

Life beyond the mirror: a reply to Anderson & Gallup

M. D. HAUSER*†‡ & J. KRALIK†

*Departments of Anthropology and †Psychology, ‡Program in Neurosciences, Harvard University

(Received 21 February 1997; initial acceptance 4 April 1997;
final acceptance 19 June 1997; MS. number: sc-1101)

We believe that none of the comments made by Anderson & Gallup (1997) diminish the force of our significant mirror test results on cotton-top tamarins, *Saguinus oedipus*, and the implications we draw from them. Furthermore, their critique suggests that a number of subtle, yet critical features of our experiment have been misunderstood. Consequently, we briefly reiterate some of the key design elements in our response, but urge the interested reader to see our original paper (Hauser et al. 1995).

Experimental Procedures and Data Analysis

Anderson & Gallup claim that we did not mention the fact that 'the original mark test was devised to verify what appeared obvious from earlier behaviour in the presence of the mirror, namely self-recognition as indicated by spontaneous use of the reflection to inspect otherwise visually inaccessible body parts'. The introductory paragraph of our original paper reads (page 10 811): 'Having observed chimpanzees use a mirror for self-directed behaviour, he [Gallup] developed an experimental procedure to test formally for self-recognition'. Anderson & Gallup appear to consider the pre-experimental observations of potential mirror-guided self-directed behaviour to be a critical component of the mirror dye-mark test. Although such observations are potentially significant, they are not critical for demonstrating self-directed behaviour. We believe that the controlled version of the mirror dye-mark test that we implemented is sufficient to demonstrate mirror-mediated self-directed behaviour.

Correspondence: M. D. Hauser, Departments of Anthropology & Psychology, Program in Neurosciences, Harvard University, 11 Divinity Avenue, Cambridge, MA 02138, U.S.A. (email: hauser@wjh.harvard.edu).

Salience

Salience is a complex problem, with many different components that have yet to be tested formally. For example, previous tests of salience were not, as far as we know, sensitive to morphological features that may be critical in species identification. In our experiment, we systematically altered the colour of the cotton-top tamarins' white hair. Although there are many species of tamarins, only the cotton-top has a massive tuft of white hair on the top of its head. In contrast to Anderson & Gallup's conclusions, the positive results from our experiment show that it is incorrect to dismiss salience as an important factor in a species' ability to pass the mirror test. Furthermore, salience was manipulated systematically in our tamarin experiment. Results showed that the traditional Gallup dye-mark was insufficiently salient to produce mirror-mediated head-touching or increased mirror-directed staring by the tamarins, but colour-dyed hair was. This direct test of salience was a primary component of our experimental design.

Scoring

Anderson & Gallup suggest that the details of our scoring procedure were too vague, thus making the interpretation unclear. Mirror-mediated body and head touching were operationally defined as follows. An animal was facing and staring into a mirror. The animal then touched a part of its body or a part of its head while continuing to face and stare into the mirror. When a tamarin touched its body while looking in the mirror, it was sitting on the nestbox or branch adjacent to the mirror and was 0–1 m away from the mirror. Even though these measures of behaviour are not elaborate, they are clear and cannot be 'meaningless', especially given the pattern of results obtained. Using these definitions, two independent observers (neither

Hauser nor Kralik), with high inter-observer reliability, scored from direct observations 13 head touches while looking in the mirror by individuals with prior mirror exposure and colour-dyed hair; no head touches while looking in the mirror were scored in any other condition by these observers. In addition, although we did not record any instances of a tamarin 'visually inspect[ing] and/or sniff[ing] the fingers after contacting the mark', we did report several instances of animals using the mirror to look at visually inaccessible body parts in addition to using the mirror to look at the reflection of their own body movements. These results show, therefore, that the observed self-touching and inspecting were responses to the animals' mirror reflection; and these self-touches and inspections cannot be dismissed, even in the face of Anderson & Gallup's unreviewed interpretations of some of the video records.

Controls

Anderson & Gallup's critique suggests that the logic of our experimental design has been misunderstood. We consider six of the most significant criticisms they raise; we encourage readers of this discussion to see Table 1 and pp. 10 811–10 812 of our original report for the details.

(1) Anderson & Gallup suggest that '... capture, separation and anaesthesia' may have caused distress in the tamarins, and 'the resultant high arousal is unlikely to be compatible with mirror-mediated self-inspection'. Subjects in our laboratory move between their home cages and transport boxes without any signs of distress. To administer the anaesthesia, subjects were restrained and this clearly caused some stress. Observations from our holding cage procedure revealed, however, that subjects were not stressed at the start of our experimental conditions. Moreover, the main results of head-touching while looking in the mirror and increased non-aggressive staring only in the coloured hair condition suggest that the tamarins were responding to their mirror reflection regardless of their state of arousal, which appeared to be the same across all experimental conditions.

(2) Anderson & Gallup suggest that because subjects stared into the mirror for longer durations, 'increased head-touching while looking in the mirror is simply an artefact of increased

overall attention towards the mirror'. During experimental conditions, we observed 57 cases of head-touching while not looking in the mirror, and 84% of these occurred during the coloured hair condition, by individuals with prior mirror exposure. Thus, the average rate of head-touching while not looking in the mirror for the remaining experimental conditions was 0.6/session. We agree with Anderson & Gallup then that attentional factors could be involved, but disagree with their claim that it is an experimental artefact. We suggest that self-directed mirror-mediated behaviour may emerge only if subjects can attend to the image and recognize the relevant details. If, as in previous studies of monkeys, subjects consider the image to reflect an opponent, and a potentially aggressive one, then this will interfere with their attentional system and thus interfere with their ability to recognize the relevant features of the image. Perhaps then, self-directed mirror-mediated behaviour may emerge only when subjects attend to the image. Moreover, it is critical to remember that the increase in attention in the tamarin experiment appeared to be due to the individual's own hair being dyed (e.g. colour on the mirror did not increase attention towards the mirror). Our proposal then is that there was increased attention to the mirror because there was increased attention to 'self'.

Anderson & Gallup further state that without data on head-touching while not looking in the mirror 'nothing can be said about the significance of the mirror's presence'. This is not quite correct. The original experiment reported that during baseline recording, the average rate of head-touching over a 20-min period was 0.2. Thus, mirror-mediated head-touching in the coloured hair condition versus all the other conditions in the experiment indicates that the head-touching occurred because the tamarins saw, via the mirror, that their own hair was dyed. This conclusion holds (a) even if the results are attention-mediated and (b) no matter what the results are for head touches while not looking in the mirror or touches to other parts of the body. Note, however, that there was not an increase in touches to other parts of the body, to other objects or to other tamarins in the coloured hair condition as opposed to the other experimental conditions. The increased touches were only to the individual's own head.

(3) It is not the case that our 'other mirror-present experimental conditions are irrelevant'.

Some of the results presented in Hauser et al. (1995) stem from a between-subjects comparison, and these are not only valid, but also extremely telling in terms of the patterns of behaviour observed. For example, one potentially critical factor was the feeling of dye in the hair. Were subjects touching their heads simply because they felt a difference in texture as opposed to seeing a change in the colour of their hair? All four subjects receiving the white hair dye (i.e. a texture change, but no colour change) failed to touch their heads while looking in the mirror, suggesting that the feeling of the dye on their hair was insufficient to elicit head-touching. Similarly, three subjects without prior mirror exposure, but with coloured hair, never touched their heads while looking in the mirror. It is, of course, theoretically possible that for the five subjects who did not receive the white hair dye condition, the feeling of the coloured hair dye alone would cause a response. However, the complete failure to elicit head-touching while looking in the mirror in the four subjects who received the white hair dye, and the three subjects who received coloured hair dye but no prior mirror exposure, makes this possibility unlikely (seven out of seven, sign test: $P=0.02$). The same statistical argument can be made for all of the other control conditions in the original experiment.

(4) In describing the pattern of results, Anderson & Gallup incorrectly state that: 'the two remaining head-touching subjects (M, N) showed only a single dyed-hair-touch each in the presence of the mirror, and the procedure was halted as soon as the "positive" response was recorded (see Table 1 in Hauser et al. 1995)'. Nowhere in the paper do we say that the procedure was halted. All subjects received a full 20-min sampling period, whether they touched their head or not.

(5) Mirror-mediated head touches by subjects L and F cannot be due to the 'control washing procedure', as Anderson & Gallup suggest. L and F did not receive the 10-min holding cage condition after anaesthesia, but these subjects were not allowed back into their home cage for observation until they appeared fully recovered. After L and F were tested, we decided to make the 10-min holding cage period mandatory to simulate and verify the recovery periods of L and F. In all of the remaining conditions of the experiment, subjects went through the general preparation procedure (including anaesthesia, body washing, etc.)

and the holding cage condition. Only subjects with prior mirror exposure and coloured hair touched their heads while looking in the mirror. These results suggest that the general preparation procedure did not cause any of the subjects, including L and F, to touch their heads while looking into the mirror.

(6) Finally, the statement that we failed to 'acknowledge the critical control period in which any responses directed to the subject's head before reinstatement of the mirror are recorded' is false. It is precisely because of previous work that our design included this critical control period (page 10 812); and no head touches were recorded during L and F's recovery period or during any holding cage condition throughout the experiment.

Staring

We respond here to two main criticisms. First, Anderson & Gallup consider non-aggressive staring, which occurs in other species who appear to fail the mirror test, to be irrelevant; but the only condition under which we observed long, non-aggressive staring was during mirror tests with individuals who had prior mirror exposure and coloured hair. Indeed, subjects stared longer and with individual stares of longer duration in this condition than in all other experimental conditions combined. Thus, our experiment itself provides evidence that staring *is* related to self-recognition.

Second, it is highly unlikely that the mirror staring was due to 'something else in the environment', as Anderson & Gallup suggest. The mirrors were positioned so that they would reflect only other parts of the cage and room, not other individuals. Furthermore, given that the mirror was in place for 3–4 weeks prior to such tests, why would subjects now, all of a sudden, and in a restricted 20-min observation period, show an increase in attention to environmental stimuli? In addition, nothing in the environment changed between the coloured hair condition and the control conditions; the only change was in the colour of the animal's hair. Thus, the positive results in the coloured hair condition versus all the other experimental conditions could not be due to 'something else in the environment'.

Self-recognition and the Gallup Mirror Test

We end by restating the main conclusion to be drawn from our original experiment: it is time to

be more specific about what cognitive functions the mirror test is actually testing. Several abilities appear to be necessary to pass the mirror test, including some form of self-recognition. It is still unclear, however, which of these abilities is most significant in distinguishing between species; and, as several studies have pointed out, it is also unclear why some individuals within a species pass the mirror test, whereas other individuals fail (Heyes, *in press*). For example, although young chimpanzees, *Pan troglodytes*, may be developmentally delayed with regard to crucial abilities for passing the mirror test, it is unclear why some older chimpanzees, who are still in their reproductive years, would fail (Povinelli *et al.* 1993).

One possible explanation for the inter- and intraspecific pattern of results is that the mirror test is an invalid measure of self-recognition (cf. Heyes 1994, *in press*). Species may fail the mirror test for a number of reasons, only one of which may be the lack of 'self' as a stimulus class. Reasons for failing the mirror test include salience of the dye-mark and significance of the trait as a species-typical marker, as we believe was shown in our tamarin experiment. Thus, many species could be failing Gallup's version of the mirror test even though they have self-recognition. If species who

fail Gallup's mirror test actually have the requisite self-recognition ability needed to pass the test, then the mirror test is measuring something other than self-recognition. If this is the case, it must be concluded that the Gallup mirror test is an invalid measure of self-recognition.

REFERENCES

- Anderson, J. R. & Gallup, G. G., Jr. 1997. Self-recognition in *Saguinus*? A critical essay. *Anim. Behav.*, **54**, 1563–1567.
- Gallup, G. G., Jr. 1970. Chimpanzees: self-recognition. *Science*, **167**, 86–87.
- Hauser, M. D., Kralik, J., Botto-Mahan, C., Garrett, M. & Oser, J. 1995. Self-recognition in primates: phylogeny and the salience of species-typical traits. *Proc. natn. Acad. Sci. U.S.A.*, **92**, 10 811–10 814.
- Heyes, C. M. 1994. Reflections on self-recognition in primates. *Anim. Behav.*, **47**, 909–919.
- Heyes, C. M. 1995. Self-recognition in primates: further reflections create a hall of mirrors. *Anim. Behav.*, **50**, 1533–1542.
- Heyes, C. M. *In press*. Theory of mind in nonhuman primates. *Behav. Brain Sci.*
- Povinelli, D. J., Rulf, A. B., Landau, K. R. & Bierschwale, D. T. 1993. Self-recognition in chimpanzees (*Pan troglodytes*): distribution, ontogeny, and patterns of emergence. *J. comp. Psychol.*, **107**, 347–372.