

Self-Recognition in Chimpanzees (*Pan troglodytes*): Distribution, Ontogeny, and Patterns of Emergence

Daniel J. Povinelli, Alyssa B. Rulf, Keli R. Landau, and Donna T. Bierschwale

Investigations of mirror self-recognition (SR) in chimpanzees (*Pan troglodytes*) have had small samples and divergent methods. In Experiment 1, 105 chimpanzees (10 months to 40 years of age) were observed for signs of SR across 5 days of continuous mirror exposure. In Experiments 2 and 3, negative SR adult and adolescent chimpanzees were saturated with mirror exposure in efforts to facilitate SR and a longitudinal study was conducted with a number of young subjects. In Experiment 4, mark tests were administered to groups of positive SR, negative SR, and ambiguous SR subjects. In Experiment 5, we explored whether previous positive SR reports in young chimpanzees were artifacts of increased arousal during mirror exposure. Results suggest that SR typically emerges at 4.5-8 years of age, at the population level the capacity declines in adulthood, and in group settings SR typically occurs within minutes of a subject's exposure to a mirror.

Nearly 2½ decades have passed since Gallup (1970) demonstrated that chimpanzees are capable of recognizing themselves in mirrors. He reported that on initial introduction to a mirror, his chimpanzees tended to react as if confronted by another chimpanzee and displayed a variety of social and aggressive behaviors toward the images. After 2-3 days, however, these social reactions declined and were replaced by apparent instances of using the mirror to explore parts of their bodies that they had never had the opportunity to see before (e.g., picking at their teeth, exploring their anal-genital areas, and extracting mucus from their nose and eyes). In addition, after anesthetizing the subjects and applying a bright-red mark to their eyebrow ridge and ear, Gallup (1970) found that the chimpanzees made no effort to touch the marks until after the mirror was reintroduced. Gallup and his colleagues, as well as other independent investigators, have rep-

licated various aspects of the effects in the 1970 report and have extended the findings to include orangutans (Calhoun & Thompson, 1988; Gallup, McClure, Hill, & Bundy, 1971; Lethmate & Dücker, 1973; Suarez & Gallup, 1981).

In contrast to the effects reported in chimpanzees and orangutans, numerous investigators (using several different techniques) have failed to produce similar results in gorillas, lesser apes, Old and New World monkeys, and a variety of other species (see Anderson, 1984, Gallup, 1991, and Povinelli, 1987, for extended reviews of the failures to produce self-recognition in other species). Gorillas appear to represent an anomaly, given their phylogenetic relationship to the other three species (humans, chimpanzees, and orangutans) that do show evidence of self-recognition. However, several systematic attempts to demonstrate self-recognition in lowland gorillas have failed to produce positive evidence for this species (Ledbetter & Basen, 1982; Suarez & Gallup, 1981). Parker (1991) observed at least 1 or 2 gorillas make contingent facial movements while in front of mirrors, and she concluded that the gorillas therefore showed evidence of self-recognition. Patterson (1984) offered compelling evidence of self-recognition by a home-raised gorilla that was involved in a variety of forms of cognitive tutoring from infancy. Although the definitive data are yet to emerge, Povinelli (in press) argued that gorillas may have undergone a heterochronic shift in maturation that has resulted in the secondary loss of the underlying cognitive structures for self-recognition but that these structures may emerge as atavisms if gorillas are reared in ways that involve abnormal inputs during early ontogeny.

In the face of witnessing chimpanzees' apparent success in identifying the significance of their mirror-image and the failure of two monkey species to do so, Gallup (1970) tentatively proposed that self-recognition in mirrors may imply the presence of some form of self-concept and that the capacity may be restricted to the great ape-human clade. Gallup (1975) later expanded this argument and proposed that self-recognition in mirrors may be impossible in the absence of self-awareness, which he originally defined as the capacity of an organism to become the object of its own attention.

Daniel J. Povinelli, Alyssa B. Rulf, Keli R. Landau, and Donna T. Bierschwale, Laboratory of Comparative Behavioral Biology, New Iberia Research Center, University of Southwestern Louisiana.

This research was supported by National Institutes of Health Grant RR-03583-05 to the University of Southwestern Louisiana New Iberia Research Center.

We thank Anthony Rideaux, without whose assistance in the training and testing of the chimpanzees and without whose logistical support the experiments reported in this article would not have been possible. Daniel Boutte generously provided the technical support for the testing of adult subjects, and John Hardcastle and Jeff Rowell provided key practical advice on animal management and training and offered continuous encouragement throughout the project. We are also especially grateful to William Greer, the Director of the New Iberia Research Center, who supported this research in innumerable ways. We thank Helen K. Perilloux for assistance in data collection and analysis. Emil Menzel, John Flavell, Melinda Novak, and two anonymous reviewers provided helpful ideas for the final text.

Correspondence concerning this article should be addressed to Daniel J. Povinelli, Laboratory of Comparative Behavioral Biology, University of Southwestern Louisiana New Iberia Research Center, 100 Avenue D, New Iberia, Louisiana 70560.

Lewis and Brooks-Gunn (1979) proposed a similar view about the relationship between the development of objective self-awareness and mirror self-recognition in young children. Mitchell (in press) and Parker (1991) have recently challenged this assumption and have offered other models to explain the emergence of mirror self-recognition. These explanations have relied only minimally (or not at all) on the presence of the form of self-identity that Gallup (1982) and Lewis (1986) have argued underlies self-recognition. Gallup and Povinelli (in press) outlined some theoretical objections to these positions and have also specified a number of ways in which some of the models can be tested under very straight-forward procedures, but so far the issue remains open.

In addition to its implications for self-awareness, the topic of self-recognition in primates has generated intriguing predictions (and data) about its possible relationship to broader metacognitive processes, such as those implicated in theory of mind research (see Gallup, 1982, 1985, 1991; Lewis, Sullivan, Stanger, & Weiss, 1989; Povinelli & deBlois, 1992; Povinelli, Nelson, & Boysen, 1990, 1992; Povinelli, Parks, & Novak, 1991, 1992). However, before these models can be properly evaluated, a number of rather basic issues related to the chimpanzee's capacity for self-recognition must first be resolved. For example, despite the repeated replication of the basic findings, little is known about the ontogeny, population distribution, and patterns of emergence of self-recognition in chimpanzees.

Developmental Effects

Human children typically show evidence of mirror self-recognition by 18–24 months of age, although some researchers suggest that it may occur as early as 15 months in some children (Amsterdam, 1972; Bertenthal & Fischer, 1978; Johnson, 1983; Lewis & Brooks-Gunn, 1979; Schulman & Kaplowitz, 1977). Much of the work suffers from limitations that are overcome by Gallup's (1970) paradigm, which was independently developed for use with nonhuman primates (see Anderson, 1984, and Gallup, in press, for critical comparisons of the two approaches). Nonetheless, there appears to be reasonably coherent and convergent evidence with procedures involving mirrors, videotape images, and photographs, that by about 2 years of age, many human children are capable of self-recognition (review by Lewis, 1986). However, a careful examination of the literature consistently reveals that there appears to be a wide variation in the exact age at which children pass traditional tests of self-recognition.

In contrast to the human data, the developmental course of chimpanzee self-recognition is currently unclear. The original report by Gallup (1970) is unhelpful in this regard because the age of the subjects tested was not known. However, subsequent studies bear at least indirectly on the issue. For example, although they were not interested in the issue of ontogeny, Hill, Bundy, Gallup, and McClure (1970) investigated the reactions of three 18- to 20-month-old nursery-reared chimpanzees by using the 10-day exposure technique developed by Gallup (1970). They reported suggestive (but

not compelling) evidence of self-recognition in 2 of the subjects. Gallup et al. (1971) studied an additional 6 subjects that ranged in estimated age from 3 to 6 years of age. Some showed self-recognition, and some did not, but with respect to ontogeny, these results were confounded by rearing history. Robert (1986) reported failing to find mirror-directed self-exploratory behaviors in an 11-month-old chimpanzee after extended access to a large mirror. Calhoun and Thompson (1988) reported evidence of self-recognition in 2 chimpanzees that were estimated to be between 3 and 4 years of age; however, such ambiguous behaviors as "tugging on the shoulder hair while watching the action in the mirror" (p. 364) were classified as indicating self-recognition. More recently, Lin, Bard, and Anderson (1992) intermittently exposed 12 chimpanzees to mirrors while they were between the ages of 10 months and 2.5 years of age and later tested them for self-recognition when they were between the ages of 1.3 and 5.1 years. Lin et al. concluded that by about 2.5 years of age, chimpanzees show evidence of self-recognition.

Of the studies that we review, the developmental investigation by Lin et al. (1992) is clearly the most comprehensive. Unfortunately, it suffers from a number of limitations. First, the behavioral evidence used to indicate self-recognition was ambiguous. Lin et al. defined several categories of "mirror-guided behaviors" including "face-directed," for which they offered "the chimpanzee watches itself scratch its face" (p. 122, Table 1) as an example. From the standpoint of inferring self-recognition, this is problematic because it fails to distinguish between self-directed activity (which occurs at some baseline level with or without a mirror present) and self-exploratory behavior (which can be defined as apparent instances of using a mirror to gain otherwise inaccessible information about the self). The latter category, self-exploratory behavior, is a more conservative and more reasonable measure of whether an animal truly understands the source of its reflection. (We must point out that this problem is not unique to Lin et al.'s investigation but rather has been present, at least to some extent, in most previous investigations.)

In addition, Lin et al. (1992) defined self-recognition as a chimpanzee's touching a wet mark that was surreptitiously applied to the animal's forehead while the animal was conscious. They defend this method because of its widespread use among developmental psychologists, despite the fact that as implemented by Lin et al., the procedure has been extensively (and repeatedly) critiqued (Anderson, 1984; Gallup, in press). Such a sham-marking method is acceptable if a control session is included in which the subject is observed before the reintroduction of the mirror so that baseline levels of mark-directed activity can be assessed (e.g., Anderson, 1983; Lewis & Brooks-Gunn, 1979; Povinelli, 1989). Lin et al. included no such control sessions in their mark tests. Instead, they marked the animals, introduced the mirror, and then compared the number of times the subjects touched the marks while they were looking in the mirror to the number of times they touched the mark while they were not looking into the mirror. It is therefore impossible to know if the animals discovered the marks on their own or with the aid of the mirror. Indeed, when one sums the average number of mark

touches in the two categories (looking vs. not looking) across the age range that Lin et al. interpret as showing evidence of self-recognition (2.5–5 years), the numbers do not differ dramatically (for not looking in the mirror, 20.5 touches, and for looking in the mirror, 25.5 touches; data reconstructed from Lin et al., 1992, Figure 6). Finally, Lin et al. marked the chimpanzees in the presence of cagemates, which raises the distinct possibility that the unmarked animals provided a number of cues as to the presence of the mark.

Patterns of Emergence

Gallup (1970), as well as several other investigators, reported that self-exploratory responses in chimpanzees emerge after two to three 8-hr days of exposure to mirrors. This finding has always been somewhat puzzling, especially when coupled with the assumption that self-recognition in humans and chimpanzees springs from similar psychological sources (Gallup, 1975, 1982). There are two lines of evidence that suggest that additional research is needed on this issue. First, although congenitally blind patients who have had their sight surgically restored initially react to mirrors as if they were seeing another person, they apparently rather quickly discover the duality between themselves and their mirror-image (von Senden, 1960). In addition, a report by Priel and de Schonen (1986) suggested that children who have never had access to mirrors may recognize themselves after less than 10 min of exposure. Although there are some key methodological limitations to the study (i.e., whether or not they were truly mirror-naïve; see Gallup & Povinelli, in press), it remains the only systematic attempt to explore the effects of prior mirror experience on the development of self-recognition in children.

Although chimpanzees are reported to take up to 3 days before self-recognition patterns emerge, there are at least two reasons to suspect that this may not reflect the true pattern of emergence in the species. First, before his original landmark chimpanzee self-recognition study, Gallup (1968) informally described the rapid onset of self-exploratory responses in a single chimpanzee that he provided access to a mirror. He reported that although the chimpanzee initially reacted with a variety of social gestures including lip smacking, head bobbing, and threats, within 3–4 min the subject began to engage in a variety of exaggerated facial gestures and self-directed behaviors including an inspection of its eyebrows and forehead. Second, although the sampling procedures that have traditionally been used to observe the ways in which chimpanzees (and other nonhuman primates) react to mirrors across time are perfectly acceptable for obtaining reliable estimates of their viewing time and average number of social and self-directed responses, they are less well-suited for capturing the exact time of onset of self-recognition. Subjects are typically observed for 30 min a day, divided into two 15-min sessions spread out across the day. Given that visual interest in the mirror wanes quickly, it may be that the initial onset of self-recognition has typically been missed in previous investigations. Thus, it remains quite possible that any differences in the pattern of emergence of self-recognition that appear to be present between chimpanzees and orangu-

tans on the one hand and humans on the other is an artifact of inadequate sampling techniques.

Population Distribution

Swartz and Evans (1991) recently challenged the assumption that most chimpanzees recognize themselves. Using procedures analogous to but differing in important ways from Gallup's (1970) protocol, they reported that only 1–3 chimpanzees out of the 11 wild-born subjects they tested showed positive evidence of self-recognition. The exact ages of the subjects were not known but were estimated (from tooth eruption data) to be from 4 to 19 years. They conclude that not all chimpanzees possess the cognitive capacities to recognize themselves in mirrors and that not all animals that show self-directed responding in front of mirrors inspect a mark that is experimentally applied to their eyebrow ridge and ear. Neither Gallup's (1970) original investigation nor subsequent work by Gallup et al. (1971) has provided data on individual subjects, but clearly Gallup's impression was that although there was individual variation in the amount of self-directed responding, most of the chimpanzees tested recognized themselves. However, Gallup et al. (1971) emphasized, in the context of investigating the effects of differential rearing, that not all chimpanzees necessarily recognize themselves. Gallup (in press) criticized several aspects of the experimental design of Swartz and Evans (1991). Further research is necessary before conclusions can be drawn about population-level polymorphisms with respect to self-recognition.

In this series of experiments, we addressed the issues described earlier in a large population of both wild- and captive-born (including mother- and nursery-peer-reared) chimpanzees. In particular, we focused on three questions. First, with proper diagnostic techniques we sought to characterize the normal age range for the ontogeny of self-recognition in young mother- and nursery-peer-raised chimpanzees. Second, with sampling techniques explicitly designed for capturing the time of onset of self-recognition, we sought to characterize the pattern of emergence of self-recognition in chimpanzees during their initial exposure to a mirror. Third, by sampling a large population of chimpanzees, we hoped to better understand the extent to which the capacity for self-recognition is distributed within the species. In addition, we examined several other issues including the relation between the presence of contingent body and facial movements on the one hand and self-exploratory behaviors and the results of mark tests on the other. We conducted additional experiments to explore possible reasons for discrepant findings from previous investigators.

Experiment 1

In the first experiment we examined the reactions of 105 socially housed chimpanzees to mirrors. Information on the onset and quantity of contingent body and facial movements and self-exploratory behaviors (as well as visual interest and social reactions) were collected with several techniques.

Method

Subjects

The subjects were 105 chimpanzees (*Pan troglodytes*) housed at the University of Southwestern Louisiana New Iberia Research Center as part of the center's breeding colony. The chimpanzees were maintained in 12 large social groups of 6–13 animals and had continuous access to large indoor–outdoor living areas (outdoor areas, 6 m wide × 4 m high × 8.5 m deep; indoor sleeping areas, 6 × 2.5 × 6 m). The subjects ranged in age from 10 months to 40 years of age. Six of the cages contained nursery–peer-reared animals that were divided by age: 2- to 3-year-olds ($n = 12$ and 13; two groups were tested), 3- to 4-year-olds ($n = 8$), 5- to 6-year-olds ($n = 13$), and 7- to 8-year-olds ($n = 7$). The remainder of the groups were breeding units of 1 or 2 adult males and 4–6 females with 0–2 offspring, which ranged from newborn to 2.5 years of age. Twenty-eight of the adults were wild-born, 2 were born in captivity, and the origins of 4 were unknown. Some of the subjects were mother-reared, some were partially mother-reared before being transferred to the nursery, and the remainder were nursery–peer-raised from birth. As far as could be determined, all of the subjects were mirror-naive.

Apparatus and Procedure

A large Plexiglas mirror (105 × 241 cm) was mounted on heavy plywood and placed into a wooden frame and served as the stimulus material.

Each group of animals was tested for 5 days (except as noted) in the following procedure. On Day 1, the mirror was placed in front (and outside) of the animals' cage just out of their reach and left for 5–8 hr, depending on the weather and cleaning schedule. Immediately on the mirror's placement in front of the cage, 1 or 2 observers (depending on the number of animals in the cage) sat on either side of the mirror, facing the subjects and equipped with data sheets and a 15-s beeper. Two methods of data collection were used. First, a mean modified-frequency focal-animal sampling procedure was used to obtain an estimation of the duration and frequency of a number of behaviors (described later) that the subjects displayed toward the mirrors. Each 5-min focal-animal sample was divided into twenty 15-s intervals, and the number and sequence of separate bouts of each behavior was recorded in each interval. We collected 5 samples a day on each subject, except in one of the 2- to 3-year-old groups and the 3- to 4-year-old group, which were sampled 4 times a day. The samples were collected according to a randomized schedule and were typically spread out across the late morning and afternoon. On average each cage contained 9 animals; thus an observer was present and observing subjects for almost 4 hr a day.

We recorded the following behaviors: proximity, visual exploration (VE), fear, social behavior, aggression, looking behind, self-exploratory (SE) behavior, contingent facial (CF) movements, contingent body (CB) movements, and out-of-sight. The categories were derived and modified from previous investigations (Gallup, 1970; Lin et al., 1992) and are defined in Table 1. For SE, CF, and CB, the observers were instructed to code both instances that met the definitions of the ethogram with which they had been trained to reliability (see later description) and weak instances of these

Table 1
Self-Recognition Ethogram

Behavior	Description
Proximity	Animal was within predefined three-dimensional area (2 × 6 × 2 m) near the mirror but was not visually oriented toward the reflection.
Visual exploration	Animal both actively and passively made visual exploration of the mirror. This was recorded on all occasions when an animal was in proximity to the mirror and its eyes were oriented toward the mirror. Animal may or may not have been looking at its own image.
Fear	Animal grimaced and made agitated responses (rocking, whimpering, etc.).
Social behavior	Animal directed social behaviors (e.g., sexual presentations, play invitations, begging gestures, head bobbing, or lip flips) at the mirror.
Looking behind	Animal attempted to look over or behind the mirror.
Aggression	Animal made aggressive displays (e.g., swaggering, charging with hair erect, or shaking fence) at the mirror.
Contingent body movements	Animal made two or more repetitions of a bodily movement while engaged in visual exploration; this often included rapid visual alternation from the body part to the mirror. (Examples are slowly waving hand, poking fingers repetitively through the mesh, moving slowly forward and backward, moving one leg slowly while standing on the other foot.)
Contingent facial movements	Animal made unusual facial contortions while it was judged to be looking at its own image.
Self-exploration	Animal used fingers or hands to manipulate parts of the body otherwise not visible (e.g., facial areas and anal-genital region). Animal had to be looking into mirror and must have been judged to be looking at its own image.
Out-of-sight	Animal was not in proximity to the mirror.

behaviors, which did not meet the formal definitions. Figure 1 provides examples of compelling instances of each of these behaviors.

Examples of weak CF behaviors included certain instances of an animal's lip smacking, sucking on the caging wire while manipulating the lips or tongue, and alternately glancing into the mirror, as well as mild facial contortions seen in other contexts. Examples of weak bouts of CB included an animal's waving its arms in front of the mirror without looking continuously into the mirror or watching itself in the mirror while performing some nonrepetitive action.

Finally, weak bouts of SE were recorded in cases when a chimpanzee was observed scratching its face, head, nose, or body while engaging in VE of the mirror and in cases when a subject actively groomed a part of its body that it could see directly while it alternately glanced into the mirror.

A focal-mirror sampling procedure was also used to record all instances of SE, CF, and CB behaviors not captured on focal-subject records. The subject, type of behavior and a brief description, estimated duration of the behavior in seconds, and the time of oc-

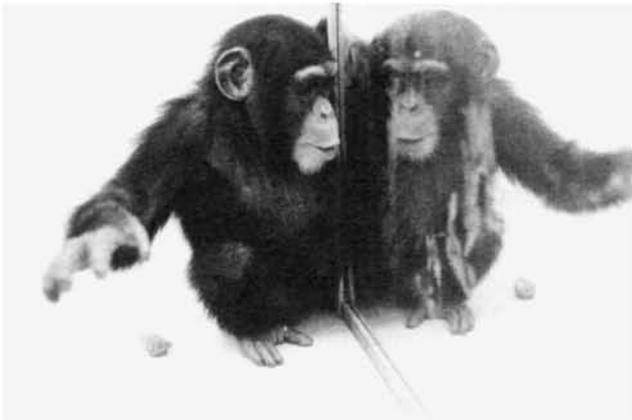


Figure 1. Examples of contingent body (top) and facial (middle) movements and self-exploratory behavior (bottom).

currence were recorded on a separate data sheet. In addition, as described earlier, it was noted for each instance if it met the definition of the ethogram (was compelling) or if it was weak. On Day 1, an observer was continuously present from the time that the mirror was introduced until at least 3 hr had elapsed and recorded all CB, CF, and SE behaviors in the focal-mirror technique. Otherwise, the focal-mirror samples were collected simultaneously with the focal-animal samples.

Two groups of 2- to 3-year-olds were tested. The original group ($n = 12$) was followed for 5 days under the procedures described earlier. The second group ($n = 13$) was studied with identical procedures except that focal-mirror and focal-animal data were collected for 2 days only. All other groups were tested for 5 days.

Reliability

Four observers participated in data collection. The observers had previously been trained to use a focal-animal, mean modified-frequency scoring system designed for collecting data on homeage behavior. The observers were retrained for this study by introducing them to the ethogram described in Table 1. All observers were trained daily for a month to use the new ethogram on both live and videotaped chimpanzees reacting to mirrors. The final 2 weeks of training involved controlled reliability checks on videotapes of chimpanzees. The following formula was used to calculate reliability: Percentage of Agreement = Total Interval Agreement ÷ Total Opportunities for Agreement. The observers were trained to 80% or better criterion during a total of 400–500 intervals, for which Daniel J. Povinelli served as the standard for comparison. During the final reliability sessions of 60 intervals, the three observers agreed with the primary observer 95%, 85%, and 86% of the time. Interobserver correlations for individual categories ranged from .80 to .98.

Results

The results are presented in two parts. First, we describe the results of the focal-mirror samples, which we used to classify those animals that recognized themselves and those that did not. Second, we examine the results of the focal-animal samples of the subjects across the 5 days of mirror exposure by group and age and in relation to whether an animal was classified as having recognized itself or not.

Classification of Self-Recognizing Animals

As a preliminary step in summarizing the data, we analyzed the focal-mirror records for the 92 subjects that were given 5 continuous days of mirror exposure. We classified all animals that showed at least 5 compelling instances of SE behavior (in total at least 30 s) as having shown positive evidence of self-recognition (SR+). Animals that showed from 1 to 4 instances (in total less than 30 s) were classified as ambiguous cases (SR?). Finally, those subjects that showed no compelling instances of SE behavior were classified as negative cases (SR-). Of the 92 chimpanzees tested, 21 were classified as SR+, 9 were classified as SR?, and 62 were classified as SR-. In addition, we examined both the focal-mirror and focal-animal records for the second group of 2- to 3-year-olds we observed for 2 days. All 13 subjects were classified as SR-. Because they received only 2 days

of exposure, these 13 subjects are not included in the following analyses, except as noted.

Compelling episodes. Figure 2 provides a summary of the mean number of bouts and mean total duration (in seconds) for both compelling and weak episodes for CB, CF, and SE behaviors for each subject across the 5 days of the study, with data derived from the focal-mirror samples. A quick inspection of the two left panels of Figure 2 shows that those animals that met our classification criteria for SR+ showed almost 5.5 times as many instances of SE behavior and that they averaged 30 times longer durations than the SR? animals. By definition, the SR- animals showed no instances of SE behavior that met our formal ethogram definitions. The same relative pattern holds true for CB and CF behaviors as well; this is important because there is nothing inherent about the data collection procedures that requires these measures be correlated (see later discussion). However, for the SR+ group, the absolute frequencies and durations of CB and CF were markedly lower than SE behavior.

In order to provide an independent confirmation of our classification scheme for the recognition status of each animal (based on the focal-mirror data), we examined the focal-animal samples and separately calculated the average number of intervals within a 5-min sample in which each subject

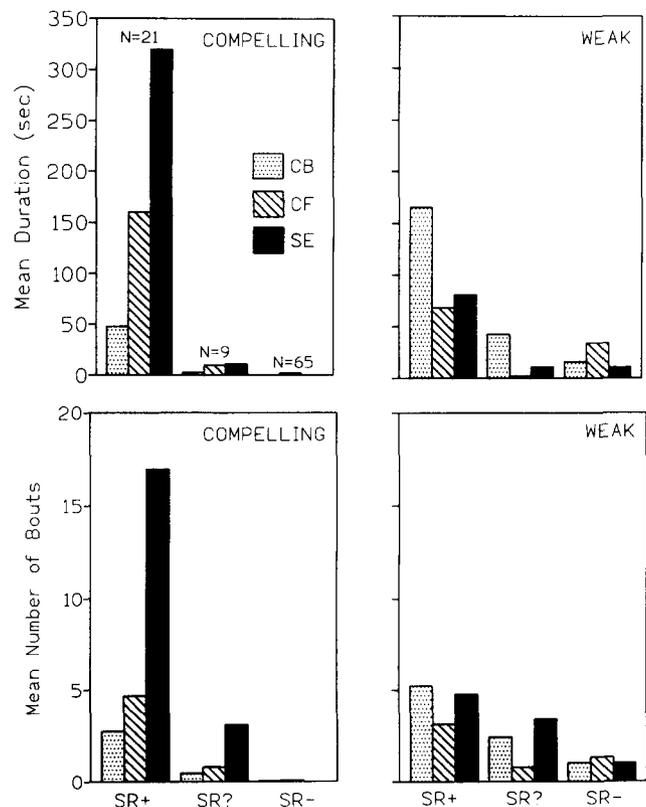


Figure 2. Mean number and duration of compelling and weak bouts of contingent body (CB) and contingent facial (CF) movements and self-exploratory (SE) behavior for chimpanzees that tested positive, negative, and ambiguous for self-recognition (SR+, SR-, and SR?, respectively).

displayed CB, CF, and SE behavior. Next, the animals were divided into the three groups (SR+, SR?, and SR-) that were constructed from the focal-mirror data. Separate one-way Group \times Behavior analyses of variance (ANOVAs) were conducted for CB, CF, and SE.

The results for CB revealed an overall significant effect, $F(2, 89) = 15.26, p < .0001$, and post hoc comparisons (Tukey's honestly significant difference tests) revealed that the SR+ animals showed significantly higher levels of CB ($M = 0.362$) than both the SR? ($M = 0.084$) and SR- ($M = 0.072$), which did not differ from each other. Similarly, the results for CF revealed an overall effect, $F(2, 89) = 12.49, p < .0001$, and the post hoc comparisons again revealed that the effect was because the SR+ animals showed significantly higher levels of the behavior ($M = 0.223$) than the SR- and SR? groups ($M_s = 0.037$ and 0.022 , respectively), which did not differ from each other. Finally, the results for SE indicated the same pattern. An overall effect was found, $F(2, 89) = 25.34, p < .0001$, and the post hoc tests confirmed that the SR+ animals showed higher levels of SE ($M = 0.499$) than the SR- subjects ($M = 0.021$) and the SR? subjects ($M = 0.116$), which did not differ from each other. These results indicate that our original grouping scheme for classifying animals as recognizing themselves produced clusters of animals that accurately predicted differences in the behavioral categories of CB, CF, and SE on the independent focal-animal scoring technique.

In order to determine if age alone was implicated in the differences in CB, CF, and SE behavior, we calculated the average frequency per 5-min sample for each of these behaviors across the 5 days of mirror exposure. For this analysis the animals were divided into six age categories. For the animals 6 years old and younger, we grouped the subjects according to their stable homecage groups, 2- to 3-year-olds ($n = 12$), 3- to 4-year-olds ($n = 8$), and 5- to 6-year-olds ($n = 13$). Because older animals were spread across many different groups, they were collapsed into two groups, adolescent or young adult (7-15 years old; $n = 18$) and adult (16-40 years old; $n = 35$) and the mother-reared offspring of the adults (10 months to 2.5 years old; $n = 6$). Separate one-way Age \times Behavior ANOVAs revealed no age group effects for SE or CF behavior. There was a significant overall effect for CB, $F(5, 86) = 2.69, p < .03$, and post hoc Tukey tests determined that the adult group showed significantly lower frequencies of CB behavior ($M = 0.082$) than did the adolescent and young adult group ($M = 0.280$), which did not differ from any other group. In general, these results reveal that age alone was a poor predictor of levels of CB, CF, and SE behavior.

Weak episodes. A quite different pattern emerged from an examination of the focal-mirror episodes of weak instances of CB, CF, and SE. An examination of the two right panels of Figure 2 shows far less discrimination between the three classes of animals in both the frequency and duration averages and in particular between those previously classified as SR+ and those classified as SR?. Also of importance is the sudden appearance of instances of CB, CF, and SE in the group classified as SR-. Taken together, these results show that many ambiguous instances of CB, CF, and SE

behavior occurred while animals were near the mirror. However, these behaviors occurred frequently in many other contexts as well.

Contingent body movements, contingent facial movements, and self-exploration correlations. In an effort to understand the relation between CB, CF, and SE behaviors, we conducted a correlation analysis with the focal-animal samples so that we might determine if high levels of SE tended to be uniquely correlated with high levels of CB or CF. The results revealed significant and similar correlations among all measures: For CB with CF, $r = .424$; for CB with SE, $r = .482$; and for CF with SE, $r = .513$ (Pearson product-moment correlations; all $p_s < .001$). However, it is evident that many subjects showed compelling CB and CF behavior without showing compelling SE behavior (see Figure 2, SR- subjects). We interpret this to mean that CF and CB behaviors are often produced by subjects that do not recognize themselves but that they are not necessarily independent evidence of the capacity (see General Discussion).

Developmental Course of Self-Recognition

In order to examine the developmental implications of our results, we grouped the animals into 5 categories: 1-5 years ($n = 48$), 6-7 years ($n = 10$), 8-15 years ($n = 12$), 16-24 years ($n = 20$), and 25-39 years ($n = 15$). Figure 3 displays the percentage of animals within each age group that were classified as SR+.

Two developmental effects are striking from the results depicted in Figure 3. First, there is a clear developmental transition from 6 to 8 years of age. Only 1 subject (Megan; age 3 years, 3 months) younger than 6 years of age showed compelling evidence of self-recognition. In addition, no subjects in this age group were classified as SR?. However, by 6-7 years of age, 25% of the subjects displayed clear and compelling SE behavior, and an additional 30% were classified as SR?. The second effect is the apparent decline in self-recognition in adult subjects. By 8-15 years of age, 75%

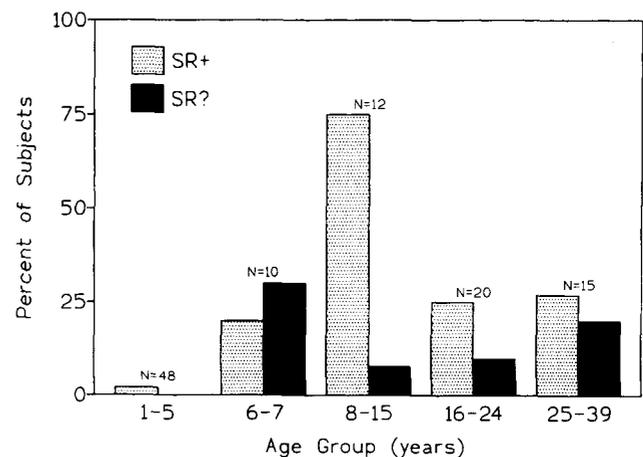


Figure 3. Percentage of subjects classified as showing compelling evidence of self-recognition as a function of age. (SR+ = positive self-recognition; SR? = ambiguous self-recognition.)

of the subjects were classified as SR+, with an additional 8% classified as SR?. By our classification scheme, only 26% of the full adults (16–39 years) showed evidence of self-recognition, and 14% were classified as SR? cases. This represents a significant difference between the adolescents or young adults (7–15 years) that recognized themselves and the full adults (16–39 years) that did, $\chi^2(1, N = 53) = 4.60$, $p < .05$.

Although all of the animals were held as part of the center's breeding colony and were healthy and showed no evidence of serious behavioral or clinical disorders, we found the decline in self-recognition between the adolescent and young adult group and the adult group counterintuitive. In order to determine if this finding was the result of previous biomedical investigations, we divided the adults into two classes, those that had participated in previous long-term projects (other than those involving short-term routine physical examinations) and those that had not. Of the 9 SR+ adults, 4 came from the group with no project history ($n = 13$) and 5 came from the group that had participated in some project ($n = 22$), a result that does not differ from chance, $\chi^2(1, N = 35) = 0.28$, *ns*. This result implicates age as the most likely factor to account for the significant difference in SR+ animals between the adolescent and young adult group and the adult group. However, an inspection of the data for the adult group revealed no correlation between increasing age and the frequency of SE behavior (Pearson's $r = .07$, $p = .704$).

Patterns and Rate of Onset of Self-Exploratory Behaviors

In contrast to previous reports, we found little or no evidence that mirror-naïve chimpanzees require 2–3 days to recognize themselves. Most SR+ subjects appeared to recognize themselves within the first 10–20 min of watching themselves, although in some cases the subjects did not begin to view themselves for a number of minutes (or hours) after the mirror was introduced. In order to quantify the rate of onset of self-recognition, the focal-mirror records were examined to determine the first recorded instance of a compelling bout of SE behavior for each of the 21 animals that were classified as SR+. Because there were only 2 SR+ animals younger than 7 years of age (ages 3 years, 3 months and 6 years, 1 month), the subjects were divided into two groups, young (3–15 years; $n = 12$) and mature (16–39 years; $n = 9$), for analysis. The average elapsed time before self-recognition for the young animals ($Mdn = 27$ min) was only about half as long as for the mature animals ($Mdn = 52$ min; see Table 2). However, the variance within the young group was high, and a Mann-Whitney U test revealed no significant difference between the groups.

Because observers were continuously present for the first 3 hr of mirror exposure, the first several instances of SE behavior were witnessed for nearly all SR+ subjects. Although the patterns of emergence differ in idiosyncratic ways among the SR+ subjects, at least one intriguing observation is worthy of further consideration. The vast majority of these

Table 2
Elapsed Time Between Mirror Presentation and First Compelling Bout of Self-Exploration

Subject	Age (in years, months)	Elapsed time (in min)
Young (3–15 years)		
Megan	3, 3	80
Paula	6, 3	50
Kirk	6, 11	1,630
Cindy	8, 5	10
Chakima	9, 7	14
Angela	10, 1	244
Misty	10, 1	28
Keetah	10, 6	26
Christy	10, 10	26
Gilda	11, 11	10
Lisa	12, 10	31
Schulman C	15, 4	5
<i>Mdn</i>		27
Mature (16–39 years)		
Wendy	16, 4	66
Margret	20, 0	52
Rose	21, 9	8
Linda	23, 4	30
Backus	24, 7	64
Zelda	26, 3	54
Brenda	29, 8	20
Dark Face	31, 8	10
Claudia	31, 3	89
<i>Mdn</i>		52

21 subjects initially showed some mild-to-dramatic social reactions, followed in some cases by CB or CF movements, before they showed compelling evidence of self-exploration. However, 2 of the younger subjects (ages 6 years, 1 month and 7 years, 3 months) displayed repeated bouts of what may be described as *recognition shock* on their initial confrontation with the mirror. Both subjects (which were tested at different times and were from different social groups) engaged in dramatic instances of slapping the top, sides, and front of their faces, pulling at the hair on their head and face, running back and forth in front of the mirror, and grimacing repeatedly while they intently monitored their reflections. During these bouts they also showed a full suite of other SE behaviors (for example, carefully scraping at their front teeth, picking at their genitals, eyes, and nose, etc). Thus, for at least these 2 subjects,¹ the onset of self-recognition appeared to be tightly linked to a profound emotional reaction.

Behavioral Reactions: Viewing Time, Proximity, and Social Behavior

To begin our analysis of the subjects' general behavioral reactions to the mirrors, we divided the animals into the six age groups described earlier. In addition to the behaviors

¹ In recent longitudinal follow-up studies, we have observed a milder form of this phenomenon in a 7-year-old male subject as well.

already analyzed (CB, CF, and SE), we selected three behavioral categories for detailed examination, VE, proximity, and social reactions (see later description). For each subject we again averaged all 5-min focal-animal samples across the 5 days of mirror exposure in order to obtain the mean number of intervals (out of 20) in which the subject displayed each of the three behavioral categories.

Visual exploration. The subjects showed high initial levels of interest in the mirror but quickly habituated, an effect previously reported by Gallup and colleagues (Gallup, 1970; Gallup et al., 1971; Suarez & Gallup, 1981). Figure 4 represents the mean modified frequencies with which the subjects within the six age groups viewed themselves, averaged across their 5 days of exposure. Separate one-way ANOVAs revealed an overall group effect for VE, $F(5, 86) = 3.13$, $p < .02$, and post hoc Tukey tests indicated that the adult group showed significantly lower frequencies of viewing themselves ($M = 2.162$) than both the 2- to 3-year-olds ($M = 3.379$) and the 3- to 4-year-olds ($M = 3.588$). However, the post hoc comparisons also indicated that the two groups that produced most of the SR+ animals (adolescent and young adult and adult) did not differ from each other in the average number of intervals that the animals spent viewing themselves ($M_s = 2.173$ and 2.166 respectively). Overall, these results confirm that VE did not uniformly increase or decrease with age and thus was not the cause of the age effect in self-recognition that we found. Indeed, the youngest 20 animals (only 1 of which recognized itself) showed the highest levels of VE behavior.

Although the values reported in Figure 4 may appear low, they are not when compared with previous results. For example, Gallup (1970) reported that his chimpanzees viewed the mirror on average only 296 s out of 30 min across the first 5 days of exposure; likewise, Gallup et al. (1971) reported that their chimpanzees spent only about 100 s out of 30 min

viewing the mirror across the first 5 days of exposure. If our mean modified-frequency values are converted into duration estimates (for which 1 interval equals on average about 7 s),² these results are well within the ranges reported by Gallup (1970; Gallup et al., 1971): Such estimates range from about 147 s (per 30 min) in our younger animals to about 100 s (per 30 min) in the adolescents and adults.

A separate analysis was performed to determine if SR+ animals differed significantly from the SR? or SR- animals in the average number of intervals that the animals spent viewing themselves. This is important, because it is possible that those animals that showed positive evidence for self-recognition did so only because they tended to explore the mirror for longer periods of time outside the context of producing CB, CF, or SE behaviors. However, the results of a one-way ANOVA revealed no significant interaction between an animal's self-recognition status and the amount of time it spent engaged in VE of the mirror. This suggests that the between-subjects self-recognition effects reported earlier are not the result of individual differences in the overall amount of time spent in VE behavior. Indeed, the group with the highest percentage of SR+ subjects showed only moderate mean viewing times (see Figure 4).

Proximity. For each subject an overall mean score for proximity was obtained by averaging the scores for each focal-animal sample across the 5 days of exposure. An ANOVA revealed an overall effect in the number of intervals the different age groups spent in proximity to the mirror, that is, near the mirror but not looking at it, $F(5, 86) = 3.29$, $p < .009$. Post hoc tests indicated that the group of 2.5- to 3.5-year-olds were in proximity to the mirror significantly less often than either the adult group ($p < .02$) or the group that comprised the offspring of the adults ($p < .02$). (This latter effect is not surprising because 5 of the 6 animals in the offspring group were still quite dependent on their mothers and hence remained in close proximity to them most of the time.) Finally, a separate analysis that grouped the subjects by recognition status revealed an overall effect, $F(2, 89) = 8.96$, $p < .001$, which post hoc tests identified as the result of the SR? animals' showing significantly higher frequencies of proximity ($M = 7.169$) than either the SR+ or SR- animals ($M_s = 3.282$ and 3.294 , respectively). Thus, the animals that were classified as SR? spent significantly more

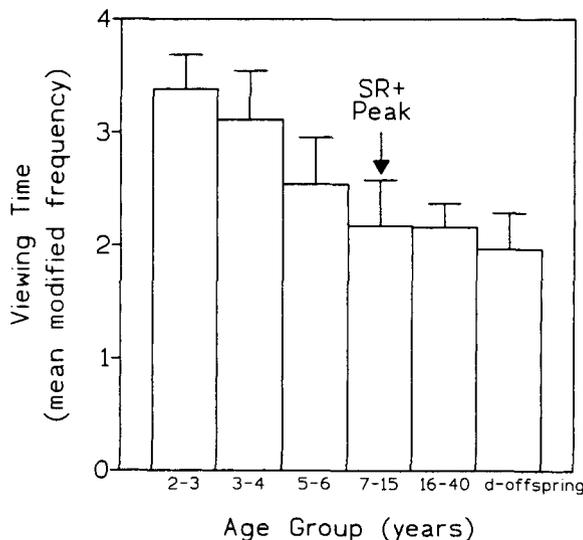


Figure 4. Visual exploration of the mirror as a function of age. (SR+ = positive self-recognition; d-offspring = dependent offspring of adults.)

² Mean modified-frequency measures are designed as a compromise between duration and frequency estimates. Our conversion is based on the fact that each interval in our 5-min samples was 15 s long and thus each mean modified-frequency tally might range from 1 to 15 s in duration. On average this was about 7 s. Although we recognize this is not necessarily a perfect estimate, it is able to provide a rough comparison to previous methods. Our rationale for using a mean modified-frequency method was that methodological advances since Gallup's (1970) original study dictated the need for a more detailed behavioral ethogram and constant surveillance for the first 3 hr of mirror exposure. Given that we were observing many subjects in each session, this made it inadvisable to collect pure duration estimates on viewing time. Although this provides some small difficulties in comparing mean viewing times to previous investigations, it escapes other methodological troubles associated with pure duration estimates.

time near the mirror, even though they were not actually looking into it.

Social reactions. On the basis of previous research, we predicted that the SR+ group would show less social reactions to their mirror images than the SR- group. For purposes of data analysis, the behavioral categories of social behavior, fear, looking behind, and aggression were collapsed to form an overall category of social reactions. A one-way ANOVA of Age Group \times Social Reactions revealed an overall effect, $F(5, 86) = 2.44, p < .05$, and post hoc tests revealed that the group of 3- to 4-year-olds showed significantly higher frequencies of social reactions than did the adults ($p < .02$) and approached a significantly higher frequency than did the 2- to 3-year-olds ($p < .09$). Further analyses revealed no significant differences in social reactions directed toward the mirror among the SR+, SR-, and SR? groups either on Day 1 or on Days 1-5, although the SR- group did show a two-fold higher frequency than the SR+ group in both cases (see Figure 5).

Rearing history and gender effects. Although we had initially hoped to analyze the data for both gender and rearing history effects, the late developmental transition made comparisons within these categories statistically meaningless. For instance, our samples of captive mother-reared versus nursery-peer-reared subjects were limited by the fact that all of the captive mother-reared chimpanzees were under 3 years old. This is well below the age at which we detected the developmental transition to an onset of self-recognition. Thus, although there may be a difference between mother- and nursery-peer-reared subjects over 3 years of age, we have no data that would bear on the issue. On the other hand, many of the adolescents and young adults had been reared by their mother for at least 1 year and in some cases up to almost 5 years before being transferred into peer or breeding groups. Sixteen subjects between the ages of 6 years, 3 months and 16 years, 4 months had rearing histories that could be clearly determined. We divided these subjects into mother-reared and nursery-peer-reared and compared the groups in terms of their self-recognition status in both the

spontaneous observations and the mark tests. There was no difference by rearing history: Five of 8 subjects in both classes were classified as SR+. The sample is too small, however, to provide a statistical analysis. In terms of the captive- versus wild-born issue, the sample is clearly biased because all but 2 of our adult subjects (16 years and older) whose histories are known were wild-born. Of these 29 subjects, 6 were classified as SR+, and 23 were classified as SR-. One of the captive-born adults was classified as SR+, and the other as SR-.

We were similarly handicapped in our investigation of gender effects. Because of the breeding configuration of the adults (1 male for every 5-6 females), our sample of adult males is extremely small ($n = 8$) in comparison with the female sample ($n = 27$). Fisher's exact test revealed no significant difference between the numbers of adult males (1 of 8) and of adult females (8 of 27) that were classified as SR+ ($p = .648, ns$). No such test was possible for the adolescent and young adult group because the younger half of the 6- to 15-year-olds were almost all males, and the oldest half were all females, and this distribution confounded age with gender.

Experiment 2

Some surprising findings emerged from Experiment 1, including both the late onset of self-recognition and our difficulty in demonstrating it in adults. In Experiment 2, we sought to address the latter issue. Although the amounts of time that the adults and the adolescents and young adults spent viewing themselves did not differ in Experiment 1, we decided to systematically explore the effects of extended exposure on SR- adults by giving them prolonged access to a mirror to determine if this would facilitate the emergence of SE behaviors.

Method

Subjects

The subjects were 5 wild-born adults (1 male and 4 females) that had participated in Experiment 1. They were part of a stable, long-term breeding group and were housed in one of the indoor-outdoor enclosures described previously. At the time of the experiment, there were also 2 infants in the group (1 male at age 28 months and 1 female at age 10 months).

Procedure

Immediately after this group's participation in Experiment 1 (which involved 5 continuous days of mirror exposure), it was given continuous access to a mirror for 15 additional consecutive days. The mirror was placed in the indoor portion of the enclosure (just out of the subjects' reach) and remained in place except during morning cleaning. Because interest in the mirror had declined to low levels, a focal-mirror sampling procedure was used. The mirror was observed for 1 hr every other afternoon, and all instances by any subject of CB, CF, and SE behaviors (including weak instances) were recorded on data sheets. The duration of each bout was estimated with a stopwatch.

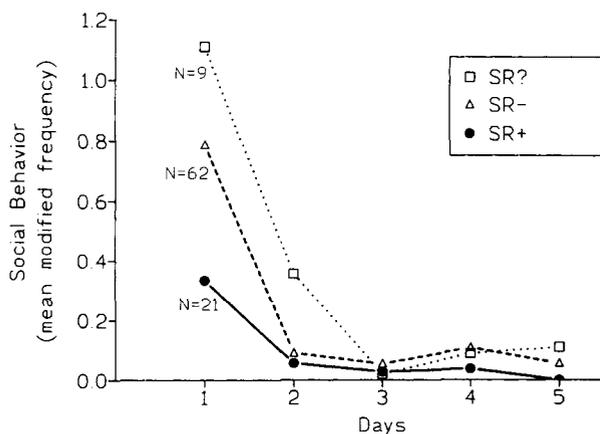


Figure 5. Frequency of social reactions to mirrors by chimpanzees diagnosed as positive, negative, and ambiguous for self-recognition (SR+, SR-, and SR?, respectively).

Results

The results of the eight hour-long focal-mirror samples collected on Days 6, 8, 10, 12, 14, 16, 18, and 20 are summarized in Table 3. None of the adults or offspring showed any instances that qualified as evidence of CB, CF, or SE behavior. However, with the exception of the youngest animal, all of the subjects showed occasional weak instances of at least one of the three behavioral categories. Informal notes recorded on the focal-mirror data sheets revealed that most of the subjects viewed themselves at rates comparable with those of Day 5 of Experiment 1 (e.g., about 3 min per hour). In addition, there were several instances of social behaviors. In general, the 15 days of extended exposure produced no changes in the subjects' reactions toward their mirror images.

Experiment 3

The first two experiments confirmed the absence of spontaneous patterns of SE behavior in many subjects. However, we wondered if additional mirror exposure might have an effect on younger subjects, even though it had no discernible effect on older subjects. In order to determine if our negative results with animals younger than 6 years old were due to the subjects' insufficient experience with mirrors, we saturated 6 of the younger subjects with additional exposure (and direct access) to mirrors and then observed them in conditions more similar to Gallup's (1970) experimental setting and administered mark tests to all subjects. In addition, several of the subjects were tested again 1 year later.

Method

Subjects

In Year 1, the subjects were the 6 oldest chimpanzees in the original 2- to 3-year-old group described in Experiment 1. Five had been classified as SR-, and 1 (Megan; age 3 years, 3 months) had been classified as SR+. Megan was unusual in that she was the only chimpanzee out of 50 subjects between the ages of 10 months and 6 years who showed compelling SE behavior in Experiment 1. At the time of this investigation, the 6 subjects were between the ages

of 3 years, 4 months and 4 years, 1 month and had been participating in several unrelated cognitive studies for some months. Twelve months later (Year 2), 5 of the original chimpanzees from Year 1 and 2 additional subjects, with whom the others had been reared for the first 3 years of life, were observed again. The 2 additional subjects had participated in Experiment 1 as well as a number of cognitive tests before their test in Year 2. All subjects had had additional exposure to mirrors in the 12 months between the two time points.

Procedure

Year 1. Immediately after the subjects' participation in Experiment 1, they were given an extended amount of additional exposure (6–8 hr a day for 2 weeks) to the mirror in the outdoor portion of their homecage setting (as in Experiment 1). In addition, the subjects were given about eight 2-hr exposure sessions with physical access to a smaller mirror (105 × 105 cm) in a small testing room in the context of free play (see Figure 6). For this additional exposure the subjects were either in pairs or groups of 4–6 animals.

The formal testing for each pair of subjects was staggered so that it began between 4 and 6 months after the subjects had participated in Experiment 1. At approximately 0830 hr each morning for 4 consecutive days, a pair of subjects was carried to a quiet testing room and placed into separate testing cages in which they had become familiar in the context of other studies. A large mirror was positioned in front of the two cages so that the subjects had full visual access to both themselves and their partners in the cage adjacent to them. The subjects could also see each other directly but were just out of range of direct physical contact.

Data were collected by two observers with the focal-animal scoring system described in Experiment 1. The behavior of each animal was sampled for 1 hr each day (six 5-min samples each morning and afternoon) under the methods described in Experiment 1. The ethogram was modified slightly by dividing the category VE into two subcategories, active VE and passive VE. The pair of animals was returned to its social group in the afternoon immediately after the second session of data collection, after 5–6 hr of continuous mirror exposure.

On the morning of Day 5, the subjects were anesthetized with ketamine (10 mg/kg) and marked on their right eyebrow ridge and upper left ear with a bright-red indelible dye that left no discernible tactile or olfactory cues. In addition, 4 of the subjects were marked on their left wrist. They were returned to their single cages and allowed to recover. During the initial period of recovery, the subjects were observed to determine the number of times they inspected the marks applied to their wrists. Once the subjects were fully re-

Table 3
Frequencies of Behaviors Shown by Negative Self-Recognition Adult Subjects During Extended Exposure

Behavior	Subject							Total
	Paco	Hope	Gwen	Sandy	Jessica	Blue ^a	Toni ^a	
Contingent body movements								
Compelling	0	0	0	0	0	0	0	0
Weak	1	0	0	1	0	2	0	4
Contingent facial movements								
Compelling	0	0	0	0	0	0	0	0
Weak	0	1	1	0	2	1	0	5
Self-exploratory behaviors								
Compelling	0	0	0	0	0	0	0	0
Weak	0	1	1	2	0	0	0	4

^a 28 months. ^b 10 months.

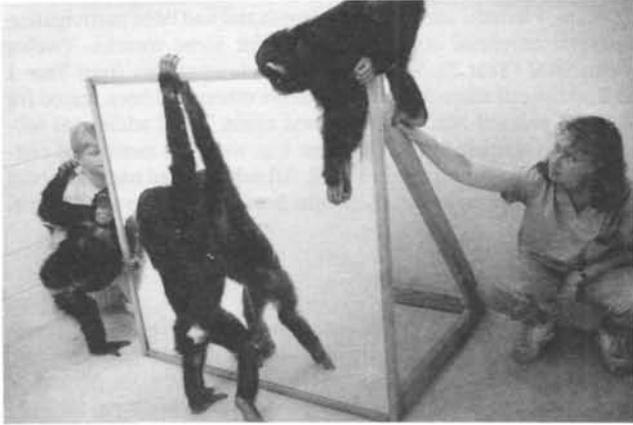


Figure 6. Experimental setting for extended exposure to mirrors in Experiment 3.

covered (about 5 hr later), they were fed and given juice. They were then each observed by a single observer for a 30-min control session without the mirror. The number of times the marked areas were contacted by any portion of the subjects' upper wrist or hand were recorded. Next, the mirror was separately introduced to each subject for a 30-min test session, and a second observer recorded the number of times the subject touched the marked areas on the eyebrow and ear. The mark touches were classified as either mirror-monitored or non-mirror-monitored in order that we might distinguish between inadvertent contacts with the marks and attempts that appeared to be the result of the subject's seeing the marks on its face and ear in the mirror.

Year 2. Twelve months after the conclusion of their first mark tests, a mirror was positioned on the outside of the subjects' indoor-outdoor cage. Data were collected by one observer with the same focal-animal and focal-mirror scoring system, methods, and ethogram as described earlier. The subjects were observed for 2 consecutive days in a 2½-hr session each day. Each subject was observed for five 5-min periods. The data were summarized and analyzed as in Experiment 1. Before observing the animals the two observers practiced the scoring system on videotapes for several days. A formal reliability check was then conducted as in Experi-

ment 1. The observers agreed 88% of the time in 245 intervals of observation. Agreement for individual categories ranged from .78 to .98.

Approximately 3 weeks after mirror exposure, the 7 animals were scheduled on 7 consecutive days for yearly physical examinations. The day before each scheduled mark test, the 7 subjects were exposed to a mirror positioned outside of the cage for 3 hr in the afternoon. The following day a subject was anesthetized and marked as before (right eyebrow ridge and upper left ear). No animals were marked on their wrist. The subject was then brought to a large testing room with a Plexiglas wall and allowed to recover. All of the subjects were familiar with the room in the context of having participated in several cognitive tasks for the preceding 5 months. Once the subject had fully recovered, a covered mirror was positioned about 30 cm in front of the Plexiglas wall and the subject was observed for a 30-min control session and all touches to the eyebrow or ear were noted as described for Year 1. The cover was then removed and the the subject was scored for any mirror or non-mirror-monitored touches for a 30-min test session.

Results

Social Reactions

The subjects' social reactions to the mirror in Years 1 and 2 are shown in Figure 7. In general, they tended to be low, and a repeated measures ANOVA revealed that the frequency of social responses did not change across the 4 days of the experiment in Year 1, $F(3, 15) = 1.11, p = .37$. This is not surprising given the subjects' extensive prior exposure to mirrors. The occasional social gestures that did occur primarily took the form of sexual presentations and begging gestures. In Year 2, social behaviors tended to be low as well. Of interest is that in Year 1 the lone SR+ subject showed among the highest amounts of social behavior, despite her immediate evidence of self-recognition when the mirror was unveiled (see later description). Also of importance is that although the social behaviors of most of the SR- subjects declined across the first 3 days in Year 1 (as in Gallup's

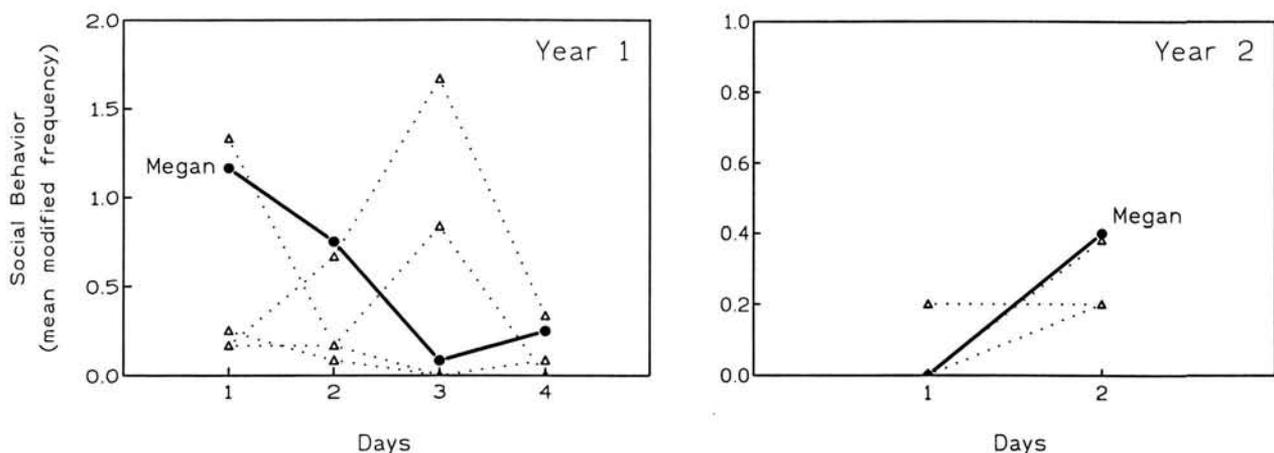


Figure 7. Social reactions in Years 1 and 2 in longitudinal study of young SR- (dotted lines) and SR+ (heavy line) chimpanzees.

[1970; Gallup et al., 1971] previous investigations), the subjects showed no compelling SE behaviors (see later description).

Visual Exploration

In Year 1, the rate of active VE was considerably lower than passive VE, but a repeated measures ANOVA revealed that both forms of VE showed a significant change across days: Active VE, $F(3, 15) = 4.31, p < .05$, and passive VE, $F(3, 15) = 6.84, p < .01$. Tukey tests revealed that for active VE the difference was because the subjects showed significantly lower scores on Day 4 than on Day 1 and that for passive VE the difference was due to significantly lower scores on Days 3 and 4 than on Day 1. Active and passive VE was not distinguished in Year 2, and the results of a two-tailed Student's t test revealed no significant difference on Day 1 and 2 rates of overall VE behavior, $t(6) = 1.66, ns$.

Contingent Body Movements, Contingent Facial Movements, and Self-Exploratory Behavior

Figure 8 provides the time course of the frequency of CB, CF, and SE across the days of the observations for each subject in Years 1 and 2 (for the SR- subjects this includes the weak instances of these behaviors). For Year 1, the responses of the 1 precocious subject (Megan; age 3 years, 9 months) who showed clear evidence of self-recognition in Experiment 1 are labeled and reveal intriguing contrasts to the other 5 subjects. Megan engaged in considerably more CB and CF movements (especially on Day 1 when the mirror was re-introduced to the subject) than did the other subjects. However, all of the other subjects showed at least an occasional instance of these behaviors as well. Indeed, 2 of the subjects displayed very clear instances of alternating their gaze between one of their limbs, as they moved it slowly back

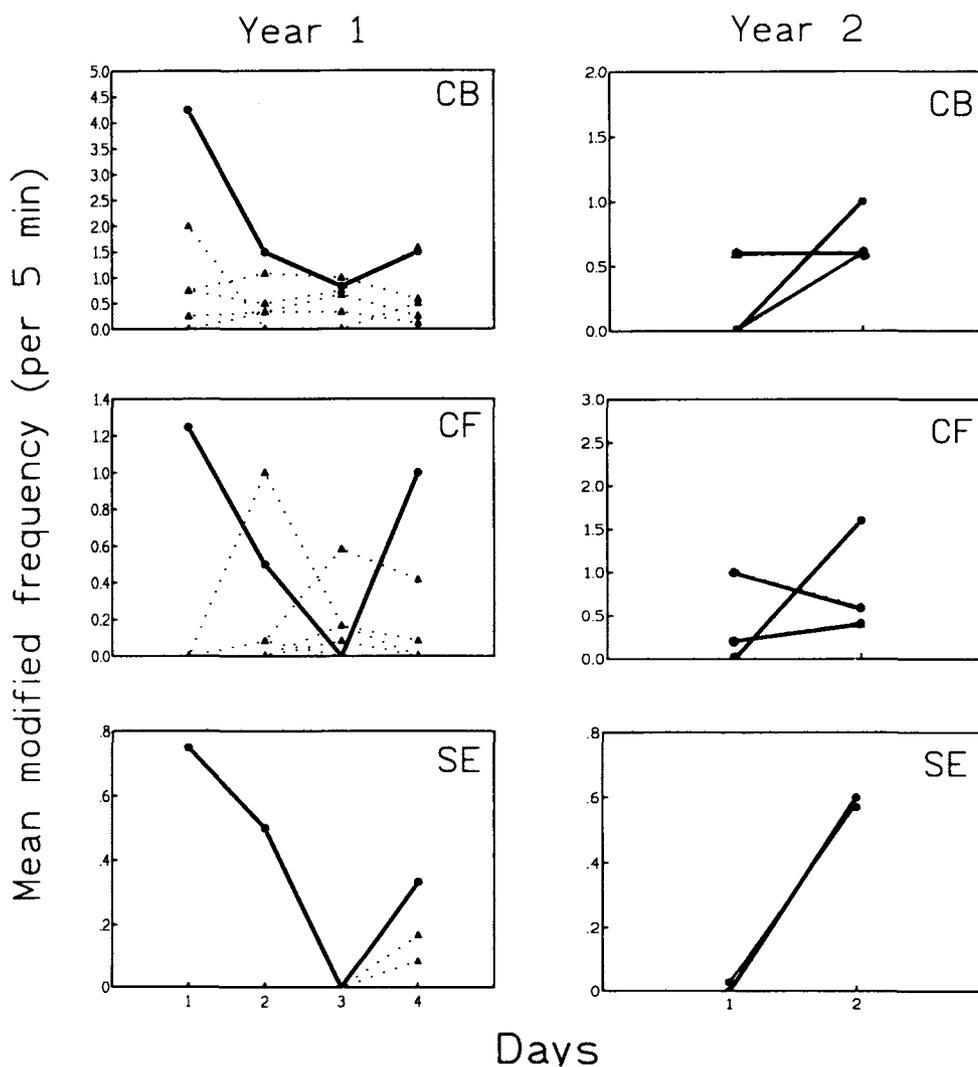


Figure 8. Mirror-related behavior in Years 1 and 2 in longitudinal study of young SR- (dotted lines) and SR+ (heavy line) chimpanzees.

and forth, and the mirror. Megan was most notably different with respect to instances of SE behavior. On Day 1, she showed numerous compelling instances of using the mirror to aid her in exploring parts of her body (nose, teeth, and genitals). However, these reactions waned quickly. Other subjects showed no compelling evidence of such behaviors, although they did show several weak instances.

In Year 2, Megan again showed several clear examples of SE behavior, as did an additional subject tested in Year 1. Also, 1 of the 2 subjects who were not part of the Year 1 investigation also showed several compelling SE behaviors across the 2 days and met the Experiment 1 criteria for SR+.

Mark Tests

The results of the mark tests for both years are presented in Table 4. In Year 1, all 4 of the subjects that were marked on the wrist touched and licked these marks repeatedly during the the initial 45 min to 1 hr of recovery. However, only Megan showed clear evidence of discovering the facial marks after she was exposed to the mirror. The other subjects, although they occasionally brushed the marks while scratching or moving around the cage, showed no evidence of mirror-monitored, mark-directed responding. Even more significant, all but 1 of the other subjects displayed a roughly equal number of brushes or scratches to the marked areas in the control and test period. This demonstrates that the mark contacts for these subjects in the test period were not contingent on the subjects' seeing themselves in the mirror. In contrast, Megan's mark-directed responses took the form of careful touches, followed by inspections of her fingers (see Gallup, 1970). This differed strongly from the inadvertent

contact made by the other subjects, which never inspected their fingers after the occasional brush or scratch while swinging in the cages. One of the subjects showed considerably more contacts during the test session than during the control session. However, none of these contacts were mirror-monitored and appeared to occur as a consequence of a general increase in agitation after the introduction of the mirror.

In Year 2, 2 of the 3 subjects diagnosed as SR+ (based on the behavioral observations prior to the mark test) passed the test (see Table 4). One of these 2 was not tested in Year 1. The other SR+ subject (see Table 4) threw a temper tantrum in the final 5 min of the control session, and in the context of rolling around inside the testing room and rubbing its head, it made numerous contacts with the marks. The data for this subject are therefore difficult to interpret. In contrast, the 4 SR- subjects (which were also diagnosed as SR- in Year 1) showed no evidence of mirror-monitored mark-directed responding, except possibly for 1 subject, which appeared to show two instances of mark-directed contacts. In additional follow-up tests 3 months later, this subject showed numerous compelling SE behaviors, whereas the other SR- subjects did not.

Experiment 4

The results of the first three experiments indicated that for certain chimpanzees, spontaneous patterns of self-exploration were not going to emerge, even after extended interaction with mirrors. The third experiment suggested a coherence between the results of the mark tests and our systematic observations of spontaneous behavior. However, in

Table 4
Touches to Mark After Extended Exposure (Experiment 3)

Behavioral diagnosis and subject	Age (in years, months)	Wrist control	Control session	Test session	
				Mirror- monitored	Non-mirror- monitored
Year 1					
Positive self-recognition					
Megan	3, 9	34	0	23	1
Negative self-recognition					
Kara	4, 1	21	4	0	5
Justin	3, 11	20	0	0	10
Apollo	3, 10	—	8	0	9
Mindy	3, 8	—	7	1	7
Brandy	3, 4	10	1	0	1
Year 2					
Positive self-recognition					
Megan	4, 9		1	14	0
Mindy	4, 9		6	5	6
Danielle	4, 5		1	9	1
Negative self-recognition					
Kara	5, 1		1	0	0
Apollo	4, 10		0	0	0
Jadine	4, 5		0	2	0
Brandy	4, 4		0	1	1

Note. The animals were not marked on the wrist in Year 2.

order to explore the statistical relation between spontaneous behavioral responses recorded in Experiments 2 and 3 and other measures of self-recognition, we applied the mark test procedure to three groups of animals: those that had shown clear evidence of SE behavior (SR+); those that had shown some instances of CB and CF movements but only a few instances of SE behavior (SR?); and those that had shown high levels of interest in the mirrors but no SE, CB, or CF behaviors (SR-).

Method

Subjects

Thirty subjects that had participated in previous experiments were selected. Seventeen of the 21 animals that were classified as SR+ in Experiment 1 served as the test subjects; 7 that were classified as SR- were selected as control subjects; and 6 of the SR? subjects were also tested. The subjects were tested between 1 and 4 months ($Mdn = 2.5$ months) after their participation in Experiment 1 and had received no mirror exposure in the interim.

Procedure

During a period in which the animals were scheduled for semi-annual physical examinations, 2 subjects from each group were separated from their group, anesthetized with ketamine (10 mg/kg), and while unconscious, they were marked on the right upper eyebrow ridge and the upper half of the left ear with a bright-red or pink acrylic paint. We selected acrylic paint as the best option after experimenting with several other dyes and paints, which were too slow in drying, did not contrast well against the dark faces of most of the subjects, flaked off, or left tactile cues after drying. In addition to the marks on the ear and eyebrow, the subjects were also marked on their left wrist. After the paint had fully dried, the animals were placed in separate clinical metabolism cages in a quiet room near their colony area and allowed to recover. The large Plexiglas mirror used in Experiment 1 was covered with an opaque tarp and placed about 25 cm away from the front of the two cages. The subjects had visual and auditory access to each other but were just outside each other's physical reach. Most of the subjects were familiar with the setting. During the 1st hr of recovery, 11 of the subjects were observed by one or two observers, and the number of times they showed a deliberate effort to touch or lick at the marked area on their wrist was recorded. As soon as 10 clear instances had been recorded, these observations ceased.

Approximately 4–7 hr later, once the subjects were fully awake and had been given food and juice, each subject was observed simultaneously by a different observer for a control period of 30 min. During this baseline period the total number of times that the subjects touched the marked areas with any part of their wrists or hands was recorded. Unlike previous investigations, however, we divided the topography of the contacts with the marked areas into two categories, touches (contacts with the marks with some portion of the subject's finger or fingers) and rubs (any other contacts with the marks by the upper wrist and hand; e.g., a brush of the mark with the back of the hand). The purpose of this distinction was to assist us in distinguishing between inadvertent and intentional contacts. Immediately after the 30-min control period, the mirror was uncovered. Again, the subject was observed for 30 min, and the number of times the marked areas were touched or rubbed was recorded. However, the contacts were further divided into two additional categories,

mirror monitored and non-mirror-monitored. To qualify as mirror monitored, a touch to the marked areas must have occurred while the subject was looking directly in front of itself into the mirror. Non-mirror-monitored touches occurred when a subject touched a mark but either was not facing the mirror or was facing the mirror but not looking into it.

Results

The results of the mark tests for each of the three groups are presented in Table 5. The results for the SR+ group revealed two very distinct response profiles among the subjects. As can be seen from the data on touches only in Table 5, about half of the subjects responded by showing clear instances of mirror-monitored mark-directed touches in the test session. These touches were often followed by visual or olfactory inspection of the fingers that had made contact with the marked areas. In addition, these subjects showed the same number of incidental touches (typically in the context of scratching their heads) in both the control session and in the non-mirror-monitored portion of the test session. In striking contrast, the other half of the SR+ subjects showed virtually no instances of mirror-monitored mark-directed touches, yet 3 of these subjects (at ages 10, 16, and 20 years) showed numerous compelling instances of CB, CF, and SE behavior, including manual manipulations of the anal-genital region, teeth, eyes, and lips. An additional SR+ subject that did not touch the facial marks showed a number of clear instances of both CF and CB behavior. These results suggest that either a dissociation can exist between spontaneous patterns of SE behavior and success on a mark test or that some animals habituated to the marks on their own wrists and to the ones on their companions and then generalized to the marks on their faces when the mirror was exposed.

The results for the SR- animals are also displayed in Table 5. Six of the 7 subjects showed no evidence of self-recognition on the mark test, nor did they display any spontaneous patterns of CB, CF, or SE behavior during the test session. However, 1 subject (age 7 years, 8 months) did display an unambiguous pattern of using the mirror to locate and then inspect the facial marks during the test session. The results for the SR? group revealed a similar pattern. Five of the 6 subjects tested negative and showed no evidence of spontaneous CB, CF, or SE behavior. However, 1 subject (age 31 years) responded to the marks in an unambiguous fashion during the test session. The reactions of the subjects to the wrist control marks showed that all subjects were interested in the directly visible marks on their wrists. Of the 11 subjects we observed during recovery, all but 2 showed 10 touches directed to the wrist mark during the early minutes of recovery from the anesthesia. Indeed, most of the subjects were responding to the marks even before they had regained complete consciousness. The other 2 subjects touched the marked areas 6 times before the observations ceased (after 45 min of observation).

In order to provide a statistical analysis of the results of the mark tests, we combined the results of the tests in Experiment 3 with the results of this experiment. This pro-

Table 5
Mark Test Results of Positive Self-Recognition, Negative Self-Recognition, and Ambiguous Self-Recognition Chimpanzees

Result of test and subject	Test session						Time elapsed since last mirror exposure (months, days)
	Control session		Mirror-monitored		Non-mirror-monitored		
	T+R	T	T+R	T	T+R	T	
Positive self-recognition							
Passed							
Cindy	3	0	24	15	2	2	2, 10
Brenda	13	8	48	40	10	5	1, 3
Rose	3	2	5	5	2	1	2, 4
Dark Face	5	1	5	5	1	0	1, 8
Linda	3	0	8	7	0	0	2, 18
Gilda	0	0	6	6	0	0	2, 18
Keetah	7	3	8	6	8	6	2, 18
Paula	1	0	5	5	1	0	4, 13
<i>M</i>		1.8		11.1		1.8	
Not passed							
Zelda	5	2	0	0	3	1	2, 4
Margret	2	1	0	0	2	1	2, 25
Kirk	27	11	1	1	5	5	2, 9
Baccus	2	1	0	0	1	0	1, 5
Wendy	4	2	0	0	3	2	1, 5
Christy	10	10	0	0	0	0	2, 17
Misty	8	7	0	0	4	3	2, 17
Angela	0	0	0	0	0	0	2, 18
Chakima	0	0	0	0	0	0	2, 15
<i>M</i>		3.8		0.1		1.4	
Negative self-recognition							
Passed							
Jeff	1	1	15	15	1	0	2, 10
Not passed							
Julius	2	1	0	0	4	1	1, 2
Bill	3	0	0	0	4	4	2, 19
Suzie	4	4	0	0	3	0	3, 5
Mildred	11	5	0	0	5	2	2, 8
Shelly	5	2	0	0	2	0	3, 2
Val	3	0	0	0	1	0	4, 8
<i>M</i>		2.0		0.0		1.2	
Ambiguous self-recognition							
Passed							
Light Face	1	1	8	8	0	0	1, 2
Not passed							
Todd	6	0	6	3	12	5	2, 27
Eric	14	13	0	0	5	2	2, 7
Liz	0	0	0	0	1	0	2, 19
Ellis	3	0	0	0	4	4	2, 29
Mikey	5	0	0	0	11	5	2, 8
<i>M</i>		2.6		0.6		3.2	

Note. T = touch; R = rub.

vided us with a total of 36 mark tests. The distribution of chimpanzees that successfully passed the tests as a function of their self-recognition classification is provided in Table 6. The results of chi-square analysis clearly revealed that the three groups differed strongly in the number of subjects that successfully passed the mark test ($p < .05$). This difference is the result of the higher proportion of subjects in the SR+ group that passed the test. A further examination of the length of the delay from mirror expo-

sure in Experiment 1 to the date of a subject's mark test showed no relation to successful performance on the test. If the SR+ subjects were divided roughly in half on the basis of the length of delay to the mark test, 4 of 8 subjects that received the mark test immediately or within 2 months, 9 days after Experiment 1 passed the test; conversely, 5 of 9 that were tested between 2 months, 10 days and 4 months, 13 days passed the test. Indeed, 4 of the 5 SR+ animals that were tested after the longest delay from

Table 6
Performance on Mark Tests As a Function
of Self-Recognition Status

Self-recognition	Successful	Unsuccessful
Positive	9	9
Negative	1	11
Ambiguous	1	5

Note. The distribution of successful mark test performances between positive and negative self-recognition animals differed strongly from chance (Fisher's exact test, $p = .02$). The distributions of successful performances between positive and ambiguous self-recognition animals and between negative and ambiguous self-recognition animals did not differ from chance (Fisher's exact tests, $ps = .34$ and 1.00 , respectively).

the previous mirror exposure passed the mark test. Thus, in general, the results of the mark tests vindicated our behavioral classification scheme, but simultaneously it advised caution against an assumption that the results of mark tests are the sine qua non of self-recognition in mirrors.

Experiment 5

The results of Experiments 3 and 4 provide clear evidence that our failure to find evidence of self-recognition in both young and adult SR- chimpanzees was not the result of insufficient mirror exposure. In addition, these experiments also indicate that SR- subjects do not typically pass a mark test (1 of 12 in our sample). However, in both experiments there were cases in which subjects appeared to show a high level of agitation after the reintroduction of the mirror in the test session of the mark tests, which resulted in a rather high incidence of non-mirror-monitored mark contacts by the SR- subjects. This led us to wonder if the presence of the mirror alone could produce increased touches to areas of the face normally targeted for the mark test. In order to empirically address the issue, we conducted this experiment.

Method

Subjects

The subjects were the same 6 chimpanzees tested in Experiment 3. In the first phase of the experiment, 1 of the subjects was removed from the data-collection procedure because of continuing levels of aggression and emotional distress during testing; this subject was included in the second phase because the behaviors disappeared. The experiment was conducted 4 months after the conclusion of Experiment 3.

Procedure

Phase 1. A pair of chimpanzees was ushered into a quiet room, and each subject was placed in a separate testing cage. No marks were applied to their faces. At the time of the study, the subjects were thoroughly familiar with the procedure and were typically tested on a variety of tasks in the room 5 days a week. The smaller mirror (105 × 105 cm) was stationed in front of the testing cages

but was covered for the baseline condition. Two observers sat facing the cages and recorded the number of times each subject touched their right upper eyebrow ridge and left ear within a 30-min baseline period with the mirror covered. In the test condition, the mirror was uncovered (open), and again the observers recorded the number of times the subjects touched the two areas with some part of their upper wrist or hand. Each subject was tested once.

Phase 2. Although Phase 1 provided some initial support for our hypothesis, we decided to probe the relation further by eliminating the confound between the order in which the subjects were exposed to the two conditions and the number of touches to the target areas. Thus, in the second phase of the experiment, the chimpanzees were tested on 4 consecutive days in an alternating design for which half of the subjects received a 30-min open-mirror condition on Days 1 and 3 and a covered-mirror condition for 30 min on Days 2 and 4, and the other half received the opposite treatment. Data on the number of touches and rubs to the eyebrow and ear were collected with the procedures described for Phase 1.

Results

The results of Phase 1 for the 5 subjects are presented in Figure 9. Three of the subjects showed higher levels of touching the upper eyebrow ridge and ear in the open-mirror condition than in the covered-mirror condition, and the others showed the same levels in both conditions. Two of the subjects showed especially pronounced increases associated with general increases in agitation when the mirror was uncovered. The results of Phase 2 are depicted in Table 7. Although the average number of touches, rubs, and scratches to the unmarked target areas were higher in the covered-mirror condition ($M = 35$) than in the open-mirror condition ($M = 24$), no consistent pattern appeared to emerge among the subjects. Indeed, only 1 of the subjects showed the same effect across both days of the two conditions; for this subject, the open mirror appeared to have a calming effect over the mirror-covered condition.

Although the results of Phase 2 did not confirm our predictions, it is important to note that in general the number of contacts with the targeted areas were quite high, despite the fact that no marks were present. In fact, the number of contacts to the target areas was on average over 10 times higher during the mirror-covered condition than it was during the comparable control period during Experiment 4 for this same

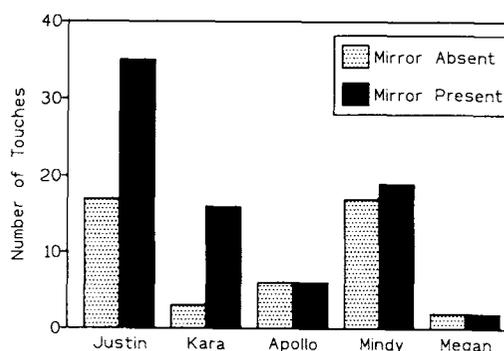


Figure 9. Number of touches to unmarked brow and ear in Phase 1 of Experiment 5.

Table 7
Number of Touches to Targeted But Unmarked Facial Areas (Experiment 5, Phase 2)

Subject	No. of contacts	
	Open mirror	Covered mirror
Kara	28	27
Justin	66	101
Mindy	14	50
Apollo	19	19
Brandy	8	8
Megan	8	2

group of subjects. Likewise, the number of contacts during the open-mirror condition in this experiment was over 4 times as high as the number of non-mirror-monitored touches in the test phase of Experiment 3. However, one must keep in mind that there was an important procedural difference between Experiments 3 and 5. In Experiment 3, the subjects had been kept in single cages for 5–6 hr a day for 4 continuous days. In Experiment 5, the subjects were habituated to remaining in the testing cages for only ½ hr for other, unrelated experimental work. This large difference probably reflected a general agitation in anticipation of returning to their homecage. Regardless of how one interprets these results in terms of the impact of the mirror on scratching behavior, they demonstrate that the incidental mark contacts in Experiments 3 and 4 need not be interpreted as evidence that the subjects had detected the marks from olfactory or tactile cues.

General Discussion

For purposes of clarity, we divide our discussion of the results of the five experiments into several sections. We begin by offering our interpretation of what the combined results of the behavioral observations and experimental mark tests suggest about the presence of self-recognition in individual subjects. Next, we outline the developmental pattern indicated by our results and explore how these results mesh with previous research. Finally, we look at the patterns of emergence of self-recognition and offer suggestions for future research in this area. The evolutionary implications of these results are discussed in the context of each of these conceptual issues.

Diagnosing Self-Recognition

Spontaneous self-exploration versus mark tests. Traditional diagnoses of self-recognition in chimpanzees have been based on the converging results of observations of spontaneous behavior on the part of the subjects toward their mirror-images and on the results of mark tests. Swartz and Evans (1991) were the first to suggest, on the basis of results from 3 subjects, that these two indicators might be dissociable. However, additional evidence for this effect was reported by Suarez and Gallup (1981), although they dis-

counted the negative results of the mark test because of procedural difficulties. Our results confirm these earlier findings and demonstrate that when appropriate sample sizes are examined, half of the subjects classified as showing compelling evidence of spontaneous self-recognition do not show mirror-monitored mark-directed responding in a 30-min test. The extent of this dissociation is best illustrated in the context of the mark tests of the 3 SR+ subjects, which showed during the test period a number of compelling instances of using the mirror to explore their anal-genital regions, teeth, and other parts of their bodies that are not usually visible but made no effort to touch the facial marks. On the other hand, although about half of the subjects that were classified as showing self-recognition according to observations of spontaneous behavior did not pass the mark test, the reverse was not the case: Of the 12 subjects that were initially classified as SR– and given mark tests, only one displayed positive results on the mark test.

Although our initial assumptions were that the results of mark tests would be isomorphic with observations of spontaneous behavior, the potential separation of the two indexes of self-recognition must now be considered a real phenomenon. But why is there such a dissociation? To be sure, methodological factors may account for some of the failures to obtain mirror-prompted mark-directed responses. First, it is possible that some, but not all, of the subjects habituate to the marks on their wrists and hence discover that such marks are of no consequence; thus, when they see marks on their faces, they make no attempt to remove them (Gallup, in press). This explanation encounters difficulty when one witnesses that SR+ chimpanzees pick at their wrist marks during the control and test period but do not touch the marks on their faces, an effect that was observed in at least 1 of the chimpanzees in Experiment 4. Such a result seems difficult to reconcile with the hypothesis that habituation to the wrist marks generalizes to facial marks. A related account of the discrepant observational and mark test results is that because the subjects could see the marks on the other subject in the cage next to them, they may have habituated to facial marks on themselves (Gallup, in press). Again, this generalized habituation argument fails to account for the individual differences between the subjects that did touch the marks and those that did not because both have seen their partner's marks. Gallup (in press) suggested a number of strategies that may reduce these potential habituation problems.

An additional concern may stem from the number of mark contacts in the control sessions of Experiment 3. Some researchers may interpret this as evidence that the marks left tactile cues. This is erroneous, as the results of Experiment 5 demonstrate unequivocally. Our chimpanzees touched these areas even when no mark was present. This can be viewed as the ambient level of touching for facial regions targeted to be marked. The discrimination between mirror-monitored contacts and touches versus rubs is an indispensable component of our procedure, and with it we were able to rule out the objection that the chimpanzees could feel the marks. More important, perhaps, is that the median number of ambient mark contacts in the control sessions of Experiment 4 for all 30 subjects combined was

only 1, and the mean was only 2.5 (see Table 4). These rates of contacts are comparable with those reported by Gallup and colleagues, which have ranged from 1 to 3 in studies with sample sizes of 2 to 4 (Gallup, 1970; Gallup et al., 1971; Suarez & Gallup, 1981).

A final methodological issue is whether the discordant results of the mark tests and spontaneous SE behaviors are the result of the length of the test period. It is possible that 30 min is simply not long enough for all of the subjects to notice the marks and respond to them. There is at least one reason for taking this criticism seriously. Although most of the subjects that showed clear evidence of detecting the marks and responding to them did so within the first 5–10 min of the test session, there were several notable exceptions. For example, 1 adult chimpanzee spent a considerable amount of time looking into the mirror during the first 20 min of the test but did not respond to the marks until the final 10 min, when she did so in an unambiguous manner. However, although there may have been some subjects that needed more time, this does not appear to be a valid explanation for the dissociation in general. Given that most of the subjects did look into the mirror when it was unveiled, we are forced to ask what the possible underlying causes of such a long delay to touch the marks might be. At present, we are left to appeal to individual variation in attention or motivation.

We cannot definitively rule out these methodological issues, and there is reason to believe that some interaction of these variables may be responsible for some subset of the variance to be explained. However, there is an additional, conceptual explanation of the dissociation to which we now turn our attention. We believe that the individual variation in motivation and attention mentioned earlier may be embedded in a background of individual differences in the retention of self-recognition across time. Thus, in addition to any methodological problems, we propose that the apparent dissociation between the spontaneous results and the results of the mark tests may spring from the fact that many chimpanzees may need to relearn to recognize themselves during the mark test. Note that except for Experiment 4, all of our chimpanzees were given mark tests after a considerable time had elapsed (1–4 months) since their last encounter with a mirror. Although this may not matter for some chimpanzees, which form a rich or stable representation of their physical appearance, it may make a difference for those that do not. If such individual differences exist, this may explain several of the surprising results obtained within the context of the mark tests. First, if some chimpanzees relearned to recognize themselves within the 30-min test period, they learned that their appearance includes red marks on the ear and eyebrow ridge, and such information essentially leaves no unique motivation for the chimpanzee to touch these areas. Second, this also explains why several of the SR+ chimpanzees showed compelling SE behaviors within the 30-min test period but no touches to the facial marks. They may have learned that the image was themselves and may have been drawn to parts of their body that they could not directly see but not to the marks, which did not violate any existing representation of their physical appearance.

Although we offer this differential retention explanation only as a tentative working hypothesis, one may ask whether or not there is any existing evidence to support it. Calhoun and Thompson (1988) reported some evidence of immediate self-recognition on the part of 2 chimpanzees that had shown evidence of self-recognition 1 year before. However, an examination of their data finds that the SE results were compelling for only 1 of the subjects and that the mark-directed patterns were not compelling for either. This may well be due to the young ages of the animals, although 1 of them did display an interesting emotional reaction after the apparent discovery of lost teeth during the year of no mirror exposure.³ However, Calhoun and Thompson's (1988) research was described from the point of view that chimpanzees need 2 or 3 days to recognize themselves. Given our results on the rapidity of the emergence of this phenomenon, it cannot be stated with any certainty whether their subjects showed retention or simply relearned to recognize themselves.

Further support for the differential retention hypothesis comes from potential differences in the ways in which those SR+ chimpanzees that touched the marks and those that did not reacted to the initial uncovering of the mirror during the test period. Although we did not collect systematic data on the point, informal notes on the mark test data sheets reveal that many of these subjects showed at least some, and in many cases, dramatic social reactions to the mirror, such as head bobbing, aggression, lip flips, and sexual presentations. Such reinstatements of social reactions on the part of subjects that showed the dissociation may be a clue as to their initial interpretation of the image. However, even if one assumes that such an effect could be replicated under controlled conditions, this still leaves two possibilities. One is that the subjects were initially confused by the image but then quickly reinterpreted it as their own image, which they had previously stored in memory from their original exposure. This account predicts that chimpanzees that relearn to recognize themselves will touch the marks on their faces.⁴ The second possibility is that the subjects literally relearned the source of the image without any aid from previously stored mental representations of themselves. As noted earlier, this predicts that the chimpanzees will not touch the experimental marks. Unfortunately, we have not yet conducted the experiments (outlined later) that allow us to choose intelligently between the two alternatives.

³ It is tempting to speculate that this phenomenon is related to our findings with the 2 subjects that showed dramatic emotional reactions to their exploration of their bodies when they first encountered the mirror (see Experiment 1 Results).

⁴ A rather striking example that tends to support this idea is the results from an adult SR+ subject. This subject's initial reactions were social in nature, but by about 7–8 min into the test, she displayed several instances of contingent body and several minutes later contingent facial behavior. Next she showed some ambiguous self-exploratory behaviors directed at her eyes and lips, finally followed at 18 min by clear and unambiguous touches to the mark above her eye, and then immediately to the mark on her ear. After this the subject engaged in a variety of clear instances of using the mirror to groom and pick at her teeth and eyes.

Contingent body and facial movements versus self-exploration. Some researchers have argued that self-recognition can be detected by several distinct behavioral indexes, including SE, CB, and CF behaviors (see Parker, 1991). Our results do not support such a conclusion. The issue at hand is whether a given behavior indicates that the subject has successfully identified the image as itself or if the subject is merely exploring the contingencies imposed by mirrors. It is difficult to deny that chimpanzees that show clear and sustained levels of SE also display compelling instances of CB and CF, and the correlations between CB, CF, and SE behaviors were high in Experiment 1. On the other hand, it is also true that CB and CF movements can be detected in young (and old) chimpanzees that do not show SE behaviors or pass the mark test. This is important because although many chimpanzees that show SE behaviors did not pass a mark test, no subjects that showed only CF or CB behaviors pass the same test. In Experiment 3, for example, 2 of the chimpanzees displayed at least two clear and compelling instances of slowly and repeatedly moving a body part back and forth while alternately glancing from the mirror to the body part. Nonetheless, no evidence of self-exploration emerged, and the subjects did not pass the mark test.

We do not find such a result surprising when viewed in the context of other concurrently developing intellectual capacities, such as Piagetian sensorimotor stages. Indeed, following other authors, we believe that Stage 4 and 5 sensorimotor behaviors, which can be characterized by an organism's ability to detect the relation between its own actions and the environmental effects caused by those actions, are sufficient to allow an organism to determine that its own movement is causing the movement it witnesses in a mirror (see Parker, 1991). Experimenting with such relations are a classic indication of these stages of sensorimotor intelligence (Chevalier-Skolnikoff, 1983). Viewed from this perspective, it is not surprising that chimpanzees incapable of self-recognition display some indications of CB movements; it is also not surprising that CF movements occur at a lower frequency, probably because they involve unseen actions.

Although on first reflection it may seem puzzling that a chimpanzee may detect the yoked contingencies of its own actions and the actions it witnesses in a mirror without understanding that image is itself, the distinction between self-perception and self-conception can help to clarify this situation (see Butterworth, 1992). The self-perceptual system of the organism (which need not be described as being aware of its own psychology or personal agency) may well be able to detect and experiment with mirrored contingencies, just as it does in many other domains. This may also partially explain why some macaques, although incapable of self-recognition, are capable of detecting at least some contingencies in mirrors (Anderson, 1986; Brown, McDowell, & Robinson, 1965; Itakura, 1987a, 1987b). However, the detection and use of such contingencies need not imply that an organism is aware of the duality between itself and what it is responding to in the mirror. For this reason we do not favor schemes that attempt to equate CF or CB movements with self-recognition proper (e.g., Parker, 1991).⁵

Developmental Results

Age of onset. The ontogenetic pattern revealed by our data suggests that the underlying capacity for self-recognition in mirrors typically is not constructed in the chimpanzee mind until 4.5–8 years of age, if it is to be constructed at all. However, our data are limited from two directions. First, although we found no difference between full mother-reared and nursery-peer- or partially mother-reared subjects in terms of age of onset of self-recognition, our sample size for fully mother-reared subjects is both small (6 subjects) and young (age range of 10–28 months). Our data leave open the possibility that fully mother-reared chimpanzees may develop self-recognition at a younger age than nursery-peer-reared subjects. On the other hand, some combination of full rearing by the mother and being raised in an environment with many objects (i.e., wild-born and wild-reared chimpanzees) may lead to an even earlier emergence of self-recognition. From the opposite perspective, however, the second limitation of our data involves cross-fostered (home-reared) chimpanzees (e.g., Hayes, 1951). Our data leave open the possibility that chimpanzees reared in an environment of abnormal stimulations, such as those found in a human home or a psychological laboratory, may advance through the cognitive stages that lead to self-recognition at a faster rate, and this may ultimately result in an earlier onset of self-recognition. Indeed, the results of the longitudinal study in Experiment 2 revealed a higher percentage of SR+ 5-year-olds than would have been expected from the data on 5-year-olds in Experiment 1. The subjects of the longitudinal investigation were also part of an intensive research program in cognitive development that involved multiple, daily testing sessions with human researchers. Of course, both of these hypotheses may be incorrect. It may turn out that the epigenetic program governing the construction of the cognitive equipment necessary for self-recognition is so tightly canalized that rearing influences make little difference one way or the other (but see Povinelli, 1993b, in press).

Our conclusion about the age of onset of self-recognition differs from that reached by Lin et al. (1992), who investigated chimpanzees reared in a manner nearly identical to the subjects of our investigation. This is true despite the fact that we applied observational and experimental techniques to a sample size nearly 5 times as large and observed the subjects across five 8-hr days, rather than three 20- to 40-min sessions. Until further data are forthcoming, we consider the

⁵ In addition, the correlations we reported in Experiment 1 between contingent body movements, contingent facial movements, and self-exploration behaviors were all about the same: For CB with SE, $r = .482$; for CF with SE, $r = .513$, and for CF with CB, $r = .424$ (all $ps < .001$). In Parker's (1991) model, the failure to find mirror self-recognition in monkeys is explicable in terms of their absence of voluntary control of facial muscles coupled with their inability to understand causality and imitation in the facial modality. However, the demonstration that contingent body movements that do not involve the face are equally correlated with self-exploratory behaviors suggests that facial movements have no special status in relation to self-recognition.

results of Lin et al.'s investigation to be the result of a failure to distinguish between compelling and weak instances of self-exploration,⁶ along with an inappropriate marking procedure during the experimental sessions.

What of the developmental findings other than those of Lin et al. (1992)? As we indicated earlier, the exact ages of the chimpanzees in Gallup's (1970) study were not known, although Gallup's (1975) description of the 2 females that participated in the original study indicated that they were undergoing sexual swelling, which strongly suggests that these animals were at least 6–8 years of age. In addition, Gallup (personal communication, October 10, 1991) provided the original data sheets for all 4 test subjects in the 1970 experiment, and the number of touches to the marked areas range from 4 to 10. Body weights for these subjects were 15, 20, 22, and 24 kg. On the basis of long-term records of the age–body weight relation in our colony of chimpanzees and making a 15%–20% downward adjustment because Gallup's (1970) subjects were wild-born, we estimate that these animals were between 5 and 8 years of age. Gallup (1970) did not report distinguishing between mirror-monitored and non-mirror-monitored contacts or between touches and rubs. From these considerations we conclude that not all of those chimpanzees or those tested subsequently recognized themselves (see later discussion), and the ones that did may easily have been as old as 5–8 years of age.

One way of casting the developmental question in a more predictive fashion is to consider the relationship between certain traditional indicators of cognitive achievement and self-recognition. In terms of human development, there is reason to suspect that the onset of self-recognition and the achievement (or closure) of Stage 6 of sensorimotor intelligence may be related. With the weight of some empirical findings in human infants, Bertenthal and Fischer (1978) argued that the skills needed to solve the mark test (*rouge test*) may be related to Stage 6 object permanence tasks. However, there was no consistent relation between the two. Thus, although certain sensorimotor cognitive skills may be necessary for self-recognition, they are certainly not sufficient to explain its emergence. If such a relation holds true across species (and we see no a priori reason why it ought not), then we may expect that the emergence of self-recognition behaviors in chimpanzees ought to follow on the heels of later Piagetian sensorimotor stages.⁷ If true, then to the extent that self-recognition in chimpanzees is delayed in relation to that in humans, then chimpanzees may also display a parallel delay in the development of sensorimotor intelligence.

A number of investigators have sought to determine the course of Piagetian sensorimotor stages in chimpanzees. Nearly all of them have concluded that although chimpanzees and orangutans pass through the initial four stages at roughly the same rate as human infants, Stages 5 and 6 are delayed and may not be completed until 5–8 years of age (Chevalier-Skolnikoff, 1983; Mathieu & Bergeron, 1981; Mignault, 1985), which is precisely the age at which we detected a developmental transition in which the majority of chimpanzees were diagnosed as capable of recognizing themselves in mirrors.⁸ However, it is premature to conclude anything about the exact nature of the relation between sen-

sorimotor intelligence and self-recognition, other than that if they remain linked across species in this fashion, the construction of the self-conceptual capacity underlying self-recognition may be organized in such a fashion that it requires as a basis some of the skills inherent in some subset of the later stages of sensorimotor intelligence. This in no way necessarily implies that Stage 6 object permanence tasks, for example, require the exact same conceptual skills as mirror self-recognition. Rather, it may simply be that the two are ontogenetically nondissociable, because one presupposes the other. This is of interest from an epigenetic standpoint because, if true, it suggests that evolution cannot act to produce ontogenies in which the emergence of Stage 6 sensorimotor intelligence and self-recognition are inverted.

Of particular interest to us are the occasional chimpanzee subjects, such as Megan, who show extremely precocious patterns of self-recognition. Although we originally believed that Megan was unusual in the rapidity with which she learned to recognize itself, she in fact turned out to be one

⁶ One could maintain that our methods are not sensitive to the behavioral subtleties of chimpanzees at this age. In other words, perhaps we were too conservative in our ethogram and therefore only classified the animals that showed overwhelming evidence of self-recognition as SR+. Of course, this contention must be wedded to an additional ad hoc explanation of the correlation between the results of our behavioral classification scheme and the results of the mark test in Experiment 3 (i.e., the marks were not salient enough). Nonetheless, we recognize the general thrust of the argument as plausible. We are therefore undertaking several additional experiments to determine if we were too conservative in our measurements. First, we have tried a *perm-shock* technique in which SR+ and SR- 3.5- to 5-year-olds have their hair bleached to a bright orange color. In keeping with our position, the 1 SR+ subject showed marked increases (in relation to baseline measures) in touching and rubbing its head when a mirror was exposed, but the 2 SR- subjects did not (Povinelli, 1993a). We are currently testing additional subjects in both categories. Finally, we are comparing the number of face and head touches, scratches, rubs, and contingent facial and body movements by young chimpanzees while they view either a mirror or a television monitor showing other chimpanzees. Again, our preliminary data show high levels of these ambiguous behaviors regardless of whether the stimulus is a mirror or a videotape (Povinelli, 1993a). These can be contrasted with the clear instances of self-exploration that we required for a SR+ diagnosis throughout our experiments.

⁷ We generalize from the results of Stage 6 object permanence tasks to sensorimotor intelligence in general. Although this is a somewhat questionable practice, the emergence of the stages in human infants are reasonably correlated with each other. Given the paucity of knowledge in this area with respect to chimpanzees, finer distinctions at this point are difficult.

⁸ Other data also suggest a general delay in chimpanzee ontogeny (or general acceleration in human ontogeny) that would be roughly consistent with this line of argument. For instance, Spinuzzi (1993) recently reported that chimpanzees do not show evidence of second-order classification (spontaneous constructions of same-different relations between two related sets) until the 5th year of life, a phenomenon that occurs by 18–24 months of age in human children. This is parallel to the fact that children recognize themselves in mirrors by 18–24 months, but chimpanzees do not appear to do so until 5–8 years of age.

of the subjects who took the longest to show evidence of SE. Our overall results now reveal that her uniqueness may lie along ontogenetic lines instead. This difference may not merely be in the domain of self-recognition. To the extent that there is merit to our arguments, Megan ought to have shown evidence of a compressed sensorimotor developmental schedule, in comparison with her companions. Even if the relation between sensorimotor intelligence and self-recognition collapses, Megan (and other young chimpanzees like her) may well represent the variation inherent in their species's epigenetic program governing the construction of the general cognitive capacities necessary for self-recognition. In this sense, depending on one's assumption about the polarity of this character state, she may be thought of as an example of the fodder on which natural selection acted to produce the more rapid development of self-recognition in human infants. Alternately, she may be viewed as the remnant of the primitive ancestral character state from which chimpanzees have deviated in their evolutionary history.

Decline in adulthood. Our results suggest either a precipitous decline in self-recognition between 16 and 20 years of age or a critical period effect. To be sure, many researchers will find either possibility difficult to accept, as do we. However, there are several lines of evidence that force us to give both possibilities consideration. To begin, our results essentially mirror those of Swartz and Evans (1991). They investigated a small sample of chimpanzees (11 subjects), whose exact ages were not known. They reported that only 3 of the subjects showed patterns of self-directed responding, and only 1 of the 3 passed the mark test. Our results, in a very large sample, may explain their results in the context of a peak in self-recognition in adolescence and young adulthood, followed by a decline in adulthood. Second, our extensive (albeit unsuccessful) efforts to induce self-recognition in a group of older animals seems to rule out such methodological issues as mere length of exposure. Despite 20 continuous days of mirror presentation, 5 adult subjects that were classified as SR- continued to stare at their images without any overt signs of self-recognition. In addition, 2 of these subjects participated in Experiment 4 and showed no attempts to touch the marks on their faces, although they did explore control marks on their wrists.

There is still room for methodological explanations for the lower percentage of SR+ subjects among adult chimpanzees. It is possible that the adults were less motivated or interested in the problem posed by the introduction of the mirror or, conversely, that captive adult chimpanzees are inherently abnormal. Although we recognize these as possibilities, we do not favor such interpretations. To begin, there was no difference in the mean time spent looking into the mirror between the adult group and the 7- to 16-year-old group (see Figure 4). This is important because the majority of the 7- to 16-year-olds were diagnosed as SR+, whereas only about 25% of the adults were similarly classified. In addition, there was no overall difference in viewing time between SR+ and SR- subjects. Thus, the adults showed as much interest in the mirror as the adolescents and young adults. Finally, when one compares the viewing times in our

study with those of Gallup (1975), our subjects (both adult and immature) showed nearly identical levels of interest in the mirror (see Experiment 1 Results). As for the potential abnormality of the adults, the groups tested were stable breeding units, most with 1-3 young infants in the group. Beyond the mild stereotypies found in virtually all captive chimpanzees, none of the adult subjects displayed any obvious abnormal behaviors, such as hyperaggressivity or repetitive rocking behaviors; this is probably due to the center's long-term emphasis on stable breeding units and indoor-outdoor housing. Finally, previous reports of a relation between social deprivation and an absence of self-recognition in chimpanzees are now confounded by the potential age discrepancy between the wild-born and isolation-reared groups (Gallup et al., 1971).

Some researchers may find the possibility of a decline in the capacity for self-recognition in adulthood or a critical period effect difficult to accept from an intuitive standpoint, arguing that such a conclusion is tantamount to saying that many adults do not have the capacity to reason about socially complex situations. However, a moment's reflection on other socially complex societies (say, of macaque species) exposes the fundamental flaw in this line of thinking. Macaque species move through a highly social milieu and engage in a variety of forms of manipulation of others (Whiten & Byrne, 1988). However, the best available evidence suggests that they neither recognize themselves in mirrors nor reason about intentions or knowledge (Cheney & Seyfarth, 1990; Gallup, 1982; Povinelli, Parks, & Novak, 1991, 1992). As many authors have pointed out, there are many routes to social complexity, not all of which require introspective modeling of the experiences of others. Loosely phrased, one need not possess a theory of mind to negotiate and manipulate a social universe. Indeed, if Gallup (1982) was correct about a relation between the presence of self-recognition and the presence of mental state attribution, our data suggests that for chimpanzees the peak in these abilities may occur in late adolescence and early adulthood (8-15 years; see Figure 3). Just as in humans, once social scripts are established through a theory-of-mind route, they may become routinized and largely unconscious.

Patterns of Emergence

The results of our experiments provide little support for the hypothesis that chimpanzees need 2 or 3 days of continuous mirror exposure before they come to recognize themselves. Only a single subject (age 6 years, 11 months) ultimately classified as SR+ was observed to show the first compelling SE behavior after Day 1 of mirror exposure. (However, this subject was different from most other subjects, both SR+ and SR-, in that we did not observe him visually exploring the mirror until Day 4.) Virtually all other subjects that were classified as SR+ showed compelling evidence within the first 2 hr of mirror presentation. Many subjects showed evidence within 5-20 min. Much of this variance may be explained by the differential initial reactions to the presentation itself. Some subjects stayed away from the mirror for a number of minutes before investigating it more carefully.

We believe that two factors may explain why previous researchers were led to conclude that self-recognition does not occur for 2 or 3 days. First, the sampling methods used by Gallup (1970; Gallup et al., 1971), and subsequently adopted by other researchers, may have missed the onset of self-recognition in those animals that did, in fact, recognize themselves. The subjects were observed twice a day, with each observation session lasting 15 min. If one considers how a typical adolescent SR+ chimpanzee in our study reacted, there would have been no SE behavior observed in the first 15-min session. In addition, given the rapid decline (and relative infrequency of the events), there is no guarantee that these behaviors would have been detected in the second session either. Thus, the failure to find evidence on Days 1 and 2 may have simply been due to sampling error. (We avoided this problem by using both a focal-animal and a focal-mirror sampling procedure and observed all animals for the first 3 hr of mirror exposure.) A second contributing factor for this delay may have been that Gallup's (1970) subjects were tested in the absence of other conspecifics and hence the source of their image may have remained more obscure. Several researchers have unsuccessfully attempted to facilitate self-recognition in a number of monkey species by giving them access to mirrors in the context of a social group, reasoning that if the subject sees familiar animals both beside them in real space and in the mirror, this might lead the subjects in a more immediate manner to question the source of their own reflection (see Benhar, Carlton, & Samuel, 1975; Gallup, Wallnau, & Suarez, 1980). Nonetheless, despite the fact that such a manipulation has no effect on monkeys, it may well have a prodding effect on the chimpanzee mind and result in faster rates of self-recognition.

There is a third potential contributing factor to the difference between our results and previous research. As indicated earlier, it may be that some of the subjects Gallup (1970; Gallup et al., 1971) tested did not, in fact, recognize themselves simply because they were too young. These subjects displayed social reactions for nearly 2 days, consistent with the peak in social reactions we found (see Figures 5 and 7). Given that these social reactions lasted for 2 days, baseline self-directed activities that are present in all chimpanzees of all ages with or without the presence of a mirror may have been displaced. Thus, for the first 2 days or so, the chimpanzees' visual explorations of the mirror may have been inducing social reactions. Only later, after their habituation to the image, did the chimpanzees begin to visually explore the mirror while engaged in baseline levels of self-directed behaviors. To be sure, we do not doubt that many of the chimpanzees Gallup (1970; Gallup et al., 1971) tested showed a full suite of compelling self-exploratory behaviors. However, others may simply have displayed normal baseline self-directed behaviors that would have been present without the mirror. In our experiments these would have been classified as weak instances of SE, CB, or CF. That this explanation deserves careful consideration is now apparent (with over 20 years of hindsight) from an inspection of Gallup's (1975) in-depth descriptions of his chimpanzees' reactions to the mirror procedure. Some of the behaviors classified as indicative of self-recognition (Gallup, 1975, pp. 323–324)

are examples of what would now be classified as contingent body or contingent facial movements. As indicated earlier, our own research shows that these behaviors do not necessarily indicate SE behavior or positive results on mark tests.

Some researchers may feel that our results indicate that our chimpanzees (unlike Gallup's, 1970) may have had prior experience with reflective surfaces, such as stainless steels cages or water surfaces. First, there is no reason to suspect that Gallup's (1970) subjects had less experience with such surfaces. Second, virtually all of our subjects initially reacted in ways that suggested nonrecognition (aggressive or appeasing social gestures made to the mirror), which indicates their initial absence of recognition. Finally, and most importantly, of the 50 or so animals with the most recent history of very poorly reflective surfaces (the walls of nursery sleeping cages), only 1 showed evidence of self-recognition. It may also be argued that our position ignores the fact that Gallup's (1970) chimpanzees displayed social gestures until about the 2nd or 3rd day. However, our additional experiment with 6 young mirror-experienced chimpanzees shows this objection to be without obvious merit as well. The single SR+ chimpanzee (Megan) showed immediate evidence of retention of self-recognition but also showed one of the highest levels of social behavior during the 4 days of testing (see Figure 7). This is not surprising to us, because although social gestures can in no way be used as unique evidence of self-recognition, they are not inconsistent with the presence of the capacity. Animals that show clear signs of self-recognition may continue to show apparent social behaviors (such as displaying or presenting) as part of their interest in their physical appearance and overt behavior.

Population Distribution

The final issue we hoped to address by this research program is related to whether the capacity for self-recognition in chimpanzees is a polymorphic trait within the species. Our results suggest that it may be. In no age group did we find self-recognition in all of the subjects. In the 8- to 15-year-old group, we found that 75% of the subjects showed SE behaviors. Nonetheless, even after assuming methodological limitations, one is left to account for the nearly 15% of the subjects that failed to show any compelling evidence of recognizing themselves. It is possible that some chimpanzees may simply never construct the cognitive infrastructure that supports self-recognition. In the context of the gorilla debate, Gallup (1991; personal communication, October 10, 1986) has already speculated that such patterns may be the result of populational differences of the sort that most (but not all) humans are capable of self-recognition, fewer chimpanzees and orangutans, and very few gorillas. However, the proportion of chimpanzees that at some point during their lifetime pass through a period of self-conception strong enough to support self-recognition is certainly much higher than that implied by Swartz and Evans (1991).

The peak in SR+ subjects (75%) we detected between 8–15 years of age suggests an alternative interpretation. It may be that most chimpanzees are capable of self-recognition during at least part of their lifespan but that in

comparison with humans that period is dramatically restricted in many cases. Thus, it may be that the polymorphisms in question have more to do with variations in ontogenetic programs controlling the onset and offset of the capacity, rather than its simple presence or absence. Outliers, such as Megan, certainly support this view. However, this must remain speculation until the results of longitudinal studies tracking the life-long development of self-recognition in a sizable number of chimpanzees are completed. Such research is currently underway in our laboratory.

Future Research

The results of our investigation highlight the need for additional research in this area. Indeed, we believe that further investigations may be needed to clarify all of the major issues addressed in this report. In addition, we believe that additional questions related to the underlying meaning of self-recognition can be answered by the use of new techniques to probe the chimpanzee's sense of self.

First, the dissociation between observations of spontaneous, clear, self-exploratory behavior and detection of facial marks during experimental testing needs to be investigated further. Several methodological improvements ought to be followed in all future investigations. Marks need not be applied to the wrists of the subjects as a control procedure. Our results demonstrate that all classes of chimpanzees touch marks applied to their wrists. Eliminating these will eliminate possible habituation to the inconsequential nature of such marks in general (see Gallup, in press, for alternate strategies). Also, subjects must be tested without visual access to other marked chimpanzees in order to reduce another problem of habituation (Gallup, in press). We gave our subjects visual access to each other in order to address the criticism that chimpanzees (or young children) that touch the marks after the mirror is reintroduced may be doing so merely because they see another chimpanzee (or child) that is marked and reach up to determine if they have marks also (Mitchell, in press). Our results show that there is no need to accept this copycat hypothesis because our subjects that passed the mark test all had 4–7 hr of visual exposure to their marked partner before the unveiling of the mirror. Finally, we recommend systematic comparisons of mark tests with SR + chimpanzees that are tested immediately after their initial block of mirror exposure and those that are tested after an appreciable delay. This will help to determine whether the dissociation is because some subjects fail to retain a visual representation of the self.

From a related perspective, we advocate the use of more careful observations during the mark tests themselves. In addition to data on the number and topography of mark-directed responses, the patterns of emergence of these touches in relation to social and other self-exploratory and contingent movements ought to be recorded. Such data, in combination with manipulations of the length of delay of the mark test after initial mirror exposure, will allow researchers to select between competing accounts of the underlying causes of the dissociation between observations of spontaneous behavior and the results of mark tests.

Finally, we advocate (and have begun to undertake) a parallel research program on the nature of the self-concept that supports self-recognition. Systematic comparisons between children and chimpanzees on a variety of tasks will help to discern whether we can meaningfully discriminate between mature and immature versions of the capacity for self-recognition (see Gallup & Povinelli, in press; Povinelli, 1993b). In one such experiment that is currently underway, the results of immediate versus delayed videotape tests of self-recognition in both chimpanzee and human subjects known to recognize themselves are being compared. Consider the following paradigm. A young child who is involved in a special task of some kind is videotaped while an experimenter surreptitiously places a sticker on her forehead. Although the subject is unaware of being marked, the videotape clearly depicts what has transpired. Next, the child is asked if she would like to watch herself on television, and the tape is played back for her. The results of our initial experiments with over 100 2-, 3-, and 4-year-old children suggest a developmental asynchrony between mirror self-recognition (in the mark test) and the delayed mark test just described. Although children pass the mirror test at 18–24 months, they do not pass the delayed video (or an analogous delayed photograph) mark test until they are between 3.5 to 4 years of age (Povinelli, Landau, & Perilloux, 1993). We suspect that other changes in the child between 3.5 to 4 years may alter the child's conception of self from one that is largely constrained to the present to one that simultaneously is able to integrate past and previous self-representations. This task is clearly ideal for use with chimpanzees because it allows for a nonlinguistic assessment of their self-concept through time.

We believe that the investigations suggested here will provide the basis for exploring other aspects of the chimpanzee's mind related to mental state attribution. If our results can be replicated, this will go a long way toward emphasizing that the phrase *the chimpanzee's mind* implies a homogeneity of psychological processes across members of the species that does not exist. Rather than being a source for alarm, such between-subjects and developmental variation will ultimately provide a even richer source for testing models that attempt to provide a coherence between the evolution of the capacity for self-recognition and the evolution of an understanding of the mental world (Gallup, 1982, 1985, 1991). Such research remains in its infancy, but methodological advances in developmental and comparative psychology promise to nurture such programs to adolescence soon. From this maturity, evolutionary biologists interested in the origins of psychological capacities will be able to harvest the data necessary to reconstruct the evolution of mental state attribution (see Povinelli, 1993b). Further explorations of the meaning and extent of the chimpanzee's capacity for self-recognition will play an intimate part in assisting this reconstruction.

References

- Amsterdam, B. (1972). Mirror self-image reactions before age two. *Developmental Psychobiology*, 5, 297–305.
- Anderson, J. R. (1983). Responses to mirror image stimulation and

- assessment of self-recognition in mirror- and peer-reared stump-tail macaques. *Quarterly Journal of Experimental Psychology*, 35B, 201–212.
- Anderson, J. R. (1984). The development of self-recognition: A review. *Developmental Psychobiology*, 17, 35–49.
- Anderson, J. R. (1986). Mirror-mediated finding of hidden food by monkeys (*Macaca tonkeana* and *M. fascicularis*). *Journal of Comparative Psychology*, 100, 237–242.
- Benhar, E. E., Carlton, P. L., & Samuel, D. (1975). A search for mirror-image reinforcement and self-recognition in the baboon. In S. Kondo, M. Kawai, & S. Ehara (Eds.), *Contemporary primatology: Proceedings of the 5th International Congress of Primatology* (pp. 202–208). New York: Karger.
- Berenthal, B. I., & Fischer, K. W. (1978). Development of self-recognition in the infant. *Developmental Psychology*, 14, 44–50.
- Brown, W. L., McDowell, A. A., & Robinson, E. M. (1965). Discrimination learning of mirrored cues by rhesus monkeys. *Journal of Genetic Psychology*, 106, 123–128.
- Butterworth, G. (1992). Origins of self-perception in infancy. *Psychological Inquiry*, 3, 103–111.
- Calhoun, S., & Thompson, R. L. (1988). Long-term retention of self-recognition by chimpanzees. *American Journal of Primatology*, 15, 361–365.
- Cheney, D. L., & Seyfarth, R. L. (1990). Attending to behaviour versus attending to knowledge: Examining monkeys' attribution of knowledge states. *Animal Behaviour*, 40, 742–753.
- Chevalier-Skolnikoff, S. (1983). Sensorimotor development in orang-utans and other primates. *Journal of Human Evolution*, 12, 545–561.
- Gallup, G. G., Jr. (1968). Mirror-image stimulation. *Psychological Bulletin*, 70, 782–793.
- Gallup, G. G., Jr. (1970). Chimpanzees: Self-recognition. *Science*, 167, 86–87.
- Gallup, G. G., Jr. (1975). Toward an operational definition of self-awareness. In R. H. Tuttle (Ed.), *Socio-ecology and psychology of primates* (pp. 309–341). The Hague, the Netherlands: Mouton.
- Gallup, G. G., Jr. (1982). Self-awareness and the emergence of mind in primates. *American Journal of Primatology*, 2, 237–248.
- Gallup, G. G., Jr. (1985). Do minds exist in species other than our own? *Neuroscience and Biobehavioral Reviews*, 9, 631–641.
- Gallup, G. G., Jr. (1991). Toward a comparative psychology of self-awareness: Species limitations and cognitive consequences. In G. R. Goethals & J. Strauss (Eds.), *The self: An interdisciplinary approach* (pp. 121–135). New York: Springer-Verlag.
- Gallup, G. G., Jr. (in press). Self-recognition: Research strategies and experimental design. In S. Parker, R. Mitchell, & M. Boccia (Eds.), *Self-awareness in animals and humans*. Cambridge, United Kingdom: Cambridge University Press.
- Gallup, G. G., Jr., McClure, M. K., Hill, S. D., & Bundy, R. A. (1971). Capacity for self-recognition in differentially reared chimpanzees. *The Psychological Record*, 21, 69–74.
- Gallup, G. G., Jr., & Povinelli, D. J. (in press). Mirror, mirror, on the wall, which is the most heuristic theory of them all? *New Ideas in Psychology*.
- Gallup, G. G., Jr., Wallnau, L. B., & Suarez, S. D. (1980). Failure to find self-recognition in mother-infant and infant-infant rhesus monkey pairs. *Folia Primatologica*, 33, 210–219.
- Hayes, C. (1951). *The ape in our house*. New York: Harper.
- Hill, S. D., Bundy, R. A., Gallup, G. G., Jr., & McClure, M. K. (1970). Responsiveness of young nursery-reared chimpanzees to mirrors. *Proceedings of the Louisiana Academy of Sciences*, 33, 77–82.
- Itakura, S. (1987a). Mirror guided behavior in Japanese monkeys (*Macaca fuscata fuscata*). *Primates*, 28, 149–161.
- Itakura, S. (1987b). Use of a mirror to direct their responses in Japanese monkeys (*Macaca fuscata fuscata*). *Primates*, 28, 343–352.
- Johnson, C. B. (1983). Self-recognition in infants. *Infant Behavior and Development*, 6, 211–222.
- Ledbetter, D. H., & Basen, J. A. (1982). Failure to demonstrate self-recognition in gorillas. *American Journal of Primatology*, 2, 307–310.
- Lethmate, J., & Dücker, G. (1973). Untersuchungen am ebsterkenen im spiegel bei orangutans einigen anderen affenarten [Self-recognition by orangutans and some other primates.] *Zeitschrift für Tierpsychologie*, 33, 248–269.
- Lewis, M. (1986). Origins of self-knowledge and individual differences in early self-recognition. In A. G. Greenwald & J. Suls (Eds.), *Psychological perspective on the self* (pp. 55–78). Hillsdale, NJ: Erlbaum.
- Lewis, M., & Brooks-Gunn, J. (1979). *Social cognition and the acquisition of self*. New York: Plenum Press.
- Lewis, M., Sullivan, M. W., Stanger C., & Weiss, M. (1989). Self-development and self-conscious emotions. *Child Development*, 60, 146–156.
- Lin, A. C., Bard, K. A., & Anderson, J. R. (1992). Development of self-recognition in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 106, 120–127.
- Mathieu, M., & Bergeron, G. (1981). Piagetian assessment on cognitive development in chimpanzees (*Pan troglodytes*). In A. B. Chiarelli & R. S. Corruccini (Eds.), *Primate behavior and sociobiology* (pp. 142–147). New York: Springer-Verlag.
- Mignault, C. (1985). Transition between sensorimotor and symbolic activities in nursery-reared chimpanzees (*Pan troglodytes*). *Journal of Human Evolution*, 14, 747–758.
- Mitchell, R. W. (in press). Mental models of mirror self-recognition: Two theories. *New Ideas in Psychology*.
- Patterson, F. (1984). Self-recognition by *Gorilla gorilla gorilla*. *Gorilla*, 7, 2–3. (Available from the Gorilla Foundation, Box 620–530, Woodside, CA 94062)
- Parker, S. T. (1991). A developmental approach to the origins of self-recognition in great apes. *Human Evolution*, 6, 435–449.
- Povinelli, D. J. (1987). Monkeys, apes, mirrors, and minds: The evolution of self-awareness in primates. *Human Evolution*, 2, 493–507.
- Povinelli, D. J. (1989). Failure to find self-recognition in Asian elephants (*Elephas maximus*) in contrast to their use of mirror cues to discover hidden food. *Journal of Comparative Psychology*, 103, 122–131.
- Povinelli, D. J. (1993a). [Differential reactions to alterations in hair color by chimpanzees, which do and do not recognize themselves in mirrors]. Unpublished raw data.
- Povinelli, D. J. (1993b). Reconstructing the evolution of mind. *American Psychologist*, 48, 493–509.
- Povinelli, D. J. (in press). How to create self-recognizing gorillas (but don't try it on macaques). In S. T. Parker, M. L. Boccia, & R. W. Mitchell (Eds.), *Self-awareness in animals and humans*. Cambridge, United Kingdom: Cambridge University Press.
- Povinelli, D. J., & deBlois, S. (1992). Young children's (*Homo sapiens*) understanding of knowledge formation in themselves and others. *Journal of Comparative Psychology*, 106, 228–238.
- Povinelli, D. J., Landau, K., & Perilloux, H. K. (1993). *Self-recognition in young children using delayed versus live feedback: Evidence of a developmental asynchrony*. Unpublished manuscript.
- Povinelli, D. J., Nelson, K. E., & Boysen, S. T. (1990). Inferences about guessing and knowing by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 104, 203–210.

- Povinelli, D. J., Nelson, K. E., & Boysen, S. T. (1992). Comprehension of social role reversal by chimpanzees: Evidence of empathy? *Animal Behaviour*, *43*, 633-640.
- Povinelli, D. J., Parks, K. A., & Novak, M. A. (1991). Do rhesus monkeys (*Macaca mulatta*) attribute knowledge and ignorance to others? *Journal of Comparative Psychology*, *105*, 318-325.
- Povinelli, D. J., Parks, K. A., & Novak, M. A. (1992). Role reversal in rhesus monkeys, but no evidence of empathy. *Animal Behaviour*, *44*, 269-281.
- Priel, B., & de Schonen, S. (1986). Self-recognition: A study of a population without mirrors. *Journal of Experimental Child Psychology*, *41*, 237-250.
- Robert, S. (1986). Ontogeny of mirror behavior in two species of great apes. *American Journal of Primatology*, *10*, 109-117.
- Schulman, A. H., & Kaplowitz, C. (1977). Mirror image response during the first two years of life. *Developmental Psychobiology*, *10*, 133-142.
- Senden, M. von. (1960). *Space and sight: The perception of space and shape in the congenitally blind before and after operation*. Glencoe, IL: Free Press.
- Spinozzi, G. (1993). Development of spontaneous classificatory behavior in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *107*, 193-200.
- Suarez, S. D., & Gallup, G. G., Jr. (1981). Self-recognition in chimpanzees and orangutans, but not gorillas. *Journal of Human Evolution*, *10*, 175-188.
- Swartz, K. B., & Evans, S. (1991). Not all chimpanzees show self-recognition. *Primates*, *32*, 483-496.
- Whiten, A., & Bryne, R. W. (1988). Tactical deception in primates. *Behavioral and Brain Sciences*, *11*, 233-244.

Received October 30, 1992

Revision received January 15, 1993

Accepted January 19, 1993 ■

Search Opens for Editor of New APA Journal

The Publications and Communications Board has opened nominations for the editorship of a new journal, *Psychological Methods*, for the years 1996-2001. Candidates must be members of APA and should be prepared to start receiving manuscripts early in January of 1995 to prepare for issues published in 1996 and beyond. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. To nominate candidates, prepare a statement of one page or less in support of each candidate. Submit nominations to

Donald J. Foss, PhD
 Psychology Department
 University of Texas
 Austin, TX 78712
 or
 Foss@psyvax.psy.utexas.edu

Other members of the search committee are Jay Belsky, PhD, Bert F. Green, Jr., PhD, Douglas N. Jackson, PhD, and Robert Rosenthal, PhD.

Psychological Methods will be devoted to the development and dissemination of methods for collecting, understanding, and interpreting psychological data. Its purpose is the dissemination of innovations in research design, measurement, methodology, and statistical analysis to the psychological community; its further purpose is to promote effective communication about related substantive and methodological issues. The audience is diverse and includes those who develop new procedures, those who are responsible for undergraduate and graduate training in design, measurement, and statistics, as well as those who employ those procedures in research. The journal solicits original theoretical, quantitative, empirical, and methodological articles; reviews of important methodological issues; tutorials; articles illustrating innovative applications of new procedures to psychological problems; articles on the teaching of quantitative methods; and reviews of statistical software. Submissions will be judged on their relevance to understanding psychological data, methodological correctness, and accessibility to a wide audience. Where appropriate, submissions should illustrate through concrete example how the procedures described or developed can enhance the quality of psychological research. The journal welcomes submissions that show the relevance to psychology of procedures developed in other fields. Empirical and theoretical articles on specific tests or test construction should have a broad thrust; otherwise, they may be more appropriate for *Psychological Assessment*.

First review of nominations will begin December 15, 1993.