

学位申請論文要旨

ダーウィンの著書「ヒトおよび動物の表情」以来、表情は、ヒトを含めた霊長類のコミュニケーション行動の進化を考える際に、最も重要なものであると考えられる。様々な霊長類の中で、ニホンザルを中心とするマカカ属に関しては、表情の記載的な研究がある程度詳細に行われてきた。しかし、ニホンザルの表情の研究の多くは、野外における記載的なものであり、実験的な研究はほとんど行われてこなかった。この原因として、表情写真などの自然刺激を、よく制御された条件で実験に用いることが困難であったことがあげられる。しかし、近年の画像処理技術の発達により、より自然な表情写真を刺激とする弁別実験が可能となってきた。本研究は、表情の認知をニホンザルとヒトで比較し、その違いを明らかにした研究である。

論文1において、ニホンザルによるニホンザルの表情認知は「優位－劣位」の次元と「覚醒の次元」で記述できることが示された。またこの両次元が、「口の突き出し具合」と「眉の上昇」の2つの動きと関連していることを明らかにした。ニホンザルによるヒトの表情の認知については、笑顔がひとつのまとまりをつくるが、泣顔と怒り顔は区別されないことが示された。論文2においては、論文1において特徴的な認知が見られた笑顔と泣き顔を取りあげ、画像処理技術を用いて特定の部分を操作した表情写真を作成し、ニホンザルがヒトの笑顔と泣き顔を認知するとき、顔のどの部分が重要な役割を担っているかを、ヒトを被験者とした場合と比較し検討した。その結果、笑顔の認知に関しては、ヒトもサルも頬の部分が重要であること、また悲しみ顔に関しては、ヒトにおいて重要な眉の認知が、サルでは存在しないことが示唆された。以上の結果は、ヒトとサルの、表情認知に関する特徴を示すものである。

Recognition of Facial Expressions in a Japanese Monkey
(*Macaca fuscata*) and Humans (*Homo sapiens*)

SO KANAZAWA
Kyoto University

ABSTRACT. Recognition of facial expressions by a Japanese monkey and two humans was studied. The monkey subject matched 20 photographs of monkey facial expressions and 20 photographs of human facial expressions. Humans sorted the same pictures. Matching accuracy by the monkey was

HOW DO JAPANESE MONKEYS (*Macaca fuscata*) SEE SMILING AND SAD FACES OF HUMANS?

KANAZAWA, So

Department of Behavioral and Brain Sciences, Primate Research
Institute, Kyoto University

Running Head: How Japanese monkeys see ^{smiling} smile and sad faces

author's address: Department of Behavioral and Brain Sciences,
Primate Research Institute, Kyoto University, Kanrin,
Inuyama City, Aichi Prefecture, 484 Japan

phone: 81-568-63-0567

fax: 81-568-63-0085

e-mail: kanazawa@smtp.pri.kyoto-u.ac.jp

abstract

Four Japanese monkeys (*Macaca fuscata*) and 17 humans performed an odd-item visual search task of a variety of photos of human facial expressions. The target was either a real smiling face or a real sad face of a particular female. The distracters were the following artificial images produced by a computerized image-processing system: 1) a neutral face made by averaging the two targets, 2) faces with smiling or sad eyebrows on the neutral face, 3) faces with smiling or sad eyes on the neutral face, 4) faces with a smiling or sad cheek on the neutral face, 5) faces with a smiling or sad mouth on the neutral face. The search reaction time of both species was the longest when they had to find the smiling target among the distracters having the smiling cheeks. This suggested that the cheek region was the most important cue to recognize smiling faces for both monkey and humans. In searching for the sad target, however, while the reaction time of humans was the longest when the distracters had sad eyebrows or sad cheeks, that of monkeys was so when the distracters had sad cheeks. Namely, the most important cue in recognizing sad faces was the eyebrows and the cheek for humans, but for monkeys it was the cheek regions. These results indicate that the monkeys recognize human ^{smiling} smile faces in the same manner as humans do, but they do not use eyebrows as a cue to recognize human sad faces.

Key Words: facial expression, Japanese macaque, odd-item visual search, image-processing system, eyebrows.

Introduction

The face is undoubtedly one of the most important media for social communication in primate species, including humans. A face yields sources of various information, such as age, sex, identity, social rank, emotional state and so on. In particular, emotional communication strongly depends on the ability to recognize facial expressions. Without facial expressions we can hardly maintain social relationships among social members.

Facial expressions are produced by various movement of facial parts. For example, human sad faces are produced by downward movement of the corners of the mouth and upward movement of the inner portions of the eyebrows. Ekman and Friesen (1975) proposed the concept of "action units" as sets of fundamental movement of the facial parts that constitute human facial expressions. They suggested that the action units may be common to all the cultures. This naturally implies a genetic background in human recognition of facial expressions.

Given such a genetic background, one way to understand characteristics of human recognition of facial expressions ought to be to know how nonhuman primates recognize them. Actually, however, though several researchers have studied how monkeys recognize faces themselves, very few have studied how they recognize facial expressions.

In face recognition psychologists have shown "inversion effects" in monkeys (Rosenfeld & van Hoesen, 1979; Keating &

Keating, 1993; Phelps & Roberts, 1994 ;Tomonaga, 1994; but negative effect: Bruce, 1982). Keating and Keating (1982,1993) and Kyes and Candland (1987) showed that the eyes provide the most salient cue when monkeys recognize human faces.

Neurophysiologists have shown neural correlates of monkey facial recognition (Perrett et al., 1982, 1990; Yamane et al., 1988; Hasselmo et al., 1989). For example, Hasselmo et. al (1989) showed that same neurons in the temporal cortex responded differentially to facial expressions of monkeys and identity of monkey faces. Yamane et al. (1988) showed the major configuration of facial parts which activate the neurons in the inferotemporal cortex.

Although these previous data revealed some aspects of monkey face recognition, they lacked the data on monkeys' recognition of facial expressions comparable with those of humans'. How do monkeys see their faces when they recognize facial expressions ? Which facial parts are important in face recognition ? Are there any differences between monkeys and humans in the way how they recognize the facial expressions ?

In order to answer these questions, I conducted an experiment on the recognition of facial expressions by a Japanese monkey (Kanazawa, 1996). The experiment showed that while the monkey recognized human smile faces just like humans, he did not discriminate sad faces from angry faces. This result suggested one aspect in recognizing human facial expressions that was different between the two species. In the experiment that follows, I analyzed which physical features are important in

monkey recognition of human facial expressions. I used photos of artificially modified faces made by an image-processing system.

Method

Subjects

Subjects were four Japanese monkeys (*Macaca fuscata*) and seventeen human adults (*Homo sapiens*). The monkey subjects were one 7-year-old male (subject Q), one 6-year-old female (subject H), and two 8-year-old female (subjects J and T) in the beginning of the experiment. All monkey subjects were born in the laboratory and reared by humans with either Japanese monkeys or rhesus monkeys as cage mates. They received minimum food deprivation to maintain their performance, and their body weights never dropped below 95 % of their free-feeding weights. They had experience of operant lever pressing reinforced by presentation of pictures of monkeys. The care and use of the subjects adhered to the "Guide for the Care and Use of Laboratory Primates" (1986) of the Primate Research Institute, Kyoto University. All the human subjects were 18-year-old female undergraduate students.

Apparatus

Monkey subjects were tested in an experimental chamber, 70cm × 70cm × 70cm in size. A 14-inch CRT monitor (27cm wide and 20cm high) (KV-14MD1, SONY, Tokyo, Japan) was mounted on one wall of the chamber. A touch sensor (HYPER TOUCH, Nissha-Inter-Systems, Kyoto, Japan) covered the monitor to detect responses onto the monitor. A response

lever was located below the monitor. A feeder (S-100, Tosoku, Tokyo, JAPAN) placed on the ceiling of the chamber could deliver pieces of food into the food cup, located at the floor of the chamber. An electronic chime provided a signal for reinforcement. There was a houselight at the top of the chamber. A personal computer (PC-286VS, EPSON, Tokyo, Japan) presented pictures of faces on the monitor and controlled the whole equipment.

Another personal computer (PC-386GS, EPSON, Tokyo, Japan) was used for the experiment with humans, and the responses were recorded on the keyboard.

Stimuli

I prepared 11 photos of human facial expressions. Two of them were those of smiling and sad faces, acted by an adult female (top two in Figure 1). These will be referred to as the *natural-smile face* and the *natural-sad face*, respectively. The third one was a neutral face, made by averaging the two natural faces above using an image-processing system (Yamaguchi et al., 1995) (the photo at the center of Figure 1). This will be called the *averaged face*. The remaining 8 photos were neutral faces with one of the facial parts replaced by those of the *natural-smile face* or the *natural-sad face* (bottom 8 photos on both sides in Figure 1). Four facial parts were chosen: eyebrows, eyes, cheek, and mouth. These artificial images were made by the same image-processing system. They will be called, the *eyebrows-smile face*, the *eyes-smile face*, the *cheek-smile face*, the *mouth-smile face*, the *eyebrows-sad face*, the *eyes-sad face*, the *cheek-sad face* and

the mouth-sad face, respectively.

Insert figure 1 about here

Procedure

Monkeys. The task was an odd-item search task (Blough, 1989). Trial proceeded as follows. The lever was illuminated after intertrial intervals of 3 sec. The monkey was required to press the lever down to start up the trial. If the monkey held the lever down, four stimuli simultaneously appeared at the four quadrants on the monitor after a random interval between 1 and 3 sec.

There were one target stimulus and three homogeneous distracter stimuli on the display. The target stimuli were either the natural-smile face or the natural-sad face throughout the experiment. The distracters varied from stage to stage.

A single touch onto the target stimulus was reinforced by a piece of sweet potato. Touches onto the distracters followed by a 3-sec. timeout during which the house light was turned off. If the monkey released the lever before touching one of the stimuli, the trial was canceled and the timeout was given immediately.

In training sessions, the distracters were either monkey faces, chimpanzee faces, or human faces. For two monkeys (subject H and T) the target was the natural-smile face and for the other two monkeys (subject J and Q) it was the natural-sad

face. Test sessions started when the performance of a training session were above 90 % correct in accuracy.

In the test sessions, the two monkeys who had searched for the natural-smile face during training (subject H and T) had to search for the natural-smile face once again. The distracters were four kinds of partly smile faces (i.e., eyebrows-smile face, eyes-smile face, etc.) and the averaged face. For the two monkeys who had searched for the natural-sad face during training (subject J and Q) searched for the natural-sad face. The distracters were four kinds of partly sad faces (i.e., eyebrows-sad face, etc.) and the averaged face. Sessions consisted of 200 trials during which each of the five distracters appeared randomly at the same frequency. Four sessions were run, during which each distracter was tested 160 times.

After the first test sessions above, each subject were trained again with the other natural faces as a target. Namely, subjects H and T now searched for the natural-sad face, and subjects J and Q searched the natural-smile face. Another test series started when the performance of the training session were above 90 % correct. This second test series were conducted in exactly the same way as the first sessions. With this procedure, we expect that the subjects reaction time will be longer when they have to choose the natural-smile (or sad) face from the distracters sharing important facial parts with the target.

Humans. Human subjects performed the same search task with the keyboard. The trial number appeared in the center of the monitor. When the subject pressed the space bar, stimuli

appeared on the monitor. The stimuli were four faces consisting of one target stimulus and three distracter stimuli. Subjects were verbally instructed to press one of four keys which corresponded to the location of the stimuli on the monitor. Seven human subjects searched for the natural-smile face from the partly smile face and the averaged face and the other 10 subjects searched for the natural-sad face from the partly sad faces and the averaged face. Each subjects received one test session consisting of 200 trials, during which each distracters was tested 40 times.

Results

Monkeys

The numbers of training and test sessions for each subject are summarized in table 1. All the monkeys performed consistently at the accuracy above 90% correct during the test sessions.

Insert table 1 about here

I analyzed the reaction time of correct responses. First, I picked up the median reaction time for each of the five distracters for each session. Then these medians for each distracter were averaged for the 4 test sessions. This was done for the two test series separately.

Insert figure 2 about here

By this analysis, I obtained 10 representative reaction time for each monkey subject. In order to examine the effect of the facial parts in smile and sad face recognition, I subtracted the reaction time for the *averaged face* from those for each of the other distracters. In this analysis, positive values mean that the monkey's response was delayed when particular facial parts were common between the target and the distracters. Figure 2 schematically shows how to calculate this increase in the reaction time for the *cheek-smile face*.

The results of this analysis are shown in Figure 3. The horizontal axis is the facial parts and the vertical axis is the increase in the reaction time. Solid lines represent the average across subjects, and dotted lines represent the values for each subject.

Insert figure 3 about here

Figure 3a is the result when monkeys searched for the *natural-smile face* from the 4 kinds of partly smile faces. One-way ANOVA revealed that the effect of facial parts was statistically significant ($F(3, 9) = 7.15, p < .01$). The subject was treated as a randomized block. A post-hoc multiple comparison showed that the value for the smile cheek was higher than the other parts ($LSD = .08, 5\%$ level). Namely, monkeys needed more time to find the *natural-smile face* when the distracter had the smiling cheek than when it had the other facial parts of the same expressions.

Figure 3b is the result when the monkeys searched for the *natural-sad face* from the 4 kinds of partly sad faces. As in the case of *natural-smile face*, the increase in the reaction time tended to be longer when the cheek was sad. The same one-way ANOVA suggested that this tendency was almost significant ($F(3, 9) = 3.41, p = .06$). However, there appears to be considerable individual differences among monkeys. This tendency was clear only for two of four monkeys (subjects J and H). For the other two subjects, the reaction time little changed whatever the distracters were.

Humans

Throughout the experiment, the performances of all the human subjects were consistently above 95 %. The analysis as was done for the monkeys conducted for the human subjects. Figure 4a is the results when the target stimulus is the *natural-smile face*. The same one-way ANOVA revealed that the effect of facial parts was statistically significant ($F(3, 18) = 12.90, p < .01$). A post-hoc multiple comparison suggested that the value for the cheek was higher than the others (LSD = 0.1, 5 % level). Just like monkeys, human subjects also needed more time to find the *natural-smile face* among distracters having the smiling cheek than among those having the other parts of the same expression.

Insert figure 4 about here

Figure 4b shows the result of the *natural-sad face*. The same one-way ANOVA revealed that the effect of facial parts was statistically significant ($F(3, 27) = 3.17, p < .05$). A post-hoc multiple comparison showed that the values for the cheek and eyebrows were higher than those for the other two parts ($LSD=0.11, 5\%$ level). Human subjects needed more time to find the *natural-sad face* when distracters had either sad eyebrows or sad cheeks than they had the other sad parts.

Discussion

In recognizing a smile face, both monkey and human subjects needed more time to find the target when it shared the smiling cheek with the distracters. This result suggests that for both monkeys and humans the cheeks are most important facial element in recognition of the smile face.

In recognizing a sad face, human subjects needed more time to find the target when it shared either sad eyebrows or sad cheeks with the distracters. On the other hand, eyebrows had no effect to increase the reaction time of the monkeys. Instead, at least for the two of the subjects, the cheek regions did increase their reaction time. This means that human subjects both the eyebrows and the cheeks are important facial elements when they recognize a sad face.

Figure 5 is to compare the results of monkey subjects with those of human subjects in smile face search and sad face search. As is clear from Figure 5a, in smile face recognition the reaction time of both monkeys and humans commonly increased when the

distracters was the *cheek-smile face*. On the contrary, figure 5b shows that sad face recognition is different between the two species. In particular, there is a marked difference in the effect of eyebrows. The effect of cheek may be somewhat weaker in monkeys than in humans, but this difference was much smaller than what we see in the effect of eyebrows. Clearly, the largest difference between monkeys and humans lies in the effect of eyebrows in sad face search.

Insert figure 5 about here

This difference may be understandable if we note that monkeys have no distinct eyebrows. It is probable that monkeys may lack a perceptual module (Fodor, 1983) to detect the orientation of eyebrows which gives us significant information about the emotional states of performer. Conversely, monkeys share the same movement of the cheek region with humans. For example, when monkeys exhibit the "bared-teeth-display", the cheek region of the face is moved by the facial musculature called "zigomatics major" (van Hooff, 1967). This musculature is homologous to that humans use to make a smile face. Monkeys and humans may have the same perceptual module to detect the movement of the cheek region.

What is puzzling in the results of the present experiment was a failure to find any effect of eyes which were suggested to be the most important parts in facial recognition (Keating & Keating, 1982, 1993; Kyes & Candland, 1987). However,

recognizing expressions from a face actually needs more than recognizing the face itself; it needs detection of movements of various facial parts. In this sense, recognizing expressions from a face may well be a separate process in facial recognition from recognizing the face itself (Bruce, 1988). In the former, the most important information for monkeys may come from movement of the cheek region rather than the eyes which they surely let move when they express the bared-teeth-display noted above.

Another important question is how they acquire the way to recognize facial expressions. Naturally there are two possibilities: being genetically determined or being learned. The present experiment cannot answer this question. The subject used in this study had been hand-reared and had had much opportunity to learn human facial expressions. While such experience does account for the similarity we found between humans and monkeys in the recognition of smiling faces, however, it does not explain the failure of monkeys to detect the difference in the eyebrows.

In the literature, at least some social recognition has been suggested to be innate. For example, Sackett (1966) showed that isolated-reared monkeys spontaneously came to be scared of the aggressive faces. Fujita (1990; 1993) showed that in rhesus monkeys the visual preference for the subjects' own species was not altered by the social experience with different species, though in Japanese monkeys it was.

Given such variability in the factors determining social recognition in monkeys, we cannot presume if the monkeys' recognition of facial expressions examined here has its genetic

background or not. We have to test monkeys reared in their natural group in the future.

References

- Blough, D. S. (1989). Odd-item search in pigeons: display size and transfer effects. *Journal of Experimental Psychology: Animal Behavior Processes*, 15, 14-22.
- Bruce, C. (1982). Face recognition by monkeys: Absence of an inversion effect. *Neuropsychologia*, 20, 515-521.
- Bruce, V. (1988). *Recognizing faces*. Lawrence Erlbaum Associates Ltd.
- Ekman, P., & Friesen, W. V. (1975). *Unmasking the face*. Prentice-Hall.
- Fodor, J. A. (1983). *The modularity of Mind*. Cambridge, MA: MIT Press.
- Fujita, K. (1990). Species preference by infant macaques with controlled social experience. *International Journal of Primatology*, 11 (6), 553-573.
- Fujita, K. (1993). Development of visual preference for closely related species by infant and juvenile macaques with restricted social experience. *Primates*, 34 (2), 141-150.
- Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioral Brain Research*, 32, 203-218.
- Kanazawa, S. (1996). Recognition of facial expressions in a Japanese monkey (*Macaca fuscata*) and humans (*Homo sapiens*). *Primates*, 37(1), 25-38.
- Keating, C. F., & Keating, E. G. (1982). Visual scan patterns of rhesus monkeys viewing faces. *Perception*, 11, 211-219.

- Keating, C. F., & Keating, E. G. (1993). Monkeys and mug shots: Cues used by rhesus monkeys (*Macaca mulatta*) to recognize a human face. *Journal of Comparative Psychology*, 107 (2), 131-139.
- Kyes, R. C., & Candland, D. K. (1987). Baboon (*Papio hamadryas*) visual preference for regions of the face. *Journal of Comparative Psychology*, 101 (4), 345-348.
- Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, 47, 329-342.
- Perrett, D. I., & Mistlin, A. J. (1990). Perception of facial characteristics by monkeys. In Stebbins, W. C., & Berkley, M. A. (Ed.), *Comparative perception vol. 2 complex signals* (pp. 187-215). Wiley-Interscience Publication.
- Phelps, M. T., & Roberts, W. A. (1994). Memory for pictures of upright and inverted primate faces in humans (*Homo sapiens*), squirrel monkeys (*Saimiri sciureus*) and pigeons (*Columba livia*). 108(2), 114-125.
- Rosenfeld, S. A., & Van Hoesen, G. W. (1979). Face recognition in the rhesus monkey. *Neuropsychologia*, 17, 503-509.
- Sackett, G. P. (1966). Monkeys reared in isolation with pictures as visual input: Evidence for an innate releasing mechanism. *Science*, 154, 1468-1472.
- Tomonaga, M. (1994). How laboratory-raised Japanese monkeys (*Macaca fuscata*) perceive rotated photographs of monkeys: Evidence for an inversion effect in face perception. *Primates*, 35(2), 155-165.

- van Hooff, J. A. R. A. M. (1967). The facial displays of catarrhine monkeys and apes. In Morris, D. (Ed.), *Primate Ethology* (pp. 7-68). London: Weidenfield & Nicolson.
- Yamaguchi, M., Hirukawa, T., & Kanazawa, S. (1995). Judgement of gender through facial parts. *Perception*, 24, 563-575.
- Yamane, S., Kaji, S., & Kawano, K. (1988). What facial features activate face neurons in the inferotemporal cortex of the monkey ? *Experimental Brain Research*, 73, 209-214.

Author note

acknowledgment: I wish to thank Mr. Sumiharu Nagumo, P.R.I., Kyoto Univ., for technical assistance. Especially I thank Dr. Kazuo Fujita, Faculty of Letters, Kyoto Univ., for his guidance and valuable suggestions of experimental design and analysis. Thanks are also due to Dr. Tetsuro Matsuzawa, P.R.I., Kyoto Univ., for his helpful advice.

All correspondence should be sent to

So Kanazawa

Department of Behavioral and Brain Sciences,
Primate Research Institute, Kyoto University,
Kanrin, Inuyama City, Aichi Prefecture, 484,
Japan

phone: 81-568-63-0567

fax: 81-568-63-0085

e-mail: kanazawa@smpt.pri.kyoto-u.ac.jp

Table 1

The number of training and test sessions for each subject.

	subject H	subject T	subject J	subject Q
training 1	11	7	14	18
test1	4	4	4	4
training 2	11	6	9	12
test2	4	4	4	4

Figure captions

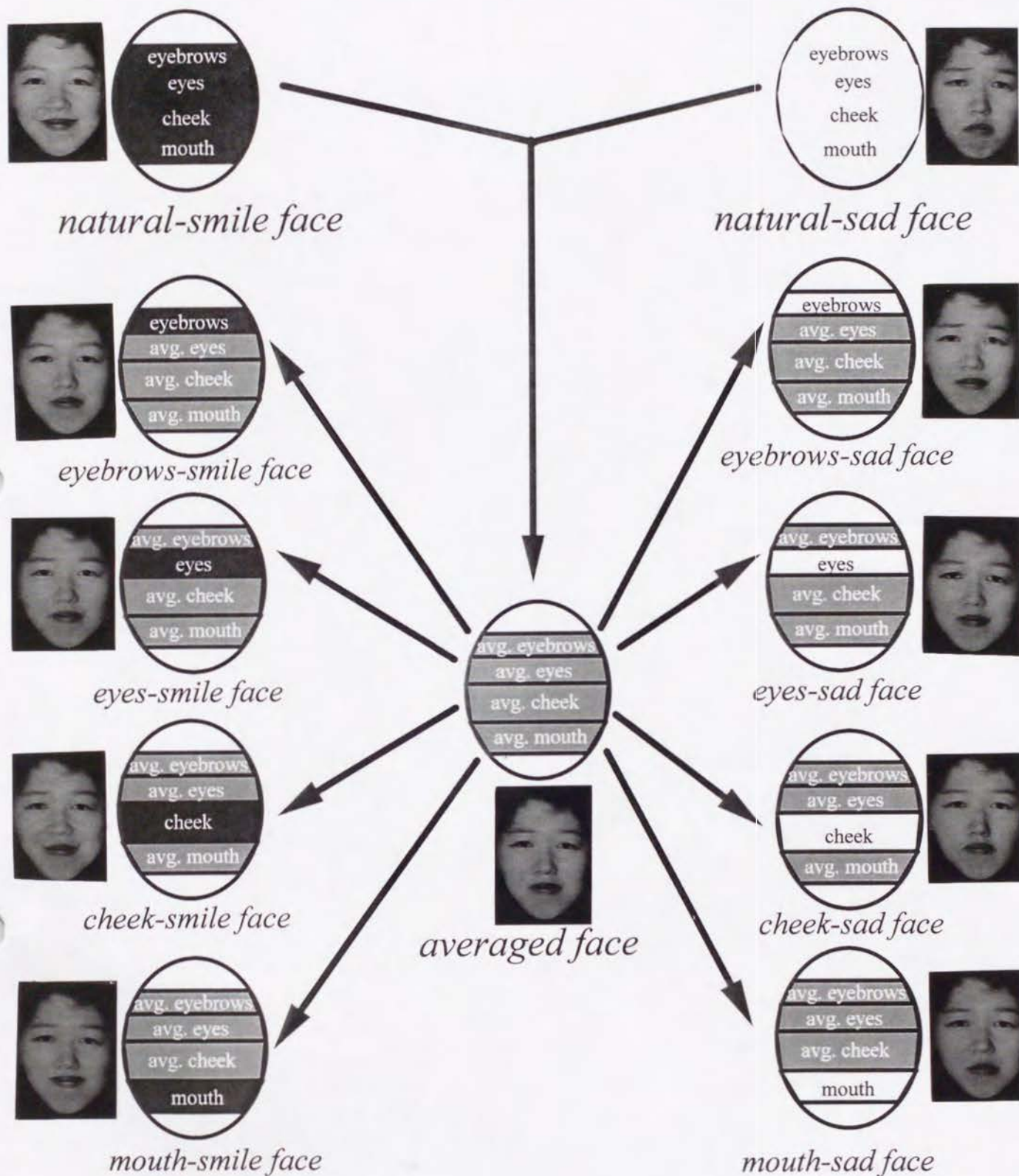
Figure 1. How to make the experimental stimuli. The top left black oval represents the *natural-smile face* and the top right white oval represents the *natural-sad face*. The gray oval at the center represents the *averaged face*. The left four faces, which are arranged in a line, are partly smile faces having only one part of the four parts of the *natural-smile faces* on the *averaged face*. The right four faces are partly sad faces having only one part of the four parts of the *natural-sad faces* on the *averaged face*.

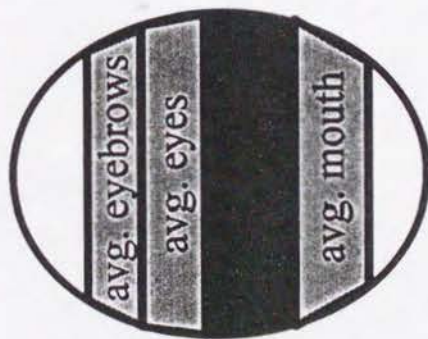
Figure 2. An example of data analysis. See text.

Figure 3. The increase in the reaction time for the monkeys to find the target when the target shared each facial part on the horizontal axis with the distracters. a) when the subjects searched for the human smile face. b) when the subjects searched for human sad face. Dotted lines represent the individual subjects, and solid lines represent the average.

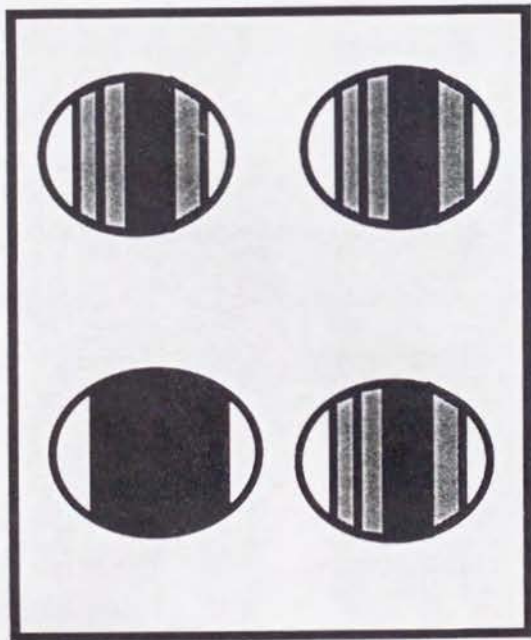
Figure 4. The increase in the reaction time for human subjects to find the target. Others are the same as in Figure 3.

Figure 5. Comparison of the increase in the reaction time between monkeys and humans. Top: the target was the *natural-smile face*, bottom: the target was the *natural-sad face*. Open squares are humans and filled diamonds are monkeys.





RT increase for
cheek-smile face



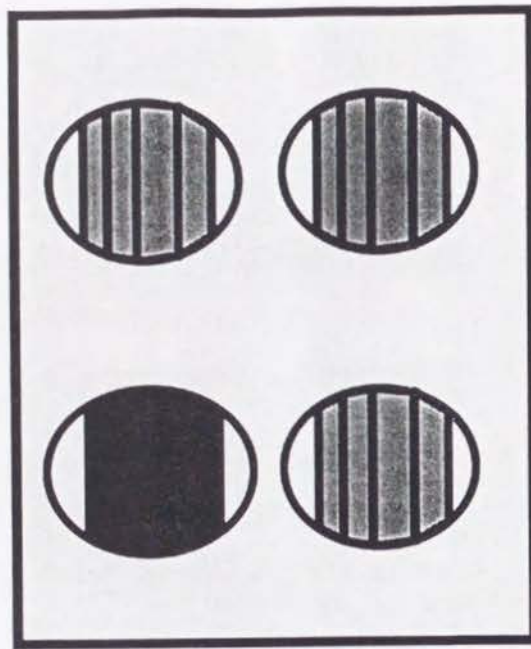
RT for

smile face search

with

cheek-smile face

distracters



RT for

smile face search

with

averaged face

distracters

