

## Self-recognition in primates: irreverence, irrelevance and irony

C. M. HEYES

*Department of Psychology, University College London*

*(Received 11 December 1994; initial acceptance 20 March 1995;*

*final acceptance 24 May 1995; MS. number: sc-1055)*

Mitchell (1993, 1996; Parker et al. 1994) and I (Heyes 1994a, b, 1995) doubt that Gallup's mark test (e.g. Gallup 1970) is a valid measure of mirror self-recognition, and that self-recognition may be explained with reference to self-conception or self-awareness. However, the present discussion focuses on points of disagreement between Mitchell and me. While we are both irreverent about self-recognition, Mitchell reveals his ultimate fidelity to the tradition of research in this area by relying on evidence that cannot resolve the issues, and proposing an alternative to the self-concept explanation of self-recognition no less speculative than that which it was designed to supersede.

Mitchell (1996) defends the view that sound evidence of mirror self-recognition is provided by descriptions of self-directed behaviour, such as picking at teeth, which appeared to the observer(s) to have been executed while the animal was looking at its mirror image. He claims that 'objectively', 'looking is part of the animal's "behaviour"' and that consequently it would be both impossible and uninformative (could it be both?) to use a control procedure in which frequency of self-directed behaviour is compared in the presence and in the absence of a mirror.

What is the import of Mitchell's claim that looking is an observable property of behaviour? Perhaps he means that, in relation to self-recognition, the evidence of an observer's senses is infallible. If what an observer sees of an animal's behaviour with a mirror leads him or her to think that the animal was using the mirror image of its body, receiving and processing information from that source, then he or she could not be mistaken. This interpretation is consistent with the weight that Mitchell assigns to anecdotal reports of self-

directed behaviour, but it is, none the less, deeply implausible.

'Folk' and cognitive psychology both regard the reception and processing of information as internal events which may be inferred from behaviour, but not observed directly. It is understood that an individual can 'look right through' an object, or maintain visual orientation to a stimulus without attending, and consequently that 'looking' does not necessarily imply 'seeing' or using information from the object. In contrast, there are theories of mind (e.g. behaviourism and the ecological view of perception) which suggest that mental states and processes are manifest in behaviour (Heyes & Dickinson 1990), but Mitchell did not explicitly adopt one of these perspectives.

On the other hand, if Mitchell accepts that observers make fallible judgements about mirror use, then it is not clear why he places so much faith in anecdotal evidence and objects to the control procedure that I described (Heyes 1994a, b) as an initial check on the validity of these judgements. It is surely not impossible for observers to record self-directed behaviour that occurs when an animal is facing a mirror (experimental condition) and when it is facing a non-reflecting surface of the same size and in the same position (control condition). The former, but not the latter, may seem to the observers to occur while the animal is using its mirror image, but neither the observers themselves, nor we who receive their reports second hand, would have reason to rely on their impressions unless self-directed behaviour in the control condition was of a lower frequency, or a different form, than in the experimental condition.

In summing up his objections to the use of a mirror-absent control procedure, Mitchell's question suggests a mistaken view of the function of experiments (studies involving control procedures) and an implicit equation between self-report and anecdotal data: 'If I groom a blemish

Correspondence: C. M. Heyes, Department of Psychology, University College London, Gower Street, London WC1E 6BT, U.K.

10 times on my face, then move my face in relation to a mirror to look at the blemish and groom it once, and then groom the blemish 10 times in the same way without a mirror, does that mean that I did not use the mirror as a guide for self-directed behaviour?' (Mitchell 1996). The answer, of course, is 'no', but it does not follow that comparison of animals' self-directed behaviour in mirror-present and mirror-absent conditions would be uninformative. Experiments cannot show definitively that animals lack any given capacity, and they are not designed to do so. In the face of a null result, it is always possible (although not necessarily persuasive) to argue that the wrong animals were tested, or that they were tested under the wrong conditions to reveal their potential. Rather, the outcomes of experiments can support or fail to support a hypothesis, and thereby render it more or less plausible, or rational to believe, that the animals in question have the capacity at issue. Thus, if a study were conducted in which repeated, careful observations of Mitchell's, or a chimpanzee's, behaviour revealed no difference in the form or frequency of grooming in the presence and the absence of a mirror, then we could conclude, not that Mitchell or the chimpanzee certainly cannot use a mirror, but that this study had provided no reason whatever to believe that they can.

Mitchell suggested that my anaesthetic artefact hypothesis is rendered implausible by the results of two experiments on self-recognition in primates (Robert 1986; Hyatt & Hopkins 1994). In the first, Hyatt & Hopkins (1994) reported that bonobos, *Pan paniscus*, engaged in more self-directed behaviour when the reflective side of a mirror was exposed to them than when the reflective surface was turned away. Since the anaesthetic hypothesis is a putative explanation for the results of mark-tests, and no marks were applied to the bonobos in Hyatt & Hopkins' study, the results have no bearing on the hypothesis. More generally, while the design of the study was such that it could have revealed an effect of a mirror on behaviour (see above), it neither achieved this, nor provided evidence of self-recognition in bonobos because (1) no data were presented, and (2) even if the putative effect occurred, it may have arisen from general arousal induced by exposure to a range of reflections. To show that self-recognition was responsible, rather than a more general mirror effect, it would be necessary to demonstrate that

the activating effect of the mirror was specific to self-directed behaviour, and dependent upon exposure to reflections of the subjects' own bodies.

In the second study claimed by Mitchell to undermine the anaesthetic hypothesis, Robert (1986) marked the faces of a chimpanzee and an orangutan while they were sleeping, and subsequently did not find that they touched the marks more when a mirror was present than when it was absent. This result is entirely consistent with the anaesthetic hypothesis, which suggests that positive mark test outcomes are due to the effects of recovery from anaesthesia in animals with a relatively high baseline frequency of face-touching behaviour (Heyes 1994a, 1995). Although Mitchell denies sharing their views, he uses the same tack as Gallup et al. (1995) in attempting to explain away Robert's data: he points out that her subjects were very young, and cites putative evidence that self-recognition does not usually emerge until later in ontogeny. However, neither the study cited by Mitchell (Lin et al. 1992) nor that cited by Gallup et al. (Povinelli et al. 1993) has demonstrated that a capacity for mirror self-recognition, rather than a propensity to engage in mirror-independent self-exploratory behaviour, increases with age (see Heyes 1994a, page 912; Heyes 1995, for details). Thus, the developmental account of Robert's data is unsubstantiated, and the anaesthetic hypothesis continues to offer a more straightforward explanation.

Mitchell (1996) alleges that I (Heyes 1994a) offered two explanations for self-recognition in apes, which were mutually contradictory and, for independent reasons, false. In fact, (1) the two hypotheses had distinct targets (self-recognition or mirror-guided body inspection versus mark test failure in monkeys) and therefore could not have been contradictory; (2) the first was offered as a part of a 'task description', a clarification of the cognitive demands of mirror self-recognition, rather than a comprehensive explanation; and (3) the evidence cited by Mitchell lacked even the potential to invalidate either hypothesis.

According to Mitchell, my first hypothesis was that all any animal needs to pass the mark test is to be able to distinguish, across a fairly broad range, sensory inputs resulting from the physical state and operations of its own body from sensory

inputs originating elsewhere. In fact, I suggested that 'to use a mirror as a source of information about its body', 'All that the animal needs to know about itself is that its body is distinct from the rest of the world' (Heyes 1994a, pp. 914–915, italics added). Thus, Mitchell misunderstood both the phenomenon that I was discussing (mark test success versus mirror-guided body inspection), and the breadth of my claim (all necessary knowledge versus all necessary self-knowledge). Since I was not claiming that a capacity to distinguish sensory feedback from inputs is sufficient for mark test success, Mitchell's evidence that 18-month-old human infants have this capacity and yet fail the mark test (Watson 1994) is irrelevant.

Mitchell took my second hypothesis to be that 'organisms fail the mark test because they do not recognize the contingency between their own movements and those of the mirror image', and declared it 'demonstrably false' because 5-month-old human infants can detect an intermodal proprioceptive-visual contingency, yet fail the mark test (Bahrick & Watson 1985), and monkeys 'appear to recognize the contingency between self and mirror (Boccia 1994)' (Mitchell 1996). In fact, this hypothesis was one of several which I discussed as possible reasons for mark test failure in monkeys (Heyes 1994a, pp. 916–917). In this case, I was pointing out that a necessary condition for mirror-guided body inspection is detection of the contingency between one's own body movements and those of the mirror image, and therefore a failure to detect this contingency is one possible cause for mark test failure. Since I was not claiming that this kind of contingency detection is sufficient for mark test success, or that its absence is always responsible for failure, Bahrick & Watson's (1985) findings have no bearing on the hypothesis. There is no reason to assume that the causes of mark test failure must be the same in human infants and monkeys or, come to that, constant within either of these groups.

In contrast, solid evidence of the relevant kind of contingency detection in monkeys would be informative, it would disconfirm my hypothesis, but Boccia (1994) did not provide such evidence. She found only that four out of 15 pigtail macaques, *Macaca nemestrina*, showed 'contingency testing' behaviour when in the presence of a mirror. That is, the observer thought he or she saw each of these animals 'looking intently at the

mirror image while slowly moving the head or a limb' (Boccia 1994, page 354). The reader was not told how many times, in how many hours of observation, the observer got this impression, or how often the slow movement occurred without orientation to the mirror. In short, there was no evidence, beyond the observer's impression that the animal was 'looking', that the slow movement was in any way related to the mirror.

Neither of the hypotheses discussed by Mitchell represents what might be described as my 'explanation' for mirror self-recognition. I came closest to offering an explanation when I said that 'mirror-guided body inspection involves the use of novel, displaced visual feedback to guide action' (Heyes 1994a, page 909), but, just as it is misleading to call mirror-guided body inspection 'self-recognition', it is misleading to describe this as an 'explanation'. The term self-recognition is unsatisfactory because, if it were shown, using a properly designed experiment (see Heyes 1995, for a suggested design), that an animal can use a mirror to derive information about its body, we would not automatically have evidence that the animal 'recognizes' itself, or even its body, in the mirror. Contrary to what has been claimed by Gallup (1982), Mitchell (1993) and Whiten & Byrne (1991), to use a mirror in this way it would not be necessary for the animal to recognize, or form a second-order representation of, the representational relationship between the mirror and its body. In principle, the animal could simply use the information available in the mirror image, without having any thoughts about why that information is available or useful.

I would hesitate to describe my characterization of mirror-guided body inspection as an explanation because, following Marr (1982) and others, I regard its formulation as an early, rather than a terminal, stage in the investigation. Even if I am right in thinking that the capacity to utilize novel, displaced visual feedback is the key cognitive requirement for mirror-guided body inspection, it would still be necessary to develop formal and informal models of how this feedback is used (e.g. of how the processing differs from that of familiar and/or direct visual feedback), and why, if at all, species differ in their capacity for its exploitation.

Preparation of the manuscript was supported by a grant from the Biotechnology and Biological Sciences Research Council. I am grateful to Chris

Mitchell and Phil Reed for their comments on an early draft of the paper.

### REFERENCES

- Bahrack, L. E. & Watson, J. S. 1985. Detection of intermodal proprioceptive-visual contingency as a potential basis of self-perception in infancy. *Dev. Psychol.*, **21**, 963-973.
- Boccia, M. L. 1994. Mirror behavior in macaques. In: *Self-awareness in Animals and Humans* (Ed. by S. T. Parker, R. W. Mitchell & M. L. Boccia), pp. 350-360. New York: Cambridge University Press.
- Gallup, G. G. 1970. Chimpanzees: self-recognition. *Science*, **167**, 86-87.
- Gallup, G. G. 1982. Self-awareness and the emergence of mind in primates. *Am. J. Primatol.*, **3**, 341-343.
- Gallup, G. G., Povinelli, D. J., Suarez, S. D., Anderson, J. R., Lethmate, J. & Menzel, E. W. 1995. Further reflections on self-recognition in primates. *Anim. Behav.*, **50**, 1525-1532.
- Heyes, C. M. 1994a. Reflections on self-recognition in primates. *Anim. Behav.*, **47**, 909-919.
- Heyes, C. M. 1994b. Social cognition in primates. In: *Animal Learning and Cognition* (Ed. by N. J. MacKintosh), pp. 281-305. New York: Academic Press.
- Heyes, C. M. 1995. Self-recognition in primates: further reflections create a hall of mirrors. *Anim. Behav.*, **50**, 1533-1542.
- Heyes, C. M. & Dickinson, A. 1990. The intentionality of animal action. *Mind Language*, **5**, 87-104.
- Hyatt, C. W. & Hopkins, W. D. 1994. Self-awareness in bonobos and chimpanzees: a comparative perspective. In: *Self-awareness in Animals and Humans* (Ed. by S. T. Parker, R. W. Mitchell & M. L. Boccia), pp. 248-253. New York: Cambridge University Press.
- Lin, A. C., Bard, K. A. & Anderson, J. R. 1992. Development of self-recognition in chimpanzees. *J. comp. Psychol.*, **106**, 120-127.
- Marr, D. 1982. *Vision*. New York: W. H. Freeman.
- Mitchell, R. W. 1993. Mental models of mirror-self-recognition: two theories. *New Ideas Psychol.*, **11**, 295-325.
- Mitchell, R. W. 1996. Self-recognition, methodology and explanation: a comment on Heyes (1994). *Anim. Behav.*, **51**, 467-469.
- Parker, S. T., Mitchell, R. W. & Boccia, M. L. (Eds) 1994. *Self-awareness in Animals and Humans*. New York: Cambridge University Press.
- Povinelli, D. J., Rulf, A. B., Landau, K. R. & Bierschwale, D. T. 1993. Self-recognition in chimpanzees: distribution, ontogeny, and patterns of emergence. *J. comp. Psychol.*, **107**, 347-372.
- Robert, S. 1986. Ontogeny of mirror behavior in two species of great apes. *Am. J. Primatol.*, **10**, 109-117.
- Watson, J. S. 1994. Detection of self: the perfect algorithm. In: *Self-awareness in Animals and Humans* (Ed. by S. T. Parker, R. W. Mitchell & M. L. Boccia), pp. 131-148. New York: Cambridge University Press.
- Whiten, A. & Byrne, R. W. 1991. The emergence of meta-representation in human ontogeny and primate phylogeny. In: *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading* (Ed. by A. Whiten), pp. 267-281. Oxford: Basil Blackwell.