MIRROR, MIRROR ON THE WALL WHICH IS THE MOST HEURISTIC THEORY OF THEM ALL? A RESPONSE TO MITCHELL*  

GORDON G. GALLUP, JR  
Department of Psychology, State University of New York at Albany, Albany, NY 12222, U.S.A.  
and  
DANIEL J. POVINELLI  
University of Southwestern Louisiana, New Iberia Research Center, 100 Avenue D, New Iberia, LA 70560, U.S.A.  

NEED FOR CONTROLS  
Before we evaluate the theories that Mitchell outlines to account for mirror self-recognition, it is important to correct some misleading information in his paper. While Mitchell equates the techniques developed by Gallup (1970) and Amsterdam (1972) for assessing self-recognition (e.g., he refers to them as "the Gallup–Amsterdam technique," p. 298), in fact they are quite different. Table 1 highlights some of the basic distinctions between the two methods. For those interested in a further analysis of these differences and their implications, Gallup (in press) provides a more detailed discussion of some of the problems inherent in Amsterdam’s methodology.  

Mitchell also misconstrues the use of the so-called "mark test." Like many others (e.g., Boccia, 1991; Thompson, 1991), Mitchell makes the mistake of equating self-recognition with successful performance on the mark test. However, the use of unobtrusively applied facial marks was developed only as a means of validating impressions which arose out of seeing animals use mirrors to respond to themselves in ways which suggested that they realized that their behavior was the source of the behavior being seen in the mirror (Gallup, 1970). In the absence of spontaneous instances of mirror-mediated self-exploratory behavior there is no reason to conduct a mark test. Not only does the use of the mark test require a particular context (i.e., using mirrors to respond to the self), but as detailed in Table 1 it also requires the use of some essential control conditions. Lacking any apparent concern for these kinds of methodological details, Mitchell credits a bottlenosed dolphin and several species of macaques as having shown self-recognition by passing the mark test. Yet none of the studies he cites included any of the control conditions needed to provide conclusive evidence of self-recognition, nor do they report any

Table 1. Comparisons between the Gallup (1970) and Amsterdam (1972) techniques for assessing self-recognition

| Inclusion of data on patterns of self-directed responding to mirrors before conducting the mark test. | Gallup (1970) | Amsterdam (1972) |
| Unobtrusive application of facial marks. | Yes | No |
| Use of marking material free from tactile and/or olfactory cues. | Yes | No |
| Marks placed on facial features only visible in a mirror. | Yes | No |
| Inclusion of control subjects without prior mirror experience. | Yes | No |

unprompted instances of mirror-mediated self-directed responding (for more details see Gallup, 1991, in press).

At another point, Mitchell asserts that as far as humans are concerned “experience with mirrors is not necessary to pass the mark test” (p. 302). Putting aside the fact that a literal interpretation of this statement would imply that humans somehow have innate knowledge of the properties of mirrors, the study he cites (Priel & de Schonen, 1986) in support of this position is less than convincing. They report that children living in tents in rural areas of Israel without mirrors show self-recognition at about the same age as their mirror experienced urban counterparts. Although Priel and de Schonen (1986) discarded some children based on reports from mothers about the presence of things in the home that might have provided them with access to their reflections (e.g., toys, tin cans, bottles), their study provides no assurance that the children in the rural sample did not have prior experience with their reflections in automobile hubcaps, chrome bumpers, and rearview mirrors, as well as reflections emanating from a variety of other objects outside the home.

Mitchell also expresses “skepticism about the mark test as evidence of self-recognition” (p. 304) on the grounds that the infant that touches a mark on its face may merely be trying to ascertain whether it has a mark on its face which corresponds to the mark on the face of the unfamiliar infant it sees in the mirror. But again, appropriate control procedures can be used to preclude this kind of dilemma. For instance, in Gallup’s (1970) original study of self-recognition both mirror experienced and mirror naive chimpanzees were anesthetized and marked on facial features which could only be seen in mirrors. Contrary to Mitchell’s contention, none of the mirror naive chimpanzees showed any mark-directed behavior on the test.

MIRROR CORRESPONDENCE

Mitchell also asserts that “mirror correspondence may not be understood prior to mirror-self-recognition” (p. 300). From our perspective, it would be more
accurate to say that understanding mirror correspondence does not suffice to produce self-recognition. Even monkeys which seem incapable of recognizing themselves in mirrors, can readily learn to respond to objects (other than themselves) in mirrored space (e.g., Anderson, 1986; Brown, McDowell, & Robinson, 1965; Itakura, 1987a, 1987b; Tinklepaugh, 1928). The same holds true for elephants (Povinelli, 1989), who like monkeys seem unable to recognize themselves in mirrors but can use mirrors to locate otherwise hidden objects. Thus, a more accurate assessment would be that both prior experience with mirrors and an understanding of at least some aspects of mirror correspondence are probably necessary but not sufficient conditions for self-recognition. On the other hand, this is not to imply that the ability to recognize one’s own reflection in a mirror means that one has mastered all of the nuances of mirrors and mirrored space. Even mirror sophisticated adults occasionally make momentary misinterpretations of their own reflections when they suddenly and unexpectedly catch sight of themselves in mirrors (see Gallup, 1968). Likewise it is one thing to learn the contingency inherent in mirrors, but quite another to learn the perceptual-motor skills needed to respond to mirrored cues. This distinction is clear from the initial difficulty that even adult humans have in mastering mirror drawing and mirror reversal tasks.

THE “INDUCTIVE” MODEL

It is important to clarify Mitchell’s use of the terms “inductive” and “deductive” as applied to the alternative theories he outlines. Mitchell clearly believes that in one of his accounts an organism inducts self-recognition, whereas in the other the organism deducts it. The problem is that Mitchell’s use of these terms reduces to a semantic trick. For instance, in reference to the “inductive theory” Mitchell states that “the organism knows that most mirror-images are images of what they are contingent with and resemble. The organism infers that this mirror-image may be an image of what it is contingent with and similar to, which is the organism’s action” (p. 300, italics added for emphasis). Why not substitute “all” for “most” in the first sentence, and “must be” for “may be” in the second? Under these conditions the organism would not “infer,” but rather “deduce” the source of its mirror-image. Indeed, a page earlier Mitchell states that “once the property of mirror-correspondence is recognized and utilized, an organism understands that objects other than itself both resemble and are the source of their mirror-image” (p. 299). Does this hold for all objects, or just some? If it is true for all objects, and we see no reason why some should be treated as special cases by either the organism or Mitchell, the organism would not be making an induction at all; rather its cognitive structure in this domain—its feeling of necessity—would force it to conclude that the mirror-image was the visual manifestation of its kinesthetic sense of self.

Having said this, it nonetheless seems to us that as presented the kinesthetic–visual hypothesis is false. Even young infants meet the criterion outlined by Mitchell: that is by 13 months of age many infants are capable of “understanding” mirror correspondence (Robinson, Connell, McKenzie, & Day, 1990) and are capable of imitation (Meltzoff & Moore, 1977, 1989), and thus, according to Mitchell, visual–kinesthetic matching. Yet they do not recognize...
themselves in mirrors. In what appears to be an effort to salvage this model to contrast it against the deductive model, Mitchell later argues that perhaps the inductive model can be rescued if “further along in the development of imitation organisms have a kinesthetic–visual matching capacity which is mature enough to support mirror-self-recognition” (p. 303). However, this “further development” is at least partially present in newborn infants who are able to imitate head rotations, tongue protrusions, mouth openings, and lip protrusions (Meltzoff & Moore, 1977, 1989).

Mitchell also argues that infants below 18–24 months of age may not understand mirror correspondence. First, it is curious that Mitchell lists this as a problem for the “inductive” model, but not the “deductive” one. In both models mirror-self-recognition is not considered possible until such correspondence is understood. Therefore, it is an equal problem for both of his models. Second, Mitchell spends a good bit of time (rightfully so), pointing out the methodological problems of the literature on children’s understanding of mirror correspondence, but then cites Maria Boccia’s work with pigtailed macaques as unambiguously positive evidence for understanding mirror correspondence. Yet her monkeys were tested without any better controls than the Robinson et al. (1990) infants. Worse yet, Japanese macaques (Itakura, 1987a, 1987b) and elephants (Povinelli, 1989) are cited as also understanding mirror correspondence. Both species are capable of learning to use mirror cues to respond to objects in mirrors, but this does not mean that they “understand” mirror correspondence. Surely 14-month-old infants, if given the amount of previous mirror exposure in a certain setting that Povinelli’s (1989) elephants received, or the elaborate training that Itakura’s (1987a, 1987b) monkeys received, could show similar behavior.

THE DEDUCTIVE MODEL

Now let us turn to the deductive model. Simply because stage 6 object permanence implies that an organism may possess “some primitive deductive abilities” (p. 307), this in no way means that it must follow the deductive route laid out by Mitchell. An organism with deductive capabilities need not use them often, properly, or even at all. Indeed, Charles Darwin’s mixture of induction and deduction in arriving at the theory of natural selection (Gruber, 1981), is an excellent example of how an organism with all of the components necessary for deducing a principle, may follow a variety of routes instead. Thus, the distinction between an inductive and deductive model might be highly idiosyncratic, if meaningful at all. There is no reason to believe that the deductive model is intuitively more appealing simply because it is labeled “deductive”.

Contrary to Mitchell’s contention that “the deductive abilities of the organism inform the organism that the body in the mirror is a contingent accurate image of its own body” (p. 308), a simpler version of this model would be that the organism treats its body as any other object located in space. Thus, the organism need not have any representation of itself before or after self-recognition occurs that is fundamentally different from its representation of other objects. In the Piagetian model, the stage 3 response to a partially hidden object is treated as indicating that objects are conceived of as wholes which are composed of parts, with the invisible
totality being reconstructed from the visible portion. Thus, organisms with a stage 3 understanding of object permanence can perceive part–whole relationships. In this way, it can be seen that the entire concept of “body-part objectification” is unnecessary. Given this capacity for part–whole reconstruction, there is no reason why “body-parts” should be any different from parts of other objects. Unless, of course, one makes an a priori assumption that animate objects are unique in the realm of objects. Mitchell provides no reason why they should be conceived of in this way. In contrast, we can conceive of one reason why organisms which attribute mental states might conceive of bodies as being fundamentally different—it is because unlike ordinary objects they can be repositories of intentions, desires, knowledge, and beliefs [(Premack & Dasser, 1991); but also see Eddy, Gallup, and Povinelli (1993) for some boundary conditions].

Furthermore, Mitchell’s two stipulations of what is meant by body-part objectification further expose the superfluous nature of the concept (p. 308). The first stipulation is (by his admission) subsumed by stage 6 object permanence. The second (that body-parts are conceived as “objects” even in unusual circumstances) is not necessary for the deductive process he outlines to occur. If an organism can reconstruct a whole from a part, it only needs to know that $x$ (e.g., the hand area of the body) is continuous with $y$ (the wrist area of the body). It does not have to “objectify” the hand or the wrist, it only needs to expect to see wrist-like areas continuous with hand-like areas. Although we write this as if they were objects, organisms need not conceive of these things as such in order to expect to see wrist-like areas near hand-like areas. They need only form expectations based on association, and as a consequence the entire issue of body-part objectification is quite unnecessary for the deductive model. Thus, if this model is correct self-recognition has nothing to do with self-awareness at all, except in the trivial case where self-awareness is merely defined as self-recognition.

**MATURE VERSUS IMMATURE SELF-RECOGNITION**

Although self-recognition might appear to be an all or nothing phenomenon, in fact the child’s perception of its own reflection probably undergoes a series of changes which continue long after it has learned to recognize its self in a mirror. As evidence for the continuing development of the capacity to objectify external referents of the self, consider a study by Cameron and Gallup (1988) on shadow recognition in human infants. Unlike studies of mirror self-recognition which show that most children learn to recognize themselves between 18 and 24 months of age, Cameron and Gallup found that shadow self-recognition was developmentally delayed, with many children failing to show clear evidence of recognizing their own shadows until they were at least 3 years of age or older. Evidence for being able to appreciate shadow correspondence, however, emerged much earlier with over 25% of the infants showing evidence of understanding that shadows have objective, tangible counterparts in reality before 18 months of age.

An issue related to this in Mitchell’s paper is the idea that there may be changes in self-recognition which occur during development. Following Zazzo (1982), Mitchell states that between 5.5 and 6.5 years of age children lose their capacity for self-recognition. This is not based on hard data. On the other hand, there is
evidence for changes in the child’s recognition of itself across time (Mounard & Vinter, 1985). These changes may reflect some interesting and important changes in the child’s conception of itself, but to claim that children lose the capacity for self-recognition between 5.5 and 6.5 years of age is disturbing. In addition, how Mitchell concludes from this that the mark test necessarily “derives from different cognitive capacities in different species” is unclear. We do not deny this as a possibility, and indeed following a suggestion first offered by Alison Gopnik, we have recently begun a formal examination of what we refer to as the mature–immature hypothesis. In short, we are examining whether or not, as the self-attribute system becomes richer through time, the chimpanzee’s appreciation of its mirror image becomes concurrently richer as well. Although young chimpanzees are clearly capable of self-recognition at some point, it remains unclear, for example, how psychologically important they consider their visual appearance. One method of testing this is to confront both young and adult chimpanzees who recognize themselves with radical alterations in their appearance. We have already piloted a project with a precocial 3-year-old chimpanzee using a “perm-shock” technique in which the animal’s hair was bleached following self-recognition.

The important point here is that self-recognition probably does not “disappear” as an individual matures, but rather undergoes redefinitions as the self-concept continues to undergo experience and maturationally driven changes.

**SELF-AWARENESS**

Mitchell makes some important points in the concluding section of his paper concerning various views of the self. Indeed, as Povinelli (1991) has pointed out, these fundamental distinctions of how the self-attributional system operates have long-standing historical antecedents. Allport (1948), for instance, lists at least eight conceptually distinct aspects of how the self (and hence self-awareness) has been treated in psychology. These cover all of those outlined by Mitchell, plus several others.

Having said that, however, let us turn to some of the distinctions in question. Mitchell distinguishes between a first sense of self-awareness in which an organism is capable of thinking about “its thoughts and internal representations and experiences” (p. 312), and a second sense in which an organism has “an understanding or knowledge of its personal history and attributes, motives, thoughts, actions, hopes, desires, and body identified as its own” (p. 313). This distinction, although literally possible, is only instructive insofar as we are concerned with the extent to which an organism’s memory is incorporated into the self-attributial system. The key issue that links these forms of self-awareness is the existence of a sense of self in the first place. No amount of redefinition will escape this problem. An organism capable of self-awareness in the first sense possesses a capacity to reflect and direct its conscious experience at least in the present, and in the second sense it has extended it to include more remote points in time. Mitchell argues that this need not be true, that in the first sense an organism need not have a “self” understood as such, but rather, as he puts it, “the organism’s awareness is (recursively) aware of itself” (p. 313). But the intersection between unconscious
Self-recognition and evidence of mental state attribution

Although Mitchell concludes that Gallup's attempt to link self-awareness (in the first and second sense) with an organism's awareness of the mental experiences of others has "failed" (p. 316), the data will ultimately decide the fate of that model. Like any other, Gallup's model is no better than the set of falsifiable predictions it makes. While Mitchell discounts the model because the awareness of the minds of others in humans appears gradually, this was acknowledged long ago by Gallup and Suarez (1986). Moreover, Povinelli (1991, 1993; Povinelli, Parks, & Novak, 1991) has outlined how the model can be rigorously tested by systematic comparisons between the ontogeny of mental state attribution in children and other great apes. Indeed, it may be that at each stage of development in the ontogeny of the child, knowledge of self and others is intimately linked and essentially nondissociable (Gopnik & Meltzoff, 1991).

It is unfortunate that Mitchell has chosen to rely on anecdotal evidence to conclude that "certainly chimpanzees, orangutans and the gorilla Koko actively pretend and/or deceive" (p. 316). If intentional deception in apes based on the capacity to (1) understand and (2) manipulate beliefs were a closed issue, then demonstrating these phenomenon should be straightforward using laboratory procedures. So far, it has not (Povinelli, 1991; Povinelli, Nelson, & Boysen, 1990; Premack, 1988; Woodruff & Premack, 1979). Nonhumans may attribute knowledge and beliefs to others, but we still await the empirical evidence that provides definitive documentation of this. On the other hand, there is considerably better evidence that chimpanzees attribute intention and desire to others (Povinelli, Nelson & Boysen, 1992; Premack & Woodruff, 1978; see Povinelli, 1991, for a review). Indeed, Mitchell argues that an awareness of another's perspective introduces problems for humans not faced by apes in relation to creating an ideal image in a mirror. Yet if the evidence of mental state attribution in apes is valid, as Mitchell appears to believe from the anecdotal evidence he cites, then why exactly, is this alleged difference present? And furthermore, is it an inherent psychological difference or a cultural one? Just as it is true that the same behaviors might be produced through different psychological mechanisms, culturally distinct behaviors can be based on identical underlying psychological mechanisms.

Conclusion

The fact remains (despite Mitchell's willingness to uncritically accept anecdotal reports and incomplete results from studies lacking appropriate controls) that no-one has produced definitive evidence of self-recognition in anything other than members of the great ape–human clade. Provided they can achieve stage 6 object permanence, Mitchell's account of this difference is that these other species simply lack body-part objectification. Our account of this difference, on the other hand, treats self-recognition as an indicator of self-awareness and as such makes a number
of testable predictions about a variety of other cognitive capacities which ought to
covary with the emergence and subsequent development of self-attribution along
both ontogenetic and phylogenetic lines. The so-called deductive theory fails to
either predict or account for a growing body of evidence which points to the
existence of some striking but heretofore unanticipated differences in social
intelligence between great apes and monkeys, as well as some equally interesting
changes in cognitive function among human infants which begin to emerge at or
after the point they come to recognize themselves in mirrors (e.g., Brownell &
Carriger, 1990; Lewis, Sullivan, Stanger, & Weiss, 1989; Povinelli, in review).

To qualify as scientific, theories must be testable. But the real utility of a theory
is proportional to the extent to which it actually generates research. It will therefore
be of interest to see if Mitchell’s position on body part objectification entices him,
or anyone else, to seriously attempt to test his “deductive” model. To do so, one
need only teach dogs or rhesus macaques (both capable of stage 6 object
permanence) to distinguish between several body parts (e.g., legs and tails, hands
and feet) and then conduct a formal test of self-recognition. We await the results of
such a test with keen anticipation.

REFERENCES

Amsterdam, B. (1972). Mirror self-image reactions before age two. Developmental Psychobiology,
5, 297–305.

Anderson, J. R. (1986). Mirror-mediated finding of hidden food by monkeys (Macaca

presented at the symposium on Cognitive Correlates of Self-awareness and Self-
recognition in Monkeys, Apes, and Humans, Sonoma State University, Rohnert Park,
California.


der differentiation during the second year. Child Development, 61, 1164–1174.


Eddy, T. J.; Gallup, G. G., Jr., & Povinelli, D. J. (1993). Attribution of cognitive states to
animals: Anthropomorphism in comparative perspective. Journal of Social Issues, 49,
87–101.


limitations and cognitive consequences. In G. R. Goethals, & J. Strauss (Eds.), The self: An
interdisciplinary approach (pp. 121–135). New York: Springer.

S. Parker, R. Mitchell, & M. Boccia (Eds.), Self-awareness in animals and humans. New York:
Cambridge University Press.

humans and other primates. In J. Suls, & A. G. Greenwald (Eds.), Psychological perspectives

relation to understanding the mental states of others. Paper presented at the symposium on
Mirror, mirror on the wall

Cognitive Correlates of Self-awareness and Self-recognition in Monkeys, Apes, and Humans. Sonoma State University, Rohnert Park, California.


