluding, perhaps, such distantly related species as bees and cockroaches at which our intuitions might balk). But this seemed arbitrary and dissatisfying. On the other hand, it was possible to explain away the impressive natural social interactions of animals by invoking such phrases as "innate dispositions," "hardwired behaviors," "social learning," and "parsimony." Although I recognized that this was problematic, I nonetheless blithely assumed that one day we would discover a simple one-to-one correspon-
dence between the evolutionary emergence of social understanding and some as-of-yet unidentified class of social behaviors. In the meantime, experimental methods were needed to pry nature apart at the seams and discover how the minds of chimpanzees (and other animals) really worked.

But slowly, under the weight of literally dozens of experiments that we have conducted to diagnose the nature of chimpanzee social understanding, I have begun to change my opinion about the relation between laboratory-based assessments of chimpanzee social understanding and their spontaneous social interactions. This has not been completely an empirical conversion. Rather, I have begun to consider the merits of a different view of the relation between the evolution of complex social behavior and the evolution of social understanding. This, in turn, has led me to rethink the relation between the spontaneous social behavior of animals and their behavioral reactions in carefully designed experiments. Finally, this has led me to construct a framework that reconciles the two, and simultaneously demonstrates why using the spontaneous behavior of animals to infer their understanding of mental states is problematic. In short, this framework provides a principled alternative to the argument by analogy in the case of the representation of mental states. This framework suggests that there may be a far more complex and subtle relation between behavior and cognition than Hume, Darwin, or Romanes ever suspected.

So here is the problem that shall occupy me for the remainder of this chapter: A human and a member of some other species engage in a particular behavior to cope with a social situation; the behaviors accomplish the same end result, and the organisms even tailor the behaviors similarly to respond to similar dynamic features of the situation. Given this degree of similarity, are we justified in assuming that the second-order mental states that often accompany the human execution of the behavior are present in the other species? Does similarity in behavior guarantee comparable similarity in psychology? The problem is a general one, but the case of humans and chimpanzees is especially poignant because here the functional and structural dimensions of similarity converge. After all, not only do these two species possess similar behaviors that serve similar functions, but also, because of their close phylogenetic relationship, they look very similar as they execute them.

**WHAT CHIMPANZEEs UNDERSTAND ABOUT SEEING**

The dimensions of behavioral similarity I explore revolve around actions that in humans are often described as “referential”—actions such as pointing and gazing. For a number of years, my colleagues and I have been examining how chimpanzees understand such behaviors. The subjects of most of these
studies have been seven chimpanzees that were raised together from birth in a nursery. Admittedly, as far as apes go, our chimpanzees live a fairly unique lifestyle. They play, groom, fight, nap, and otherwise socialize with each other all day long in their indoor-outdoor compound, and at the same time they are constantly exposed to caretakers, staff, and students who interact and play with them in a variety of ways. Embedded into this spontaneous social life with other chimpanzees and humans, each of them is also tested two or three times a day in a specially designed testing facility that is connected to their compound. This facility allows us to test each ape in turn for 10–20 minutes at a time. While the others play, one of them is transferred into an outside waiting area, which is connected by a shuttle door to an indoor testing room. Typically, this animal waits outside as a trial is set up indoors. When the shuttle door is opened, the animal is free to enter the lab and respond to the task. A Plexiglas panel separates the apes from the humans. This routine of coming out of the group to be tested was established when our apes were quite young, and they have now participated in this routine so many times that it blends seamlessly into the other predictable routines of their day such as eating and sleeping.

Although we have investigated their potential understanding of a wide range of mental states, the most detailed work we have conducted concerns their understanding of the attentional aspect of gaze or “seeing.” Our interest in this topic emerged from several directions, but perhaps no one has given this issue more careful experimental consideration than Flavell and his colleagues who conducted an extensive series of studies investigating the timing of young children’s developing understanding of seeing (e.g., Flavell, Everett, Croft, & Flavell, 1981; Flavell, Shipstead, & Croft, 1978; Lempers, Flavell, & Flavell, 1977). What chimpanzees (and other nonhuman primates) understand about seeing is of central importance to the problem of this chapter, because in some sense or another “attentional structure” is a defining feature of primate social organization. Many years ago, Chance (1967) noted that primate dominance hierarchies revolve around who is paying attention to whom. Social primates seem to constantly monitor the behaviors of others up the dominance hierarchy, apparently computing whom the more dominant animals are approaching, threatening, or even at whom they are looking. Thus, both monitoring the gaze of others and looking into the eyes of others are a part of the social life of many primates, although the nature and consequences of these exchanges range from extremely hostile (as in many species of Old World Monkeys such as macaques; see Perrett et al., 1990), to extremely affiliative (as in chimpanzees; see de Waal, 1989).

We recently provided the first empirical demonstration of gaze-following in nonhuman primates (see Fig. 11.1). Because of the importance that has been placed on the emergence of gaze-following in human infants (see
Baron-Cohen, 1994; Butterworth & Cochran, 1980; Butterworth & Jarrett, 1991; Corkum & Moore, 1994), we conducted a series of studies to examine first, whether our apes follow gaze at all, and second, the level of sophistication of whatever gaze-following system was present (see Povinelli, Bierschwale, & Čech, in press; Povinelli & Eddy, 1996a, 1996c [Experiment 12], 1997). This research has demonstrated a strong similarity in the gaze-following system of humans and chimpanzees. First, chimpanzees display the gaze-following response whether it is instantiated by movements of the head and eyes in concert, or simply the eyes alone. Second, 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. Furthermore, chimpanzees do not even need to witness the shift in another’s gaze direction.
in order to follow it into space outside their immediate visual field. Rather, the gaze-following response can be triggered when chimpanzees encounter another's head and eye direction that is statically oriented above and behind them. Third, chimpanzees extract specific information about the direction of gaze from others (at least when that gaze occurs in concert with head and upper torso movement). Finally, chimpanzees may also possess at least a tacit understanding of how another's gaze is interrupted by solid, opaque surfaces. In general, these findings reveal that chimpanzees follow gaze with a sophistication comparable to that exhibited by 18-month-old human infants (see Butterworth & Cochran, 1980; Butterworth & Jarrett, 1991; Corkum & Moore, 1994). More recently, other researchers have demonstrated at least a general capacity for gaze-following in other nonhuman primates (Emery, Lorincz, Perret, Oram, & Baker, 1997; Tomasello, Call, & Hare, in press).

Although the sophistication of gaze-following in chimpanzees and other nonhuman primates may satisfy the part of us interested in the dynamics of social complexity, the philosopher and psychologist in us remains curious, wanting to know what these phenomena reveal about these species' understanding of attention as an internal mental state. As we have seen, Hume would have considered the problem answered as soon as we described the behavior: To the extent that humans reason about attention in such situations, so do other organisms who exhibit the behavior. But in truth there are at least two very different possibilities. First, chimpanzees who follow gaze may do so because they appreciate the underlying attentional states. On the other hand, it is possible that chimpanzees are merely looking where others are looking, without entertaining the idea that the gaze of the other organism is connected to internal visual or attentional experiences (see Povinelli & Eddy, 1996a). Although I run the risk of being misunderstood, there is a need to have succinct labels for these alternative accounts and so I hereafter refer to them as the low- and high-level models, respectively. To summarize, the high-level model stipulates that chimpanzees form concepts about internal mental states (such as attention) and use these concepts to assist them in interpreting the behaviors that unfold around them. In contrast, the low-level model stipulates that chimpanzees reason about behavioral propensities, not mental states.

DO CHIMPANZEES KNOW THAT OTHERS SEE?
A LONGITUDINAL PROJECT

We have addressed the problem of what chimpanzees understand about seeing from a number of different perspectives, but in the interests of space, I focus on one particular procedure that we have used to diagnose our animals' understanding of seeing over a period spanning their late
Year 1: Seeing/Not Seeing
(age 5-6 years)

FIG. 11.3. (a) Conditions used to test chimpanzees for their understanding of seeing/not seeing (blindfolds, buckets, hands-over-eyes, back/front). (b) Results of first four trials of blindfolds, buckets, hands-over-eyes, and back/front conditions. The dotted line indicates level of performance expected by chance responding. **p < .01.
another who could not? The high-level model predicted that the subjects would gesture to the person who could see them from the very first trial forward; the low-level model predicted that the subjects would initially respond at random, and only gradually learn to choose the correct person.

Figure 11.3b presents the results of these initial studies. As can be seen, the subjects responded according to the low-level model in three of the four conditions (blindfolds, buckets, hands-over-the-eyes). The animals entered the test unit, measurably paused, and then gestured to one of the experimenters. However, much to our surprise, they were just as likely to gesture to the person who could not possibly see them, as to the one who could! In direct contrast, the animals performed excellently from trial 1 forward on the back/front condition. Here, the animals had no trouble gesturing to the person facing forward. Two possible explanations for this discrepancy presented themselves. One was that the back/front condition was simply the most obvious case of seeing and not seeing—the one easiest for the apes to recognize. However, a more mundane possibility presented itself as well. The apes might have been gesturing to the person facing forward because this was precisely what we had taught them to do in the training phase: Enter the test unit, approach a person who happens to be facing forward, gesture in front of him or her, and get a reward.

We attempted to distinguish between these possibilities by confronting the apes with another, equally natural posture that they experience with us and each other daily: someone looking over the shoulder toward them (see Fig. 11.4a). This allowed a precise test of the high- and low-level explanations of the back/front results from the first experiment. In this new condition the general frontal aspect of both experimenters was not visible, but one of the two could clearly see the animal. We administered this condition, along with the back/front condition as a control, using our standard probe trial techniques. Again, the animals entered the testing unit, paused in reaction to the novel postures, but then proceeded to gesture to the person looking away from them as often as to the person looking over the shoulder in their direction (see Fig. 11.4b).

We conducted an extended series of studies which followed up on these surprising results. We created several new conditions for these studies including one involving someone holding a screen in front of the face versus someone else holding one over the shoulder (see Fig. 11.5a). Despite an extended amount of initial experience with the screens in the context of play, the animals initially performed randomly. Slowly, however, as we continued to administer these screen trials, their performances began to improve, until finally they were performing at levels exceeding chance (see Fig. 11.5b). Our two models explained these results in very different ways. The high-level model could explain these results by arguing that the subjects had finally learned to select the experimenter who could see them.
Year 1: Looking-Over-Shoulder
(age 5-6 years)

FIG. 11.4. (a) Looking-over-the-shoulder condition. (b) Results of first four trials of looking-over-the-shoulder condition. The dotted line indicates level of performance expected by chance responding.
FIG. 11.5. (a) Screens condition. (b) Results of screens condition with repeated experience. The dotted line indicates level of performance expected by chance responding.
In contrast, the low-level model argued that the subjects had simply learned a rule such as, "Gesture in front of the person whose face is visible."

In order to test these competing interpretations we readministered all of the conditions to the subjects. Both models converged in that they predicted that the apes would perform well in the buckets, hands-over-the-eyes, and back/front conditions, but they differed in their expectations about the blindfolds condition. The high-level model expected that the apes would do well on this condition because it was another particularly salient example of seeing/not seeing. But if the low-level model were correct, and the subjects were merely selecting the experimenter whose face was visible, they ought to continue to respond randomly on the blindfolds condition; after all, in this condition an equal amount of the two faces were visible (see Fig. 11.3). And, in full support of the low-level model, this is exactly what our apes did.

Finally, we conducted several additional tests contrasting predictions of the low- and high-level models (such as one involving a distracted experimenter versus an attending one, someone with eyes open versus eyes closed, and several conditions designed to examine the relative importance of the eyes versus the face). As before, the low-level model continued to generate the most accurate predictions concerning the animals’ responses. However, even in these new conditions the animals showed evidence of learning. Indeed, by the end of the initial 14 experiments we conducted, the subjects were able to learn new conditions within 3–4 trials involving differential feedback. Indeed, several of the final experiments provided confirmation for our idea that the apes were relying on a series of hierarchical rules related to the front, face, and eyes of the experimenters. First, they sought to approach and gesture to someone facing forward. Clearly, the back/front condition easily allowed them to satisfy this rule. On the other hand, if this were not possible (e.g., if both persons were facing forward), then the apes moved on to the next most important rule which was to gesture to the person whose face was visible. In the case of buckets and screens, this rule was easily satisfied. Of least importance was the eyes rule, and indeed, at this point there was only limited evidence that they were even using it. They did perform well on the blindfolds condition (which could be solved by the eyes, but not the face rule), but they showed much weaker evidence of understanding the eyes open/closed condition. Nonetheless, by the end of these studies our seven young chimpanzees were behaving exactly as if they understood something about seeing as a mental event. They would approach two familiar caretakers, look carefully at each one, and then gesture in front of the person who could see them. Indeed, by the end of these experiments, our apes were performing in much the same manner as the 2-, 3-, and 4-year-old children that we tested on these same procedures had performed on their very first trials (see Povinelli & Eddy, 1996c, Experiment 15).
11. SOCIAL UNDERSTANDING IN CHIMPANZEES

We have now arrived at the heart of the problem I promised to address in this chapter: Does the fact that two species engage in similar behavior imply similar accompanying mental states? And does it matter how the behavior arises? For example, is there a psychological distinction between a 2- or 3-year-old child who arrives at our laboratory, participates in our seeing/not seeing test, and performs perfectly from trial 1 forward, and our chimpanzees, who, after months of differential feedback, finally learned to do likewise? It is important not to trivialize what our chimpanzees had learned. Although our investigation of their understanding of seeing was best predicted by the low-level model, by the end of these tests the apes were using the direction of the face and the presence or absence of the eyes as the bases for their choices. Although the low-level model explained their use of these features in terms of a psychological system that reasons about physical postures, here the low- and high-level models converge. After all, even though the high-level model envisions the operation of a psychological system that interprets these postures in terms of underlying mental states, even this system must make the relevant postural discriminations in the first place. Another way of thinking about this is to realize that whatever the underlying differences in interpretation, our chimpanzees and 2- to 3-year-old children were attending to the same physical stimuli when they were making their decisions.

Perhaps the problem is best stated from the point of view of the children. Prior to visiting our testing center, these children have had numerous semi-structured experiences of seeing/not seeing (in the context of playing with parents and peers). Thus, long before participating in our tests, the children have been confronted with games, and even “real” social experiences, involving the distinctions between seeing and not seeing. Although their reactions to these situations were not yoked to receiving or not receiving a sticker, surely the range of their responses were linked to a range of consequences from their parents and peers. The question I raise is simple. How do these kinds of experiences received by children compare to the kinds of experiences received by our chimpanzees in testing situations? And even more directly, what can such comparisons tell us about the psychological structures that cause and/or accompany the final behavioral outcomes?

At this juncture, two separate arguments present themselves. On the one hand, it is possible to question whether the 2- to 3-year-old child’s performance on our task really reflects an understanding of seeing/attention in the first place. After all, our tests with the children were not nearly as extensive as those with the apes, primarily because even the youngest children performed well from trial 1 forward. However, those initial tests consisted of the conditions that could be solved by the face rule. Perhaps a low-level model of their behavior would better predict their reactions on
more complicated tests. Although possible, there are independent confirmations of the 2- to 3-year-old's understanding of seeing/attention (see Baldwin, 1991, 1993; Gopnik, Meltzoff, & Esterly, 1995; Lempers, Flavell, & Flavell, 1977). So, perhaps too hastily, I move to the second argument, which is of more central importance to this chapter. Given the amount of experience that children have with explicitly created instances of seeing/not seeing by their parents and siblings, how do we know that it is not precisely this experience that allows the child to create the idea of visual attention in the first place? If so, then perhaps our apes' final trials on our tests are more comparable to the children's very first trials. In other words, perhaps chimpanzees simply need sufficient experience to allow them to construct a concept of seeing-as-attention. Thus, although the low-level model best predicted our apes' behavior at each critical testing juncture, this might have been the case only because our tests were chasing the apes' concurrent construction of the idea of attention.

Extreme modularity theorists will object immediately, claiming that the insularized nature of social understanding excludes this possibility. Yes, they admit, children have such seeing/not seeing experiences in play, but these experiences are incidental to the ontogenetic unfolding of their understanding of mental states in others. In this view, their knowledge of the mental state of attention matures biologically; it is not constructed through experience. Thus, only the apes' initial trials are diagnostic; what they learn through trial and error is not relevant to the question of cognitive development. Apes either naturally and spontaneously develop an understanding of seeing, or they do not; changes in task performance that occur as the result of feedback should be assigned to an "as-if" category of understanding.

But such objections derive most of their force through empirically underdetermined theoretical positions. Although there are good reasons for considering stronger and weaker versions of modularity in the development of cognitive structures, there is nothing about our knowledge of cognitive development that forces the modularity view. Indeed, several theorists have argued that experience does play a crucial role in cognitive development. For example, Premack (1988) argued that training chimpanzees to use a symbol for same/different judgments about objects, altered the natural state of the species' cognitive structures in such a manner that they could now engage in abstract analogical reasoning. More recently, Tomasello (1995) has resurrected the argument that exposure to human culture dramatically affects the cognitive development of apes—in this case, the apes' understanding of joint attention (see Tomasello, chap. 4, this volume). If opinions matter, I certainly favor accounts of cognitive development that emphasize the complex epigenetic interactions that occur during development. But in the final analysis, the analogy between the construction of bodily structures (including the brain) on the one hand, and the construction of concepts
related to mental states, on the other hand, remains unclear. In fact, the ambiguity of this analogy manifests itself in the context of interpreting the results of our seeing/not seeing experiments. For example, some have interpreted our results as evidence that apes do not understand seeing, whereas others have concluded that our apes' final performance demonstrates that they do. With respect to the latter claim, our apes' ability to learn to gesture to the person who could see them can be taken as evidence that they understand the attentional aspect of seeing in one of two ways: (a) they finally figured out what we were (awkwardly) asking them, or (b) they finally constructed an explicit concept of seeing (which they normally do not need in order to get along in chimpanzee society). In either case, our results can be interpreted to show that reasoning about the mental dimension of seeing is not beyond the capacity of chimpanzees.

In summary, then, there are at least three distinct ways in which to characterize the nature of the social understanding connected with our chimpanzees' final, successful performances on the seeing/not seeing tests that we gave them.

1. A first possibility is that before participating in our tests, our apes did not possess a concept of attention. However, through the differential feedback they received, and the through our refinement of the tests, they came to construct such a concept, and indeed, learned one of its sensory bases (e.g., seeing).

2. A second possibility is that our apes entered the tests with a general, amodal conception of attention (perhaps interpreting attention as being governed by proximity, as opposed to sensory channels such as seeing, hearing, touching, etc.). However, again through the feedback that our tests provided, they constructed the notion of visual attention.

3. A final possibility is that our apes neither entered nor exited from our tests with an understanding of the mental state of attention. Rather, they constructed an "as-if" understanding of seeing-as-attention. On this view, our feedback procedures sculpted their behaviors into a form that matched our own.

Although there may be theoretical reasons for favoring one account over another, we found very little in our empirical results to exclude any of these possibilities (Povinelli & Eddy, 1996c, p. 134).

**Year 2: Failure of Retention**

Thirteen months after we completed the initial series of tests described previously, at least one means of distinguishing between these alternatives serendipitously presented itself. In the context of preparing our apes for a
different set of experiments concerning their understanding of joint attention, we returned to our seeing/not seeing protocols, and tested them on the eyes open/closed procedure. To our surprise, rather than finding this easy, the animals were just as likely to gesture to the person who had their eyes open as to the person who had their eyes closed. Indeed, even after 48 trials of this condition, the subjects were still not responding above chance (see Fig. 11.6). At first, we assumed that this was because eyes open/closed was the most subtle condition of all of those we had previously used, and therefore the animals may have never developed a robust understanding of it. Intrigued, we decided to test the apes on the condition with which they had received the most previous experience: screens. Again to our surprise, however, their performance only gradually crept up across the 12 trials we administered to levels just above chance (57% correct; see Fig. 11.6). We were thus forced to consider the possibility that despite the fact that our apes had been 89% correct on their final series of the screens trials a year earlier, they had apparently not consolidated this understanding into a form that would endure a year of participating in other tests.

Let me emphasize the significance of this failure of retention by pointing out that our apes were not simply lounging in the sun during the year that intervened between these two longitudinal time points. On the con-

![Year 2: Seeing/Not Seeing](image)

**Fig. 11.6.** Results from tests conducted during Year 2 for eyes open/closed, screens, and back/front conditions.
trary, they had participated daily in at least a dozen other experiments, all of which were designed to probe their potential understanding of attention (or other mental states; see Povinelli & O’Neill, in press; Povinelli, Bierschwale, & Čech, in press; Povinelli, Davidson, & Theall, 1997; Povinelli & Eddy, 1996a, 1996b, 1997; Povinelli, Perilloux, Reaux, & Bierschwale, 1998). Indeed, in many of these studies, the role of visual attention played a prominent role. Although these tests provided no better evidence that our apes possessed a genuine understanding of attention (or any other mental state), our fortuitous findings on the longitudinal seeing/not seeing task raised a much broader, and much more interesting, question: How did the experiences on these various tests interact with each other?

One might naturally assume that much like young children attending school, such testing experiences would build cumulatively, mutually reinforcing each other in a manner that would assist our apes in homing in on, and better understanding, the questions we were posing to them (using a variety of procedures). Indeed, we carefully planned the nature and sequence of our tests with this assumption in mind. To some extent, of course, this must be true. But our animals’ failure to rapidly understand even the easiest seeing/not seeing tasks at the second longitudinal time point caused us to realize that we needed to more directly confront the exact manner in which their testing experiences interact with each other.

Certainly skills accumulate. We have ample evidence that abilities the apes learned at one time point were retained years later: Once they learn how to do something, they remember for a very long period of time. But then why did they exhibit such poor retention on the seeing/not seeing tests? There are several possible explanations. One is that despite a full year of experiences on tests that should have helped to clarify the central construct that we were asking them about (i.e., the mental state of attention), our apes failed to integrate these new experiences with their older ones. More intriguing still, these new experiences might actually have interfered with what they had learned a year earlier. If our apes had never deeply understood why they were rewarded after gesturing to the person with the bucket on the shoulder as opposed to the person with the bucket over the head, for example, then such rules might never have been well-consolidated—especially given that we did not overtrain them on any of the conditions; the maximum number of trials they received in any of the conditions was 28 (in the screens condition). Given this rather weak understanding to begin with, rules or relations learned during intervening tests may have displaced or interfered with these older structures. In more stark terms, if our apes had no concept of attention, then all of our tests must have seemed like a bizarre collection of arbitrary reactions to certain social stimuli with nothing more concrete than our reinforcement procedures uniting them.
Year 3: Pryng Apart Chimpanzees’ Understanding of Seeing

We were so struck with our apes’ absence of retention at 7 years of age, that we returned to these tests when they were 8 to 9 years of age (see Reaux et al., in press). At this point, they had reached the end of adolescence—indeed, within a year, the first baby would be born to the oldest female in our group. It seemed reasonable to suppose that if chimpanzees develop a notion of seeing it ought to have emerged by this point. Our strategy for these final longitudinal tests was twofold. First, we exposed the animals to the main conditions we had administered two years previously (screens, buckets, blindfolds, eyes open/closed, distracted/attending, and back/front). We did this to determine if they would understand these conditions immediately (perhaps indicating a qualitative change in their understanding), or whether they would be forced to relearn them. Second, we intended to design some novel transfer tests that would allow us to make some final inferences about whether they had developed an appreciation of the attentional aspect of seeing.

First, the results of their first four trials of each of the old conditions (screens, buckets, blindfolds, eyes open/closed, distracted/attending) are depicted in Fig. 11.7. Perhaps the most striking aspect of these results is that, in the majority of these conditions, the apes displayed little evidence that they preferred to request food from the person who could see them. However, the apes did perform at levels exceeding chance in the buckets condition (and, as usual, were perfect on the back/front trials). There was no easy way to characterize these results. The results of the buckets condition might suggest that the apes were able to use the face rule, but their poor performance on the screens condition (in which the face rule could work just as well) did not fit with this idea.

Next, we administered eight more trials of each of these conditions. As can be seen from the filled bars in Fig. 11.7, this experience helped—at least for most of the conditions. With this additional experience, the animals’ performance improved to levels exceeding chance in 5 of the 6 conditions. The exception was the eyes open/closed condition. (Why the animals had greater difficulty on eyes open/closed is unclear. Perhaps it was simply a more subtle discrimination. On the other hand, it is possible that the apes thought the eyes might open at any moment). At any rate, on the whole these results seemed to fit the procedural rule model that had captured the apes’ performances over the two previous years. The apes seemed to have constructed a hierarchical rule structure in which the front, face, and eyes of the experimenters (in descending order of importance) served as the bases for where the apes directed their begging gestures. Recall that this model envisioned that the apes were trying to satisfy certain rule structures. First, they sought to approach and gesture
Year 3: Seeing/Not Seeing
(age 8-9 years)

Conditions

FIG. 11.7. Results from initial tests during Year 3 for screens, buckets, blindfolds, eyes open/closed, distracted/attending, and back/front. The dotted line indicates level of performance expected by chance responding. *p < .05, **p < .01.

to someone facing forward. Clearly, the back/front condition easily allowed them to satisfy this rule. On the other hand, if this were not possible (e.g., if both persons were facing forward), then the apes moved on to the next most important rule which was to gesture to the person whose face was visible. In the case of buckets and screens, this rule was easily satisfied. Of least importance was the eyes rule, and indeed, at this point there was only limited evidence that they were even using it. They did perform well on the blindfolds condition (which could be solved by the eyes, but not the face, rule). But with the exception of a female ape named Megan, the animals did not appear to understand the eyes open/closed condition yet. Furthermore, the subjects had relearned the distracted/attending condition, which could not be satisfied with any of these rules (because in this condition both of the experimenters' eyes and faces were visible).

Collectively, these results suggested that the apes were in the process of learning a set of hierarchical rules, but because these rules did not work all of the time (e.g., in the distracted/attending condition), they were also
learning condition-specific rules, which strictly relied on a discrimination between the two postures within those conditions. After reflecting on such condition-specific learning, we attempted to determine how the subjects were reasoning about those conditions on which they were already performing well. Our first approach was to mix together the correct (seeing) and incorrect (not seeing) option from each of several of the different conditions (see Fig. 11.8a). We reasoned that if the apes had learned a set of condition-specific relational rules, then their performance would be expected to decline if a correct option from one condition was mixed with an incorrect option from another. On the other hand, if the apes had extracted a context-independent understanding that certain configurations were correct (e.g., "gesture to the person holding a screen over the shoulder"), they could be expected to perform well on these mixed conditions. Finally, we considered the conditions that the majority of the subjects did not yet understand, such as eyes open/closed. We reasoned that by mixing the incorrect option from the eyes condition (eyes closed) with the correct option from one of the conditions on which they were performing well (blindfolds), we might uncover whether the apes understood these correct options as abstractly correct, or whether they were understood as correct only within the context of other options known to be incorrect.

The results of these tests are depicted in Fig. 11.8b. The most striking aspect is that the mixture of correct and incorrect options from conditions on which they were performing well posed no problem for the animals—in these cases they performed at levels exceeding chance (see results for +screens/-buckets and +screens/-distracted in Fig. 11.8b). In direct contrast, when we mixed the correct option from a condition on which they were performing well, with the incorrect option from the one on which they were not, their performance was random (see results of +blindfolds/-eyes and +buckets/-eyes in Fig. 11.8b). Together, these results provided further evidence that the apes were learning a rule about gesturing to a person whose face was visible, and that this rule could be flexibly deployed when we mixed the different options from such conditions together. However, their inability to move beyond this rule on the mixed conditions that involved the incorrect, eyes-closed option, suggested that the apes were treating the conditions as problems to be solved by comparing physical postures, not by reasoning about who could see them. Megan, however, did not fit this pattern. She performed reasonably well on both the final series of eyes open/closed trials (6/8 correct) as well as the mixed conditions involving eyes closed (6/8 correct), a performance significantly above chance (12/16, binomial test, $p = .038$).

Megan's performance intrigued us because she, unlike the others, had clearly learned the eyes open/closed condition during the first time point in this longitudinal project (see Povinelli & Eddy, 1996c, Experiment 13,
FIG. 11.8. (a) Mixed conditions used in Year 3. (b) Results of mixed conditions in Year 3. The dotted line indicates level of performance expected by chance responding.
p. 100)—although she did not learn it at the second time point, and showed a learning curve during the final time point. Such a subtle discrimination was impressive, and we knew that some researchers might interpret this as evidence that she, at least, had learned something about seeing per se. Her performance thus provided a focal point for our final test between the low- and high-level models. Recall that the low-level model stipulated that the apes were learning a set of procedural rules in which the front, face, and eyes served as the bases for their choices in descending order of importance. If this were true, then even for Megan the frontal aspect of a person would be more important than whether his or her eyes were open. To this end, we constructed the new condition shown in Fig. 11.9a, in which the correct option from looking-over-the-shoulder (a face rule condition) was combined with the incorrect option from eyes open/closed. This presented the animals with a choice between someone facing forward (but who could not see them because the eyes were closed) versus someone facing away from them, but looking over the shoulder with eyes wide open looking in their direction. We also tested the apes on three other conditions: back/front, eyes open/closed, and looking-over-the-shoulder (a condition the apes had not experienced in 2½ years). The low-level model predicted that the subjects would succeed on the looking-over-the-shoulder condition because they could use the face rule. However, the low-level model also predicted that on the new mixed condition the subjects would prefer the incorrect option! The reason for this counter-intuitive prediction is because of the postulated dominance of the front rule: In this condition the subjects would apply the front rule, and because it could be satisfied, would proceed to gesture to the person facing forward, even though the eyes were closed, and not to the person looking over the shoulder with clearly visible eyes and face.

The results of this final test are depicted in Fig. 11.9b, and provide striking confirmation for the low-level model. First, as expected, the subjects tended to prefer the correct person on the looking-over-the-shoulder trials. Second, and most important, the subjects performed significantly below chance in the mixed condition, meaning that they preferred to gesture to the person who was facing forward but could not see them—exactly as the low-level model had predicted. Although this was striking enough, there are additional aspects of the data set that are equally revealing. Megan, as well as Brandy and Kara, were almost perfect on the looking-over-the-shoulder and eyes open/closed conditions (8/8, 8/8, and 7/8, respectively, in the two conditions combined). Thus, these three animals exhibited a strong understanding of the conditions from which the novel, mixed condition had been composed. Yet when confronted with these conditions mixed together, Megan, Brandy, and Kara exhibited a strong preference for the incorrect option, selecting the person with eyes closed on 4/4, 4/4, and
FIG. 11.9.  (a) Final condition used in Year 3 consisting of eyes closed (−) and looking-over-the-shoulder (+) options. (b) Results from eyes closed (−) and looking-over-the-shoulder (+) condition. The dotted line indicates level of performance expected by chance responding.
3/4 trials, respectively. However one chooses to interpret these results, they certainly do not support the idea that Megan, or any of the other animals, selected the eyes open option because they understood that this person could "see" them.

THE REINTERPRETATION HYPOTHESIS

Chimpanzees may or may not possess a mentalistic appreciation of certain aspects of behavior. To be sure, their spontaneous behavioral patterns mirror our own so closely that we are tempted to assume that their general understanding of others is similar to our own. But the results that I have reviewed here counsel caution in following Hume, Darwin, and Romanes into assuming that behavioral similarity guarantees psychological similarity. In the remainder of this chapter, I outline an alternative view of the relation between the evolution of spontaneous social behavior and the evolution of social understanding that we have labeled the "reinterpretation hypothesis." (In the interests of space, the indictment of the argument by analogy is kept broad, but a detailed, formal critique can be found in Povinelli & Giambrone, in press.)

In order to appreciate the alternative that I am about to outline, it is necessary, curiously, to conceptually decouple the evolution of behavioral complexity from the evolution of cognitive abilities related to theory of mind. Recent thinking in this area has been in the opposite direction, with many theorists supposing that theory of mind evolved in order to cope with on-line social problems—the so-called social intelligence hypothesis (for various treatments of the social intelligence hypothesis, see Baron-Cohen, 1995; Byrne, 1995; de Waal, 1982; Gallup, 1982; Humphrey, 1976; Jolly, 1966; Whiten, 1996; Whiten & Byrne, 1988; see also contributions to Byrne & Whiten, 1988). But I explore an alternative possibility, namely, that although intense sociality (characteristic of many species of primates) drove the evolution of increasingly complicated social behaviors, these behaviors were governed by psychological systems essentially no different from the earliest mammals. To be sure, various lineages of primates may have evolved cognitive specializations in various domains to cope with the unique challenges posed by their way of life. But I suppose that social complexity—mediated through and manifested by gaze-following, deception, selective retaliation, reconciliation, and appeasement—all evolved long before abilities related to theory of mind. Thus, I wish to draw a distinction between the emergence of behavioral complexity generated through fairly low-level psychological mechanisms, and an understanding or interpretation of such behaviors in terms of mental states such as attention, desire, knowledge, and belief.
Next, let us consider the evolution of theory of mind. It is possible that only one species, our own, evolved a cognitive specialization in theory of mind. Perhaps it evolved as a separate, domain-general faculty; on the other hand, it may have evolved as a more specific system to aide in already complicated social computations. Finally, and perhaps most likely, it may have evolved in connection with the evolution of a specialized language capacity. But, no matter how it arose, this new psychology may not have replaced the psychological structures that predated it. A fundamentally new specialization in theory of mind may have been created without destroying the rich array of social behaviors that has emerged during the course of primate evolution. Psychological states (ultimately reducible to brain states) may, in a complicated and interdependent manner, carry the ancient alongside the new. Just as developmental constraints that were established in the Cambrian are reflected in the bodily structures of modern lineages, so, too, might we suppose that innovations in psychological structures accommodated the old. If true, this new psychological system may have allowed (indeed, forced) humans to reinterpret ancient social behaviors (behaviors that evolved long before we did) in explicitly mentalistic terms.

Gaze-following is an excellent case in point. This is a phenomenon shared by humans, chimpanzees, and many (perhaps most) species of monkeys. Humans are certainly capable of interpreting the shift in gaze of others in an explicitly mentalistic manner. That is, a shift in gaze can be equated with a shift in attentional or referential focus. Indeed, some researchers have interpreted gaze-following by 9- to 12-month-old human infants as reflecting precisely this kind of understanding (see Baron-Cohen, 1994). Others have demanded more rigorous demonstrations, but even using these methods they have argued that by 18 months or so there is evidence for such understanding (see review by Baldwin & Moses, 1994). Although I have reservations about both positions, surely there is some period during the preschool years in which a mentalistic understanding of gaze is consolidated. That is, at some point, gaze is understood as a behavioral ambassador for the mental state of attention.

This raises two problems for the argument by analogy, one developmental, the other evolutionary. From the developmental perspective, we may have already identified a case where similarity in behavior does not guarantee similarity in attending psychological states: gaze-following in 9- to 12-month-old infants. Although it is true that even in adults there may also be many cases in which the behavior of gaze-following is triggered before the representation of the other person's attentional state, there may also be other cases in which the adult's gaze-following behavior is directly caused by the representation of the other's internal state (e.g., "What is she looking at?"). In contrast, it is at least possible that the 9- to
12-month-old's gaze-following behavior may never be triggered by a representation of attention—perhaps because they cannot yet represent the behavior in mentalistic terms. Indeed, from the developmental perspective, this argument may really just be a general statement of Karmiloff-Smith's (1992) notion of "representational redescription." She argues that the dissociations often found in human development between children's ability to produce and their inability to understand it, are evidence for a process whereby procedural information is progressively redescribed in increasingly explicit ways.

The second case is the evolutionary problem. Instead of comparing an infant human to an adult human, let us compare an adult chimpanzee to an adult human. Although I have only presented a small portion of our work in this chapter, collectively it suggests that despite their sophisticated gaze-following abilities, chimpanzees have little or no understanding that gaze is linked to internal mental states. Thus, although a chimpanzee who follows your gaze tugs on a Humean inclination to assume that he or she is trying to figure out what you are "looking at," there is considerable reason to suppose that he or she does not understand the concept of "looking at" in the first place. Furthermore, the longitudinal project I have described in this chapter reveals that although we can get our chimpanzees (like 2- to 3-year-old children) to selectively gesture toward someone who can see them as opposed to someone who cannot, this similarity between their behavior belies profound underlying differences in the attending psychological interpretation of the act.

This hypothesis has at least one clear advantage over the social intelligence hypothesis: It neatly reconciles the evolution of social complexity with the data summarized in this chapter. The superficially confusing mosaic of similarity and difference no longer need be seen in conflict. In the view that I am offering here, ancient social behaviors are shared in common by humans and other primates such as chimpanzees. Thus, we expect and see an astounding degree of spontaneous behavioral similarity. But humans do more. Humans interpret this behavior. It may seem as if I am advocating a form dualism in which the representation of mental states accompany, but play no causative role in, behavior. On the contrary, I assume that humans evolved the capacity to represent the mental states of self and other because of its useful, causal connection to behavior. But as the case of gaze-following should make clear, two species may share a very similar behavior, controlled by similar low-level mechanisms, but interpret that behavior in very different ways. Does this mean that the human trait of representing the attentional aspect of gaze is superfluous? No, because although it may not be directly linked to the act of following gaze, the representation of attention (and other mental states) may play a profound role at a higher level of behavioral organization. Once humans evolved the capacity to represent other organ-
isms not just as behaving beings, but as beings who attend, want, and think, it became necessary to use already existing behaviors to respond to the world of the mind, as well as the world of behavior. This distinction may be especially important in those psychological arenas in which we seem to differ most from other species such as culture, pedagogy, and ethics. Determining the correct causal connections between individual behavioral elements and the mental states that ultimately control their expression may simply be a project beyond the reach of introspection alone (for a more thorough description of the causal role I ascribe to higher order intentional states, see Povinelli & Giambrone, in press).

Thus, the error in the argument by analogy appears twofold. First, the argument mistakenly assumes that an accurate introspective assessment of second- (or higher) order intentional states that accompany our own behaviors is likely to yield an accurate inference as to their cause. Second, the error is compounded by the assumption that because mental states accompany our behavior, similar mental states must accompany similar behavior in other species. Although humans form internal representations of mental states such as desires, knowledge, and beliefs, and these representations are now connected in causal ways to our behaviors (albeit in ways too complicated for introspection to reveal), our basic behavioral repertoire evolved long before those higher level representations were possible.

ACKNOWLEDGMENTS

This research was supported by NSF Young Investigator Award SBR-8458111 to D. J. Povinelli. This research 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, in particular Anthony Rideaux, James Reaux, Donna Bierschwale, Laura Theall, and Timothy Eddy. Photographs are by Donna T. Bierschwale and Corey G. Porché.

REFERENCES


11. SOCIAL UNDERSTANDING IN CHIMPANZEES


