Face recognition in primates: a cross-species study

Olivier Pascalis, Jocelyne Bachevalier *

Department of Neurobiology and Anatomy, University of Texas Health Science Center, 6431 Fannin, Houston, TX 77030, USA

Received 8 April 1997; received in revised form 18 November 1997; accepted 18 November 1997

Abstract

Recognition for human faces, monkey faces, and objects was assessed in both adult humans (Homo sapiens) and monkeys (Macaca mulatta) with a visual paired-comparison task. The results demonstrated that while both species showed strong novelty preference for objects, human participants showed novelty preference for human faces but not for monkey faces, and vice versa for the monkeys. This 'species-specific effect' in face recognition is discussed in relation with data on both the 'other-race effect' observed in humans and the effects of experience or training on face recognition processes in primates. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Recognition; Faces; Humans; Monkeys; Comparative study

1. Introduction

Face recognition is known to play a crucial role in the establishment and maintenance of social communication in primates (Chevalier-Skolnikoff, 1973; Redican, 1975; Petit and Thierry, 1992). It provides a rapid and powerful mechanism by which an individual can distinguish individuals from the same, or a different, species to ensure its survival. Humans possess a neural mechanism specialized for the processing of faces (Haxby et al., 1995; Sergent et al., 1995) and recent experiments have indicated that such a specialized system seems also to be present in non-human primates. Thus, behavioral experiments have shown that not only can monkeys process a two-dimensional picture of a face as a face prototype, and not simply as a visual object (Rosenfeld and Van Hoesen, 1979; Keating and Keating, 1982; Overman and Doty, 1982; Sands et al., 1982; Boysen and Bernstein, 1986; Perrett et al., 1988; Boysen and Bernstein, 1989), but they can also recognize faces of individuals of their own species (Bruce, 1982; Overman and Doty, 1982; Wright and Roberts, 1996).

A growing interest has emerged in recent years in determining how face recognition processing compares in human and non-human primates. Despite the paucity of comparative studies using
identical, well-controlled, experimental designs, face processing appears to share some similarities in the two species. For example, the orientation of a face has a profound effect on its recognition in humans, with inverted faces being recognized with greater difficulty than upright faces. This inversion effect is exclusive to face processing since it is not observed when objects are used as stimuli (Ellis and Shepherd, 1975; Philips and Rawles, 1979) and, therefore, has led to the view that face processing in humans depends upon a specialized neural mechanism (Yin, 1969). Research in monkeys has provided mixed results. Bruce (1982) found no inversion effects with a concurrent discrimination task using pictures of monkey faces as stimuli. However, several other studies have successfully demonstrated an inverted effect in monkeys. First, Overman and Doty (1982) using a match-to-sample task found a significant decline in performance for inverted human and monkey faces but not for scenes. More recently, using a match-to-sample task, Phelps and Roberts (1994) found an inversion effect in human subjects and in one monkey, but this inversion effect appeared only for human faces and not for monkey faces or scenes. This latter finding has been replicated by a recent systematic analysis of the inversion effect in both humans and monkeys (Wright and Roberts, 1996), suggesting the presence of similar face processing mechanisms in monkeys and humans.

A second characteristic of face processing observed in both humans and monkeys pertains to the feature analysis of a face. Thus, when examining familiar faces, human participants perform better when analyzing the internal features rather than the external features of a face (Ellis et al., 1979). Correspondingly, when viewing face stimuli, monkeys spent more time scanning the internal features rather than the external features of a face (Keating and Keating, 1982, 1993).

A final example of similar face processing in primates is the presence of specialized cortical areas which subserve this mechanism. Both lesion, recording, and brain imaging experiments in humans have revealed specific cortical areas in the ventral part of the occipitotemporal junction which are involved in face recognition (for review see Tovee and Cohen-Tovee, 1993). Likewise, electrophysiological studies in non-human primates have demonstrated neurons in the superior temporal sulcus of the inferior temporal cortex which discharge specifically to face stimuli. These cells respond to a variety of photographs of human and monkey faces, as well as to the recognition of specific individuals, facial emotional expressions, and facial orientations (Rolls, 1984; Perrett et al., 1988; Desimone, 1991).

While the similarities listed above argue in favor of a common neural mechanism for face recognition in primates, it remains unclear whether a common conspecific facial recognition template exists for all primate species. Although both monkeys and humans showed good recognition of primate faces, subjects in these earlier studies received substantial training with human and monkey faces before being tested with new stimuli in a recognition test. It is thus possible that practice or experience with a large number of primate faces may render the subject experts in recognizing primate faces in a more general sense. There exists evidence in the literature that suggests that practice and experience might have an important influence on our ability to discriminate faces. One line of evidence is given by the ‘other-race’ effect for face recognition in humans, showing the difficulty that human subjects have in recognizing faces from a different ethnic group (O’Toole et al., 1994). Another line is provided by the study of face recognition in infancy (de Schonen and Mathivet, 1989). There it was proposed that face recognition processes in human infants develop as a result of prolonged exposure to this class of stimuli. Therefore, face recognition skills gained after extensive practice with face stimuli in discrimination or recognition tasks may provide the subject with extensive experience with different primate species and may also encourage the subject to use a specific strategy to solve the task that can be different than the strategy usually used to process the same stimuli in a more ‘naturalistic’ environment (Ridley and Baker, 1991). To assess whether the common ‘primate template’ found with discrimination and recognition tasks (Keating and Keating, 1993; Morris and Hopkins, 1993; Phelps and Roberts, 1994; Wright and Roberts, 1996), might also be found in tasks not
requiring a problem-solving strategy, we tested monkeys’ and humans’ ability to recognize monkey faces, human faces, and objects using a visual paired-comparison task (VPC). The VPC task has predominantly been used to assess recognition memory functions in infancy (Fagan, 1973; Pascalis and de Schonen, 1994). It is an habituation paradigm which relies on a prominent preference to explore novel stimuli in the environment. In a sense, one could argue that the recognition process measured with the VPC task is comparable to the recognition mechanism that humans use in everyday life when exploring stimuli from their surroundings. Because this task has the advantage of not overtraining the participants with a class of stimuli before recognition memory is measured, as was the case with problem solving tasks (discrimination and match-to-sample tasks), the presence of similar primate face recognition in humans and monkeys in that task should indicate similar neural processing for faces in the two species and provide some additional information on whether practice or experience with faces of a species might affect face recognition process.

2. Materials and methods

2.1. Subjects

2.1.1. Human participants

Participants were 12 healthy Caucasian research fellows (six males and six females) from the Department of Microbiology of the University of Texas (Houston, TX). These participants had never worked with monkeys before and had no special expertise in monkey face recognition. They ranged in age from 22 to 35 years (mean = 28.45, S.D. = 4.82). Participants were only informed that they would participate in a visual test.

2.1.2. Monkey subjects

Subjects were four adult rhesus monkeys (Macaca mulatta), three males and one female, aged 10–12 years (mean = 10.67, S.D. = 0.81). These animals had been raised with humans, including animal caretakers, veterinarians, and researchers. Although these monkeys had been exposed to human faces, the number of human faces they saw during their life was limited. They were housed in individual cages in large animal rooms accommodating 20–30 monkeys, and received daily periods of social interactions with other peers. Thus, they had been exposed to a large number of conspecifics prior to the present experiment. In addition, while these animals had served as normal controls in other learning and memory experiments, none of them had been trained in tasks using faces as stimuli.

2.2. Stimuli

The stimuli consisted of 35-mm black and white slides of unknown faces and variegated, nonsense objects, photographed in front of a uniform background and made with a scanner. Black and white slides were used instead of colored ones to allow a better control of brightness and contrast between stimuli. To avoid contamination of the results with the effects of recognizing faces from other races or other species, slides of faces consisted of Caucasian humans or rhesus macaques. All faces were represented without the neck. In addition, for human faces, individuals posed without jewelry, glasses, or intense make-up, and with the most neutral expression possible. Examples of each category of stimuli are displayed in Fig. 1. The size of the pictures as well as their brightness were kept uniform on each slide and, when projected onto the screen, each picture was 15 cm high and 10 cm wide. Only one stimulus was projected in the center of the screen for the familiarization period, and two stimuli were projected side by side separated by a 12-cm interval during the retention tests.

For human participants, four different pairs of stimuli were used in each condition, e.g. four object-pairs, four monkey-face pairs, and four human-face pairs (two pairs representing male subjects and two pairs representing female subjects). Each human participant received one trial at each of the three conditions and, for each trial, a stimulus-pair was selected from the pool of four stimulus-pairs and randomly assigned to each participant. For the monkeys, 15 pairs of objects, 15 pairs of monkey faces, and 15 pairs of human
faces were used, and the 15 stimulus-pairs of each condition included the four stimulus-pairs used with humans. For each condition, monkeys received 10 trials made of 10 stimulus-pairs drawn from the pool of 15 pairs and randomly assigned to each subject.

2.3. Behavioral procedure

A visual paired comparison task (VPC) was used to assess recognition memory in both humans and monkeys. In this task, the participant was first exposed to a visual stimulus and allowed to passively explore it during a familiarization period. For both human participants and monkeys a long familiarization period was used to ensure complete sensory processing of the sample. After a brief delay, during which the participant’s view of the stimulus was prevented, the participant was confronted with the familiarized stimulus presented together with a new stimulus, for two successive retention periods in which the position of the stimuli on the screen was reversed to minimize left/right looking preference. Recognition memory was inferred from the participant’s tendency to fixate significantly longer at the novel stimulus during the retention periods.

2.3.1. Human participants

Human participants sat approximately 60 cm in front of a screen onto which slides were back projected. For each trial, a sample stimulus was presented in the center of the screen for a cumulative familiarization time of 10 s and then removed. After a 10-s retention period, the familiar stimulus (FS) and a new stimulus (NS) were projected side by side for two consecutive 5-s retention tests. During the two retention tests, the left/right positions of the stimuli were reversed and counterbalanced across participants. Following a 30-s period, another trial began in the same way but with a different pair of stimuli. For all participants, the stimulus conditions were presented in the same order: object, monkey face, human face.

2.3.2. Monkeys

A transport cage made of transparent plastic was used to test the animals. The cage was placed into a Wisconsin General Testing Apparatus (WGTA) with its front side facing a screen (30 cm away) onto which the stimuli were back projected. At the center of the front side of the cage a sipper tube was attached through which orange juice was delivered during training to keep the animal facing and looking at the screen while drinking.

An adaptation period of 4–5 days was necessary to acclimate the monkey to sit quietly in the front of the cage and to drink from the sipper tube while looking at stimuli. Thereafter, formal training began as in humans. That is, during the familiarization period the FS was presented in the middle of the screen and remained on until the monkey had looked at it for a total cumulative time of 20 s. (The 10 s difference between the familiarization time in humans and monkeys was dictated by our observation that human participants, who did not receive any incentive for con-

Fig. 1. Example of stimulus-pairs for each category of stimuli.
continuously looking at the stimuli, became rapidly bored; therefore, to reduce distractibility the familiarization time was slightly reduced). Again, after a retention period of 10 s, the FS and an NS were simultaneously projected onto the screen for two consecutive tests in which the left-right position of the stimuli were reversed and counterbalanced across trials. Following a 30-s period, another trial began in the same way but using a different pair of stimuli. Ten trials were used for each stimulus condition and were randomized during a daily session.

For both humans and monkeys, a video camera with a videotimer (to the nearest 1/100 s) was fixed above the screen to record the participant’s eye movements onto a TV monitor and a videotape during the entire testing session. Stimulus fixation was assessed by corneal reflection of the stimuli. An observer hidden from the subject’s view examined the subject’s eyes on a control TV monitor and controlled the time for the different stimuli presentation during the familiarization and retention tests, for the retention and intertrial intervals. The familiarization period began when a subject first looked at the stimulus and ended when a subject had looked at the stimulus during either 10 s for human participants or 20 s for monkeys. The familiarization time was calculated on-line during the experimentation by using the display of the time-code generator on the TV monitor as a chronometer. The two retention tests started when the subject looked at one of the two stimuli and ended after 5 s had elapsed in both humans and monkeys.

Following testing, the videotapes were played back and, for each trial, the time spent by participants looking at FS and NS stimuli during retention tests was recorded using a frame by frame video recorder, and percent looking time towards the NS was calculated for each trial. Samples of videotapes were analyzed by two observers and the interobserver reliability was 90%. During videotapes analysis, the observers were not informed of the position (right or left) of the new stimulus on each slide.

3. Results

For both humans and monkeys, the percent looking towards the NS in each category of stimuli is depicted in Fig. 2.

3.1. Human participants

During the two 5-s retention tests, the average looking time at both FS and NS stimuli was 4.7 s (S.D. = 0.31) for objects, 4.7 s (S.D. = 0.32) for monkey faces, and 4.7 s (S.D. = 0.3) for human faces, and these values did not differ from each other.

To test the effects of stimulus category on recognition memory (Fig. 2), a one-way ANOVA with Condition as repeated measures was performed on the percent looking time towards the NS. The analysis revealed a significant effect of Condition \[ F(\text{Huynh-Feldt df 1.3,4}) = 43.44, P = 0.002 \]. Post-hoc tests (t-test with a Bonferroni adjusted probability of 0.05/3 = 0.0167) indicated that percent looking time at novel human faces and novel objects did not differ from each other, but both differed significantly from percent looking time at novel monkey faces (paired two-tailed t-test: \( t = 3.34, \text{df} = 11, P < 0.01 \) and \( t = 3.08, \text{df} = 11, P < 0.01 \), respectively). Thus, for both object and human face stimuli, the average look-
ing time towards NS was significantly longer than that towards FS (paired two-tailed t-test: \( t = 6.44 \), df = 11, \( P < 0.0001 \) and \( t = 2.7 \), df = 11, \( P < 0.02 \) respectively). By contrast, for monkey face stimuli, participants looked at the NS and the FS for similar amounts of time (paired two-tailed t-test: \( t = 0.8 \), df = 11, \( P > 0.05 \)).

3.2. Monkeys

During the two 5-s retention tests, the average looking time at both FS and NS was 3.0 s (S.D. = 1.1) for objects, 3.0 s (S.D. = 1.1) for monkey faces, and 3.3 s (S.D. = 1.2) for human faces. A two-way ANOVA with Condition and Trial as factors with repeated measures indicated a significant difference for Condition (\( F = 15.68 \), df = 2,9, \( P < 0.004 \)) but not for Trial (\( F = 0.658 \), df = 9,81, \( P = 0.74 \)). The interaction between the two factors was not significant (\( F = 0.322 \), df = 18,81, \( P = 0.99 \)). Post hoc tests (Bonferroni adjusted probability \( P = 0.0167 \)) revealed that although percent looking times for novel monkey faces and novel objects did not differ (Fig. 2), they both significantly differed from that for novel human faces (paired two-tailed t-test: \( t = 3.8 \), df = 39, \( P < 0.0005 \) and \( t = 3.8 \), df = 39, \( P < 0.0005 \), respectively). Thus, for both object and monkey face stimuli, the mean looking time towards NS was significantly longer than towards FS (paired two-tailed t-test: \( t = 6.27 \), df = 39, \( P < 0.0001 \) and \( t = 5.38 \), df = 39, \( P < 0.0001 \), respectively). By contrast, for human face stimuli, monkeys looked equally at the NS and FS (paired two-tailed t-test: \( t = 0.9 \), df = 39, \( P = 0.37 \)).

To establish statistical comparison between performance of humans and monkeys on the three conditions, the average percent looking towards novel stimuli across the 10 trials at each condition was calculated for each monkey. These averages were used for comparisons with the human data. A two-way ANOVA with repeated measures for the factor Condition was performed and revealed a significant interaction between species and conditions (\( F(Huynh-Feldt \ df 1.521) = 5.85, \ P = 0.015 \). Tukey HSD multiple comparisons revealed that although percent looking time to novel stimuli did not differ in the two species for the category ‘objects’, the two species significantly differed for the category ‘human faces’ (\( P < 0.01 \), for which human participants but not monkeys showed a preference for novelty, and for the category ‘monkey faces’ (\( P < 0.01 \), for which monkeys but not human participants showed a preference for novelty.

4. Discussion

Using a visual paired comparison task, we first demonstrated a strong preference for looking at novel objects in both adult humans and adult monkeys, suggesting that the VPC procedure provides a useful tool to assess recognition memory in adult primates as well as in infant primates. Surprisingly, however, while human participants showed good novelty preference for human faces, they did not for monkey faces, and the reverse was found for monkeys. That is, monkeys showed good novelty preference for monkey faces but not for human faces. Thus, when tested in a recognition task requiring little practice with face stimuli, humans and monkeys showed a clear ‘species-specific effect’. Before discussing the implication of these findings for our understanding of face recognition processing mechanisms in primates, several methodological issues that may have accounted for the results need to be addressed.

One issue pertains to the type of recognition measured by the VPC task. The VPC task is predominantly an habituation paradigm in which a preference for viewing new stimuli is measured after the subject has been habituated to a familiar stimulus. If a subject looks longer at a novel stimulus, it is inferred that he/she remembers that the other stimulus has already been seen. One question with this paradigm is whether faces can be recognized as faces rather than as objects. Indeed, in both species, subjects showed preference for novelty when using both objects or faces of their own species, and one could argue that the face stimuli could be processed in the same way objects were. If this were the case, no differences should have been found in recognizing faces of the two species. That is, the two species would have showed preference for novelty for faces of
their own species as well as for faces from the other species. Because this was not the case, we conclude that both monkeys and humans have processed the faces of the other species differently from the objects or faces of conspecifics, and, thereby, subjects of both species processed faces differently than they processed objects. One way to investigate further this specific face processing system will be to test the effect of novelty preference for inverted faces in both humans and monkeys.

Another issue pertains to the time of familiarization needed to process faces from the other species. If faces of other species require more time to be processed, the subject might still be involved in that process when the retention trials began, and, instead of looking at the novel face, they might have still been actively processing that first face stimulus and not be attending longer to the novel stimulus. This is unlikely for two reasons. One is simply that face processing is an extremely rapid process requiring only a fraction of a second. Indeed, even young 6-month-old infants show novelty preference for faces with a short 10-s familiarization time (Fagan, 1974). Second, because monkeys have had longer familiarization time than humans (20 s instead of 10 s), they should have shown better recognition of human faces than human participants for monkey faces, but this was not the case. Thus, the difference in the length of the familiarization period in monkeys and humans does not appear to account for the species-specific novelty preference observed.

Finally, one could argue that because the task is based on preferential looking, subjects might have been more interested or motivated to look at conspecifics rather than at members of another species. If participants’ greatest interest or motivation was to look at human faces, why would they have more interest or motivation to look at objects rather than monkey faces? Because novelty preference was observed for objects in both species, it is unlikely that their lack of novelty preference for other species’ faces reflected a lack of interest or motivation.

Therefore, because identical task procedures and stimuli were used in the present study to assess face recognition in humans and monkeys, the better performance observed in the two species in recognizing faces from their own species than from another species provides further evidence of a species-specific effect in the face processing mechanism in primates (Itakura, 1994; Tomonaga et al., 1993). This ‘species-specific effect’ is reminiscent of the ‘other-race effect’ reported earlier in human participants (O‘Toole et al., 1994). The presence of a ‘species-specific effect’ in face recognition in non-human primates is still controversial. While field-born Japanese monkeys preferred photographs of their own species than those of rhesus monkeys (Fujita, 1987), laboratory raised Japanese and rhesus monkeys showed preference for photographs of rhesus monkeys over those of Japanese monkeys (Fujita, 1990; Tomonaga, 1994). This latter finding suggests that macaque monkeys are able to discriminate faces of other macaque species. However, because the experiments have used the same species of monkeys (e.g. all macaques), it is possible that the similarities in the configuration of all macaque faces will enable macaque monkeys to easily recognize faces from different groups of macaque monkeys. Whether or not macaque monkeys could recognize faces from a different species, such as chimpanzees or gorillas, remains to be determined at this point.

The ‘species-specific effect’ observed in the present experiment contrasts sharply with those of earlier studies demonstrating the good ability of both humans and monkeys in recognizing faces of both species (Keating and Keating, 1993; Morris and Hopkins, 1993; Phelps and Roberts, 1994; Wright and Roberts, 1996). One important difference, we believe, between the present study and the previous one relates to the behavioral task employed to assess recognition. Thus, the VPC task in the present report does not require extensive exposure or practice with faces from other species, before recognition memory can be assessed. By contrast, problem solving tasks (e.g. visual discrimination or matching-to-sample) used in earlier reports necessitate extensive training and experience with face stimuli of both species (e.g. humans and monkeys), before recognition memory is measured. Therefore, this difference in the behavioral tasks indicates that practice or experi-
ence with a class of stimuli, such as faces from another species, might have a strong influence on the ability to recognize faces. The effect of exposure or practice on recognition ability has already been shown in human dog trainers who were experts in dogs’ configuration (Diamond and Carey, 1986). Thus, these human experts remembered better upright than inverted dogs’ pictures, whereas human participants without such practice showed poor performance in recognizing either upright or inverted dogs’ pictures.

This effect of practice or experience on face recognition has already been demonstrated for the development of face recognition in infancy. de Schonen and Mathivet (1989) and Johnson and Morton (1991) proposed that during ontogeny cortical circuits develop their specialization for faces as a result of prolonged exposure to that class of stimuli. Hence, an infant primate maturing with only one species of primates in his/her environment will develop a face-recognition mechanism specific to analyzing face configurations of that species and not of another species. This proposal is strengthened by the findings of Chance et al. (1982) indicating that the ‘other-race effect’ is not seen in young children since they showed good recognition of both ‘own-race’ and ‘other-race’ faces. Because this effect appears only with adult participants, it presumably develops with greater experience or exposure during development with faces of one’s own race than with faces of other races (Valentine, 1991). The effect of practice and experience in face recognition processes fits particularly well with recent electrophysiological studies in monkeys (for review see Desimone, 1991) indicating that face-selective cells are distributed over several temporal cortical areas that have different physiological properties and anatomical connections. Thus, these face-selective cells constitute a distributed network contributing to the recognition of specific individuals, to the analysis of emotional expression, and to the construction of face configurations from different orientations. The establishment of such a distributed system for the coding of faces is likely to be formed by practice or experience.

However, the role of exposure or experience may not be the only factor affecting cross-species recognition. Indeed, as shown in the present results, although the difficulty in recognizing monkey faces by human participants who had received no previous exposure to monkeys could be explained by a lack of practice or experience, this is not the case for the monkeys which were raised by humans since birth. These monkeys had been exposed to many exemplars of human faces during their life and had learned about human faces from their exposure to caregivers and experimenters. According to Johnson and Morton (1991) and de Schonen and Mathivet (1989), the monkeys of the present study should be experts in recognizing monkey faces as well as human faces; but the present findings indicate otherwise. Thus, a definite conclusion on the role of experience or training on face recognition cannot be reached at this point and awaits further studies aimed at investigating face recognition in both humans and monkeys that will have received previous, well-controlled, exposure with both human and monkey faces.

In conclusion, humans and monkeys showed species-specific face recognition, even though both species displayed good recognition memory for objects. This pattern of results implies that if a common face-recognition mechanism does indeed exist in all primate species, to be effective for cross-species face recognition this mechanism may require prolonged exposure or training for the establishment of species-specific face configuration. Moreover, because other mammalian species, like sheep, have neural circuits in the temporal cortex which respond preferentially to faces as opposed to other visual stimuli (Kendrick, 1991, 1994; Kendrick and Baldwin, 1987), it will be important to explore the similarity or difference in face recognition in primates and other mammals.

Acknowledgements

This work was supported in part by grant MH-49728 from the National Institute of Mental Health and John D. and Catherine T. MacArthur Foundation to J.B., and by a postdoctoral fellowship from the Fondation Fyssen to O.P. We
would like to thank Ruth Campbell, Michael Coleman, and Anthony Wright for helpful comments on a previous version of this manuscript.

References


