The evolution of primate visual self-recognition: evidence of absence in lesser apes

Thomas Suddendorf* and Emma Collier-Baker

School of Psychology, University of Queensland, Brisbane, Queensland 4072, Australia

Mirror self-recognition typically emerges in human children in the second year of life and has been documented in great apes. In contrast to monkeys, humans and great apes can use mirrors to inspect unusual marks on their body that cannot be seen directly. To study this further, Gallup (1970) exposed chimpanzees to mirrors, then placed a mark on their heads and observed their reaction upon re-exposure to the mirror image. Chimpanzees used the mirror to inspect the mark and this test has since been widely used to study visual self-recognition in a range of species. Amsterdam (1972) independently developed a similar task for human children. Toddlers begin to investigate their own head upon seeing the surprising mark in the mirror from around the age of 15 months, and by 24 months most children react in this manner (Amsterdam 1972; Nielsen et al. 2006). Passing the task has been argued to reflect self-awareness (Gallup 1998), but at a minimum appears to imply that subjects have a mental model of what they look like from the outside (Nielsen et al. 2006). A recent direct comparison of human and chimpanzee infants found that the development of self-recognition in our closest relatives is comparable (Bard et al. 2006).

Many primate species have been tested with versions of the basic task, but only members of the great ape species react as 24-month old children do (Gallup 1970; Lethmate & Ducker 1973; Suarez & Gallup 1981; Povinelli 1989; Anderson & Gallup 1997; Povinelli et al. 1997; Posada & Colell 2007). The competence of gorillas has been the most controversial with rigorous experimental attempts failing to find evidence (e.g. Suarez & Gallup 1981; Ledbetter & Basen 1982; Shillito et al. 1999). This led to the proposal that gorillas may have lost an ancestral great ape capacity for self-recognition (Povinelli 1993; Gallup 1997). However, positive results from five gorillas (out of a total of 15 tested) are now reported in the literature (Parker 1994; Patterson & Cohen 1994; Swartz & Evans 1994; Posada & Colell 2007). In fact, not all chimpanzees and orang-utans that have been tested passed the task either. In their review, Swartz et al. (1999) reported that 43 per cent of chimpanzees (42 out of 97) and 50 per cent of orang-utans (three out of six) passed the task. There are as yet no reports of formal mirror mark tests on bonobos in the literature, although three studies report self-directed behaviour in front of mirrors (Westergaard & Hyatt 1994; Walraven et al. 1995; Inoue-Nakamura 1997). The fact that not all great apes tested with the mirror mark test pass may mean that not all have the capacity for self-recognition (e.g. owing to differences in age, Povinelli et al. 1993; de Veer et al. 2003). Observed performance differences may, however, also reflect individual differences in motivation and differences in the methods and criteria employed (Bard et al. 2006). Although bonobos still need to be tested and only slightly less than half of all the individual great apes tested have passed, these data suggest that it is likely that the potential for mirror self-recognition was inherited from a common ancestor, because this hypothesis requires only one assumption about an evolutionary change (i.e. acquisition of the capacity by a common ancestor of the great apes and humans), whereas a model of convergent evolution would imply at least four separate acquisition events in the lines leading to modern great apes and humans (Suddendorf & Whiten 2001). Given recent molecular estimates (Wildman et al. 2003), this suggests that the trait is at least 13.8 Myr old (figure 1).

However, the trait may be substantially older. To establish an upper limit as to its first emergence, one needs to determine which closely related species form the ‘out-group’—that is, the species that do not share the
Ascertainment whether lesser apes do or do not have the capacity for visual self-recognition is not only important for establishing the time frame for phylogenetic emergence of the trait. It would also open the door for a comparative approach to the identification of the underlying neural substrate as well as its genetic underpinnings. This approach could deliver vital clues to the growing imaging data on human visual self-recognition (Devue et al. 2007; Platek et al. 2008). What do self-recognizing species have that non-self-recognizers do not? To determine whether lesser apes can or cannot recognize their mirror image, we conducted, to our knowledge, the largest study to date with a comprehensive set of novel control conditions and ample opportunity for gibbons to demonstrate their capacities.

2. MATERIAL AND METHODS

(a) Participants

Subjects represented three genera of gibbons (three Hylobates, seven Symphalangus and seven Nomascus) and were housed at four zoological parks in Australia and the USA (Perth Zoo, Adelaide Zoo, Gorge Wildlife Park and Smithsonian’s National Zoological Park). Species, age and sex of the 17 gibbons, we mark tested (mean age 15.88, range 5–37 years) are provided in Table 1. From an original sample of 20 gibbons, three (Hylobates leucogenys) could not be tested as they refused to approach the experimenter. All apes lived in pairs or family groups and were exposed to the mirrors in their normal group enclosures. The extent of previous experience with mirrors for each subject was difficult to establish as zoo visitors and previous keepers might have presented hand mirrors. However, from the available record, only one subject (Siam) had previously had a mirror in her enclosure (while in isolation for a period of approx. four months).

(b) Apparatus

Mirrors were constructed for each of the gibbon groups at Perth Zoo, Adelaide Zoo and Gorge Wildlife Park, and were 34 cm wide × 46 cm high sheets of mirror-finished stainless steel firmly secured to wooden backing. The mirror was attached perpendicular to the mesh for all gibbons except one white-cheeked gibbon (Phillip) and for two of the silvery gibbons’ exposure sessions in which the mirror was attached flush to the inside of the enclosure with cable ties (figure 2). A perpendicular angle was preferred because it allowed gibbons to sit on a bench in front of the mirror as well as look behind the mirror more easily.

One fixed mirror was used for all gibbons at Smithsonian’s National Zoological Park because the four gibbon groups were rotated through adjacent enclosures daily. This was a large plastic mirror (81 cm wide × 135 cm high) attached to a wooden backing and secured perpendicular to the mesh inside the enclosure next to a perch, so that subjects could sit in front of it. The mirror was covered by a wooden board attached with screws when subjects were not being tested (see the electronic supplementary material).

Commercial cake icing (white, green and red) was used for the pre- and post-test control conditions and non-toxic paint of matching colours for the mark tests. Paper stickers (1 cm in diameter) of various colours were also used in two post-tests.
Procedure

(i) Mirror exposure sessions

In their habitual groups, all gibbons were exposed to a mirror inside their enclosure over several days. Groups received a minimum of 5 hours of mirror exposure over three sessions. All mirror exposure sessions were videotaped (one gibbon, *Nomascus leucogenys*, was taped only for 2 hours 30 min). Responses oriented to the mirror during these sessions were recorded live or from videotape into mutually exclusive categories (see the electronic supplementary material).

In order to provide data that could be most usefully compared with previously tested species, we followed the behavioural categories of the only comparative primate mirror self-recognition study that included gibbons (Inoue-Nakamura 1997). We subdivided one of Inoue-Nakamura’s categories, exploration, into physical exploration of the mirror and reaching/looking behind the mirror. A behaviour needed to continue for at least 2 s with the subject’s face oriented to the mirror to be counted as one occurrence. A small quantity of the gibbons’ regular food (fruits or vegetables) was put in front of the mirror at the start of each session to encourage initial attention to the mirror.

(ii) Pre-test control conditions

Prior to testing, the experimenter (E.C.-B.) habituated subjects to touches on the head and on a limb over several days to reduce the likelihood of them noticing the placement of the mark. Marks were surreptitiously applied while subjects were feeding at the mesh. Apes were tested individually in order to make marking of less dominant gibbons easier and to avoid the possibility of the subject’s mark being removed by a conspecific. All gibbons were comfortable alone except for

Table 1. Summary of performance by subjects on pre-, mark and post-tests. (Subject demographics are provided in the subject column: male (M) and female (F); age in years; and gibbon locations represented as Perth Zoo (PZ); Adelaide Zoo (AZ); and Smithsonian’s National Zoological Park (SNZP) in the USA. After his initial test (1) one gibbon (Arjuna) was tested again 10 days later (2). Post-tests included a small amount of icing smeared on the mirror and on the subject’s head; the experimenter (E) emphasizing herself in the mirror; and placement of a sticker on the subject’s head or foot. Dashes (—) indicate that an intervention was not administered. Three subjects(*) inadvertently marked their own faces with icing or paint during the mark test and failed to remove it.)
3. RESULTS

During the mirror exposure sessions, apes most frequently engaged in close looking at the mirror \( (n=678, M=13.9, \text{s.d.}=13.59) \), followed by physical exploration \( (n=514, M=10.08, \text{s.d.}=20.54) \), reaching or looking behind the mirror \( (n=390, M=7.65, \text{s.d.}=12.21) \), social behaviour \( (n=303, M=5.94, \text{s.d.}=23.59) \), contingency behaviour \( (n=166, M=3.25, \text{s.d.}=8.20) \) and finally self-directed behaviour \( (n=23, M=0.45, \text{s.d.}=1.30) \). Every ape engaged in close looking at the mirror and all but two apes (\textit{Hylobates moloch} and \textit{N. leucogenys}) reached or looked behind the mirror (these two apes were the most cautious of the mirror and showed lower response rates across all categories relative to the other apes). This reaching behaviour may appear to an observer as if the subject was searching for ‘the other gibbon’ (figure 2). There were large individual differences in behavioural responses, but no significant correlations with age. Analyses of variance also yielded no significant differences between subadults (age 4–6 years), adults and older adults (30 and above) on any of the variables. All subjects demonstrated interest in the mirror. Analysis of a 5 min section of video tape selected at random from each groups’ exposure sessions revealed that they spent an average of 48 per cent of time at the mirror \( (M=145.1 \text{s}, \text{s.d.}=96.83 \text{s}) \).

Of almost 60 hours of behavioural recording during the mirror exposure sessions, there were only 23 instances of self-directed behaviour from seven of the gibbons while facing the mirror. Apart from some tongue manipulation displayed by one siamang (\textit{Symphalangus syndactylus}), that we included although her tongue may have been visible without the aid of the mirror, every other behaviour observed in the category of ‘manipulating a body part visible only in the mirror’ consisted of scratching or touching the head, face or back. It is possible that these instances reflect a capacity for self-recognition. However, none of these behaviours were followed by further self-directed manipulations, suggesting that they may have been merely coincidental with orientation to the mirror. The mark test gave gibbons the opportunity to provide more objective evidence.

All gibbons passed the pre-test control conditions that were designed to establish their motivation to inspect the mark. They all eagerly consumed the icing when it was offered to them by the experimenter, and they all inspected and consumed the mark when discovering the surreptitiously placed icing on their limb. This motivation was evident throughout, as subjects would frequently find and eat tiny smears of icing and paint which had been left on the mesh during pre-, mark or post-tests. Despite their demonstrated interest in the mirror, and in the icing, none of the apes showed any mark-directed behaviour, or indeed any touches to the head, in front of the mirrors during the mark test (table 1). One ape

(d) Data analysis

Touch marks could be coded unambiguously live and subsequently checked on videotape. Interobserver reliability for behavioural coding in the mirror exposure sessions was assessed through 84 randomly selected 1 min sections of video tape. Based on agreement of behaviour identification, interobserver agreement was 96.1 per cent.

Figure 2. A white-checked crested gibbon (\textit{Nomascus leucogenys}) reaching behind the mirror.

Siam and Sydney who became somewhat stressed upon separation. Therefore, these two animals were tested together but did not interfere with one another. All sessions were videotaped. To establish subjects’ motivation to retrieve a mark from their body, we first conducted two control tests. The apes were offered a small amount of visually salient icing selected to contrast with the subject’s coloration. We then surreptitiously placed icing on one of their limbs and observed their reaction upon chance discovery (cf. \textit{Gallup et al. 1980}).

(iii) Mark tests

In the experimental test, apes were surreptitiously marked above the brow with odourless, non-toxic paint of the same colour as the icing. Subjects were observed for 30 min while marked on the head with the mirror present. Mark touches were recorded live and checked on video. After being marked, subjects in Australian zoos \( (n=10) \) were given a 30 min mirror-absent condition (with the mirror turned around or covered) before the test condition, in order to control for discovery of the mark without the mirror. At Smithsonian’s National Zoological Park, the mirror-present test condition was given first.

(iv) Post-test controls

Following the mark test, we administered a range of further control tests designed to increase gibbons’ opportunities to demonstrate self-recognition (for details, see table 1 and §3).
...behaviour could reflect a false negative result, we provided the apes with post-tests that offered additional opportunities to demonstrate self-recognition (table 1). After the 30 min mark test session, all gibbons were again offered icing and they all consumed it once more. To ascertain that there were no auxiliary reasons for the mirror to inhibit mark-directed behaviour, we smeared the icing on the mirror surface itself. Fourteen subjects retrieved the visible icing on the mirror (i.e. by picking or licking it off), but all ignored entirely the larger mark of the same colour on their own head that was clearly visible in the mirror.

With considerable risk of inducing false positives, we continued to administer even more interventions in order to maximize subjects’ opportunity to demonstrate some competence (table 1). Administration of these additional post-tests was dictated by apes’ continued interest in the mirror and continued interest in approaching the experimenter. The tests included marking a limb once more, drawing attention to the experimenter behind the subject in the mirror, marking the subjects’ heads with icing or stickers, and finally marking their feet with a sticker. While all subjects that were marked on the limb (n = 6) or on the foot (n = 4) removed the mark or sticker immediately on discovery, none touched the icing, paint or sticker on their head. See the supplementary video for an example of a marked gibbon (N. leucogenys) reaching behind the mirror. Five gibbons also inadvertently marked their own faces with food during the course of the mirror testing or exposure sessions, yet none touched these marks.

One gibbon (H. moloch) showed particularly strong interest in the mirror. Thus, after the mark test, we left the mirror in his groups’ night enclosure for an additional 10 days and supplemented this with a second mirror in the day enclosure for the remaining 5 days (170 hours of access). We then conducted a second mark test but the subject again failed. Thus, none of our attempts to scaffold and increase subjects’ chances of recognizing themselves in the mirror resulted in any gibbon passing the mark test.

4. DISCUSSION
We did not find any evidence to suggest that lesser apes can recognize themselves in mirrors. The apes were clearly driven to find and retrieve a mark, and their continuing failure is hence evidence of absence rather than merely absence of evidence. In the face of substantially increased opportunity to show competence and continued failure to do so, it is now likely that lesser apes do not in fact have the capacity to recognize their mirror image. Absence of the capacity in the lesser apes and in monkeys suggests that the date of emergence falls after the line that led to modern gibbons split from the line that led to humans, and before the split from the line that led to modern orang-utans (figure 1). Recent analyses have placed this period between 18 Myr ago (Widdell & Penny 1996) and 13.8 Myr ago (Wildman et al. 2003). It remains possible that the common ancestor of all apes already had the capacity and the line leading to modern gibbons lost it. However, on the present data, it is more likely that the trait emerged after the split from the line leading to modern lesser apes, as that entails only one assumption about evolutionary change. By contrast, loss accounts require at a minimum two assumptions of change (i.e. the emergence in an earlier ancestor and the later loss in one or several lines).

One can only speculate as to why these ancestors may have first evolved this capacity. But phylogenetic reconstruction is a powerful tool that allows us to reason about the presence of cognitive capacities of long extinct ancestors even without ever laying eyes on a fossil. The large amount of data that exists on the capacities of our great ape relatives can be usefully employed in this way. Great apes share a range of relatively complex cognitive skills that seem related to the capacity for mirror self-recognition (e.g. they can pass Piagetian invisible displacement tasks, Call 2001; Collier-Baker & Suddendorf 2006) and whose phylogenetic history can be mapped (Suddendorf & Whiten 2001; Whiten & Suddendorf 2007). It is comprehensive data on the lesser apes that are largely missing from the comparative picture.

Yet, such mapping can inform the search for the neuronal and genetic underpinnings of the traits. Homologous traits are based on similar mechanisms and thus point to what neuronal or genetic characteristics are shared among the species. Identification of the out-group, the closest relatives that do not share the trait, provides important additional clues. Whatever neuronal and genetic characteristics the out-group shares with the species that have the trait are clearly not sufficient to produce it. Necessary factors underpinning the trait must hence be found among the genetic and neuronal characteristics that are not shared with the out-group, but are present in all species that have the trait. Such a comparative approach has great potential, but depends on the difficult task of establishing not only which species do, but also which species do not share the trait in question.

On a gross anatomical level, primate brains appear to differ little from each other, except in size (Semendeferi & Damasio 2000). But on a micro level there are differences with potentially significant consequences (Balter 2007). For example, humans and the great apes have spindle cells, large bipolar neurons located within layer five of the anterior cingulate gyrus and the fronto-insular cortex, whereas the other primates, including the lesser apes, do not (Nimchinsky et al. 1999). These neurons hence fit the comparative criteria that mark characteristics as potentially necessary for mirror self-recognition. Closer examination of their function in this context is warranted. Examination of the genetic basis of mental capacities can similarly benefit from such comparative criteria, as progress in mapping the genomes of apes (Mikkelsen 2005) will identify what precisely is and is not shared. Although it is too early to draw immediate conclusions about the neuronal or genetic underpinnings of visual self-recognition, the present evidence that lesser apes do not share this trait with great apes narrows down the search space considerably.

If other mammals have evolved self-recognition through convergent evolution, more extensive comparative analysis could provide additional clues. However, the underlying mechanisms involved may be fundamentally
different. Positive results for one elephant (Plotnik et al. 2006), two magpies (Prior et al. 2008) and one dolphin (Reiss & Marino 2001) have been reported in the literature. Among primates, only the descendents of a humanoid that probably lived between 13.8 and 18 Myr ago have so far reliably demonstrated that they know who it is that who looks back at them when they look in a mirror.

This research was approved by, and met the ethics guidelines of, the University of Queensland Animal Ethics Committee, the Wildlife Ethics Committee of the Department for Environment and Heritage, South Australia, the Perth Zoological Parks Authority Animal Research and Ethics Committees (Australia), and the Institutional Animal Care and Use Committee (USA).

This work was supported by an Australian Research Council Discovery grant and a Queensland-Smithsonian fellowship. We are grateful for the support of Adelaide Zoo, Perth Zoo, Gorge Wildlife Park, Smithsonian’s National Zoological Park and to Andrew Hill, Philip Kirk, Mauricio Perez, Ifi Flores and Stacey Enneking for valuable assistance with data collection and analysis.

REFERENCES


Collier-Baker, E. & Suddendorf, T. 2006 Do chimpanzees (Pan troglodytes) and 2-year-old children (Homo sapiens) understand double invisible displacement? J. Comp. Psychol. 120, 89–97. (doi:10.1037/0735-7036.120.2.89)


