

research on the phylogenetic distribution of mirror self-recognition, and it is a hypothesis that can be (and is being) empirically evaluated. Yet C&S dismiss nearly a quarter century worth of research and theory construction on this topic by trying to turn the issue into a question of degree. "All animals," they write, " – and even cells – recognize themselves to some degree. At what point, however, do individuals become able to reflect upon themselves, in a sense to interview themselves?" (p. 174). Yet, as we pointed out, this is precisely the issue at hand: the distinction between the capacity for self-perception and the capacity for self-conception (Butterworth 1992). Mirror self-recognition is a behavior, not a psychological capacity. What makes the behavior of such potential interest is that it may qualify as a method of producing interspecific evidence of self-conception (or self-awareness). Are there many ways of "recognizing" oneself? Yes, of course, but many of these forms of recognition may only be based upon an unconscious capacity for self-perception that extends in all directions throughout the web of life. The question, as I see it, is not a nebulous issue of "continuity" versus "discontinuity," but rather a question of shared derived features of portions of a developmental program controlling the expression of various components of a theory of mind (see Povinelli 1991; Povinelli & Godfrey 1993; Povinelli et al. 1991).

As an instance of how vague the issue of degree can become, consider C&S's notion that a monkey's ability to "recognize" its own dominance rank relative to others indexes some form of self-recognition. No example could make more clear the need to sharpen the distinction between the psychological capacities of self-perception and self-conception. Does the behavior of responding appropriately in a dominance hierarchy imply the underlying capacity for self-conception? It might, and C&S may ultimately wish to construct a theory that predicts this relationship. If so, the extent to which the theory will be of heuristic value will be the extent to which it predicts the presence of experimentally demonstrable related behaviors that ought to be present in organisms that display appropriate behavior in a dominance hierarchy. I would be surprised if such a relationship held true, precisely because this behavior is probably the result of self-perception, not self-conception. But, as always, the issue is an empirical one.

In contrast, Gallup (1982; 1985; 1991) has predicted (note: *predicted*) that mirror self-recognition is impossible without a primitive form of the capacity to become the object of one's attention, to reflect on one's own existence, or "to interview oneself." (The earliest form of the capacity that supports mirror self-recognition remains unclear, although new evidence suggests that it may be related to the emergence of a sense of personal agency; Povinelli 1991). Gallup's model also predicts that species incapable of mirror self-recognition do not have this capacity. Gallup (1982; 1985) and I (Povinelli 1991) have made it explicitly clear that these are predictions, not descriptions. Thus, the primate mirror self-recognition literature generated a model over a decade ago that anticipated the descriptions offered by C&S.

And what of the evidence for the existence of mental state attribution in primates? C&S worry that the laboratory setting will continue to stifle the chimpanzee's capacity for displaying its intellectual prowess. As an example, they cite one of our own theory of mind experiments with chimpanzees in which we investigated their capacity to understand the causal connection between *seeing* and *knowing*. They note that the chimpanzees only selected the correct human about 70% of the time, and that if my colleagues and I had been circus trainers we could have achieved the same results with pigs. It is fortunate we are not in the circus business, I guess, because when we attempted the same experiment with macaques, we had no luck whatsoever. Likewise, the 3-year-old children's responses were similar to those of the macaques (including a pilot subject to whom we gave extended training over a month-long period). The 4-year-

A theory of mind is in the head, not the heart

Daniel J. Povinelli

Laboratory of Comparative Behavioral Biology, New Iberia Research Center, University of Southwestern Louisiana, New Iberia, LA 70560

Will researchers outside the field of empathic ethology ever be convinced that at least some nonhuman species are aware of the existence of the mental world? I predict that the day will not arrive until a revolution in the field of comparative primate psychology occurs. Ironically, despite the claims of many, the success of the revolution-in-waiting no longer has anything to do with academic taboos against investigating mental phenomena. It has to do with methodology, and, in particular, a commitment on the part of comparative psychologists to retaining the rigorous approach of behaviorism while simultaneously discarding its dogmatic ideology (e.g., Skinner 1987). [See also *BBS* special issue on Skinner, *BBS* 7(4) 1984.] But in rejecting the ideology of behaviorism we must not resurrect the alternative ideology that spawned it in the first place: animal mentalism (e.g., Griffin 1976). Investigations of an organism's awareness of mental states can profitably proceed only insofar as they engage in hypothesis-testing, as do other branches of science.

Although I find myself agreeing with much of what Cheney & Seyfarth (C&S) (1992) have to say, I remain troubled by several issues. To begin, C&S display an apparent unwillingness to consider the logical possibility that monkeys truly lack a theory of mind, that is, that they are unconscious biological systems that have no reflective awareness of their own minds or the minds of others. As we pointed out in our original commentary (Povinelli & deBlois 1992a), this is not a conclusion – it is a hypothesis that was developed a decade ago from Gallup's

olds, in contrast, were right from Trial 1 forward (Povinelli & deBlois 1992b).

Does this contrast between species mean that chimpanzees definitely understand the *seeing-knowing* relationship? Absolutely not. As we discussed at length in our original report, disentangling learning and attribution-theoretic accounts is a difficult process with nonhuman primates, and no one experiment should be taken as evidence for the existence of a particular attributional capacity (see Povinelli et al. 1990, pp. 207–10). Nonetheless, it should be noted that the chimpanzee's 70% performance rate is as much a potential embarrassment for learning theory as it is for attribution theory. If the chimpanzees learned a simple behavioral rule (for example, "pick the person who stayed in the room") then why did they only achieve a 70% performance rate? To be sure, explanations are readily available, such as the complex and shifting nature of the cues, poor short-term memory, as well as species differences in temperament (Povinelli et al. 1990). But each of these explanations can be used to buttress a mental-state attributional account of the data as well. For example, if the chimpanzees had difficulty remembering who left the room from trial to trial, they would perform at depressed rates with or without a theory that equates seeing with knowing. Yet our macaques (with whom circus trainers have often had a field day), who are every bit as proficient as rule-learners as are chimpanzees, never made the relevant discrimination (Povinelli et al. 1991). In the face of such intermediate results, I have advocated caution in interpreting the existing evidence, and more research, using a modified paradigm that both takes into account those potential species differences and can be used to distinguish between a learning theoretic and mental-state attributional account (Povinelli 1991). We are currently conducting these (and related) studies with a large population of socially reared chimpanzees of various ages.

Finally, provided the chimpanzees we test have had a rich, social upbringing, I find C&S's concern about the laboratory setting questionable. To be sure, certain experiences in captivity may affect a chimpanzee's performance on a task, but it must be kept in mind that such effects potentially run in both directions. As for the unnaturalness of the laboratory tasks, we must ask, when a captive chimpanzee confronts someone with a paper bag covering his head, is the inference of "not seeing" any more difficult than when the chimpanzee confronts another chimpanzee whose head is obscured by a tree branch? And to turn the tables entirely, how natural a setting can it be for a 4-year-old child to confront one adult wearing a paper bag, and another adult who is not, and then ask them to respond on an apparatus that was originally designed for use with monkeys? Yet we can be assured that children will make the *seeing-knowing/not seeing-not knowing* inferences easily, and at the same age that they make it in other seemingly more "natural" settings (Povinelli & deBlois 1992b). Can the chimpanzees' difficulty be rescued by claiming that they have a weaker theory of mind and therefore need more natural settings? Again, the data from the younger children make this claim unlikely.

C&S also contend that the chimpanzees' performances on such laboratory tests would improve markedly if they were required to reason about mental states in other chimpanzees, as opposed to human actors. Admittedly, this is an empirical issue, but at least one point should be made. All evidence indicates that a theory of mind is hard-wired to apply to most animate entities (frogs, cows, elephants, trees – even circus pigs), and many inanimate objects as well (dolls, the sun) (see Miller & Aloise 1989). Indeed, early theory of mind research in young children began by using dolls as a substitute for real people (and continues to do so, e.g., Wimmer & Perner 1983). It is only later in development that humans learn (sometimes) to suppress these attributions about inanimate objects, and they do so considerably less when it comes to animate ones. Thus, if chimpanzees and humans share some portion of a developmental program that governs mental state attribution, it probably

does not matter whose mental states they are attempting to reason about.

Finally, what of the cross-cultural validity of theory of mind research? In commenting on C&S's article, Furth (1992) claims that fascination with this phenomenon has more to do with middle-class Western ideology than with reality. This is a serious critique, because if correct, it would call into question the entire program of mental-state attribution research launched by Premack and Woodruff (1978). A couple of years ago, Laurie Godfrey and I took this criticism to heart and examined evidence offered by some cultural anthropologists who claim that concepts of the private self, intention, and knowledge are peculiarly Western constructs, not found in many other cultures (Povinelli & Godfrey 1993). Under careful examination, such differences turn out to be surface distinctions in the ways in which these capacities express themselves under varying cultural influences. Do cultural differences exist in the ways in which humans conceive of the private self? Absolutely. Do some cultures not share a concept of self? Absolutely not. A variety of lines of evidence (often derived from the critics themselves) point to the cross-cultural existence of the concepts of self, private intentions, and personal knowledge (see Povinelli & Godfrey 1993). Of course, there is nothing like precise, cross-cultural experimentation explicitly designed to determine whether the development of specific capacities is stable across very different cultural backdrops. Avis and Harris (1991) have recently provided the first empirical step in this direction by demonstrating the 3- to 5-year-old age transition in understanding the concept of false belief among children of the Baka, a preliterate gatherer-hunter society in Cameroon.

During the past two decades, comparative psychologists such as David Premack and Gordon Gallup have realized that the tools of behaviorism can be used to ask questions about how nonhumans understand the mental world. Thanks to their efforts, researchers in a variety of fields are now in a position to make tremendous progress toward characterizing the minds of other species, with "progress" defined as theory construction and hypothesis-testing. C&S's most recent contributions to this field are of great importance (e.g., Cheney & Seyfarth 1990a). But we must always be on guard to remember that descriptions can never replace predictions, and predictions must always face attempts at falsification. Evolution has no vested interest in creating monkeys with minds, or monkeys without minds. Until the description-to-argument cycle is replaced by a prediction-to-data-collection cycle, we will continue to be alternately embarrassed by both the paucity and the richness of the mental contents of the complex biological systems that evolution has unknowingly produced.