Facial expression recognition in crested macaques (*Macaca nigra*)

Jérôme Micheletta[1], Jamie Whitehouse[1], Lisa A. Parr[2], Bridget M. Waller[1]

[1] *Department of Psychology, Centre for Comparative and Evolutionary Psychology, University of Portsmouth, United Kingdom*

[2] *Department of Psychiatry and Behavioral Sciences, Center for Translational Social Neuroscience, Yerkes National Primate Research Center, Emory University, Atlanta, Georgia.*

Corresponding author: Jérôme Micheletta, Department of Psychology, University of Portsmouth, King Henry Building, King Henry I Street, Portsmouth PO1 2DY, United Kingdom.
Email: jerome.micheletta@port.ac.uk; Phone: +44 (0)23 92 84 63 30; Fax: +44 (0)23 92 84 63 00

Abstract:

Facial expressions are a main communication channel used by many different species of primate. Despite this, we know relatively little about how primates discriminate between different facial expressions, and most of what we do know comes from a restricted number of well-studied species. In this study, three crested macaques (*Macaca nigra*) took part in matching-to-sample tasks where they had to discriminate different facial expressions. In a first experiment, the macaques had to match a photograph of a facial expression to another exemplar of the same expression produced by a different individual, against examples of one of three other types of expressions and neutral faces. In a second experiment, they had to match a dynamic video recording of a facial expression to a still photograph of another exemplar of the same facial expression produced by another individual, also against one of four other expressions. The macaques performed above chance in both tasks, identifying expressions as belonging to the same category regardless of individual identity. Using matrix correlations and multidimensional scaling, we analysed the pattern of errors to see whether overall similarity between facial expressions and/or specific morphological features, caused the macaques to confuse facial expressions. Overall similarity, measured with MaqFACS (macaque Facial Action Coding System), did not correlate with performances. Instead, functional similarities between facial expressions could be responsible for the observed pattern of error. These results expand previous findings to a novel primate species and highlight the potential of using video stimuli to investigate the perception and categorisation of visual signals in primates.

Keywords: crested macaques, facial expressions, FACS, matching-to-sample.

Introduction

Facial expression is one of the most common communicative systems used by primates. Facial expressions can convey important social information about the producer's internal state and possibly its forthcoming behaviour (Leopold and Rhodes 2010; Waller and Micheletta 2013). Consequently, the ability to discriminate and interpret these communicative signals is crucial to navigate the complex social environment characteristic of primates.

Few studies have investigated the ability of primates to discriminate between different facial expressions, and only in a limited number of species (reviewed in Parr 2011). Chimpanzees, and some macaques can discriminate facial expressions but their performances are affected by the similarity between matching pairs and the foil expression (Kanazawa 1996; Parr et al 1998; Parr and Heintz 2009). Further studies have shown that some components of the facial expressions were more salient that others (e.g. eyes and mouth when viewing threat faces, Gothard et al 2004). Chimpanzees could even discriminate facial expressions using a single component movement, if this movement was characteristic of the facial expression (Parr et al 2008). Macaques had difficulties when the pictures depicted similar facial expressions produced by different individuals and made more errors when facial expressions shared highly conspicuous features such as the mouth opening and teeth exposure (Parr and Heintz 2009). However in both species, the presence of distinctive features could not account for the overall performances of the subjects, which suggest than chimpanzees and rhesus macaques, as humans, use both featural and configural information to discriminate between facial expressions (Leopold and Rhodes 2010; Parr 2011).

In the current study, critically endangered and highly understudied crested macaques (*M. nigra*) were presented with two facial expression discrimination tasks, one using pictures, the other using videos. Our aims were to explore their ability to discriminate facial expressions and investigate the effect of visual similarities between facial expressions to explain the macaques' performances. Crested macaques' faces are morphologically different from rhesus macaques' (i.e. bony cheek ridges, a shelf-like brow bone and large canine pillars, see ESM for illustration) and they have a highly graded repertoire of facial expressions (Thierry et al 1989; Micheletta et al 2013). The graded and dynamic nature of their facial expressions might mean that they are less accurate than rhesus macaques when matching still representations of conspecific facial expressions. Instead, videos might be a better way to represent the real signals. Consequently, we anticipated an increase in performance in the experiment using video as stimuli.

Material and methods

Subjects

Three adult crested macaques (one male, *Bai*: aged 9 years old and two females, *Sat*: 7 years old and *Dru*: 12 years old) belonging to a social group of five individuals housed at Marwell Zoo (Winchester, United Kingdom: www.marwell.org.uk), took part in the experiment. They all had prior experience with the experimental procedure, having completed similar tasks with pictures of neutral faces of familiar and unfamiliar conspecifics with the same procedure and setup as this study (Micheletta et al. in review). They also had been exposed to videos of conspecifics in a different task prior to this study (Waller unpublished data). Tests took place in an enclosed testing area (approximately 2 x 4 x 5 m) adjacent to their normal enclosure (see ESM). The macaques had unrestricted access in and out of this area. All parts of the enclosure were equipped with vertical and horizontal wooden structures, ropes, trees and various enrichment devices. The macaques were fed daily with commercial monkey pellets, fruits, vegetables, seeds and nuts before and after the experiments. Water was freely available.

Procedures

All the tasks presented here were conducted using a matching-to-sample (MTS) format (see ESM). In the picture experiment, the macaques were first required to orient towards a sample image displayed on a touchscreen (Elo 1939L 19-inch Open-Frame Touchmonitor) by touching it three times in rapid succession. The sample appeared randomly in a central position on the top, bottom, left or right side of the screen. After this initial response, two comparison images were displayed on the screen, on the opposite side of the sample, forming a triangular configuration. The sample remained visible after the appearance of the comparison images (simultaneous MTS). The subject was then required to choose the comparison image that matched the sample (hereafter, the match). If the response was correct, it was followed by an inter-trial interval of 2 s during which the subject received a food reward (pea, sweet-corn or cereal). If the macaque chose the incorrect comparison image (hereafter, the foil), no reward was given and the inter-trial interval lasted 8 s. The video experiment was similar except that a short video clip (4-5 s) was played after the subject touched the sample image. The video ended on a still frame of the facial expression video at its prototypical intensity. Both tasks were programmed with Microsoft Visual Basic; the custom-made program also recorded subject responses. The macaques first completed the picture task before being tested on the video task.

All procedures were approved by both the University of Portsmouth and the Marwell Wildlife Ethics Committees and were in compliance the ASAB/ABS guidelines for the use of animals in research.

Stimuli

The stimuli used in both the picture and video experiments were obtained during a previous study in the Tangkoko Nature Reserve, North Sulawesi, Indonesia (Macaca Nigra Project field site). Pictures were taken with a digital SLR camera. Videos were recorded with a high-definition camcorder. Each stimulus was cropped around the face and resized to a height and width of 300 pixels. All stimuli depicted adult individuals. Both males and females were used, but within any trial, the sample, match and foil represented individuals of the same sex.

In the picture experiment, we used pictures of neutral faces (N) and 4 types of facial expressions (Fig. 1): the bared-teeth display (BT), the half-open mouth threat (OM), the open-mouth bared-teeth threat, or scream face (SC) and the relaxed open-mouth face, or play face (PF). The pictures were combined so that the sample and match showed the same facial expression produced by different individuals, while the foil depicted the other facial expressions and neutral faces (e.g. BT vs. OM, BT vs. SC, BT vs. PF, BT vs. N), making 16 unique stimulus sets (4 for each facial expression). Each stimulus set was repeated twice within a session in a randomized order (32 trials in total). Subjects received these trials until their performances exceeded chance in a session (binomial z-score: > 67.32 % or 22/32 correct responses). Once they completed the picture experiment, they were tested in a second experiment using videos. The video experiment was similar, but the samples were videos of N, BT, OM, SC and yawns (Y) while the match and foil were pictures (Fig. 1). The videos were silent, to ensure that macaques matched stimuli on their visual appearance only. This experiment consisted of 20 unique trials repeated twice within a session in a randomized order (40 trials in total). Subjects received these trials until their performances exceeded chance in a session (binomial z-score: > 65.49 % or 27/40 correct responses). In both tasks, the stimuli displayed adult conspecifics. Within a trial, the sex of the individuals represented in the sample, match and foil was the same but the identity of the individual was different. This ensured that the macaques could not match the stimuli based on sex or identity instead of expression category.

Analysis

For each experiment, we assessed whether similarity between sample and foil expressions affected performances. We estimated similarity using maqFACS (Parr et al 2010), which catalogues the muscle movements underlying facial expressions. Originally, maqFACS was developed for rhesus macaques, whose faces are morphologically different from crested macaques (see ESM), so caution should be taken with these estimates as we might have missed species-specific movements. Nevertheless, the basic FACS system can be used to make broad descriptions across primate species (e.g. Dobson 2009a; Dobson 2009b; Santana et al 2014). Overall appearance similarity between facial expressions was calculated as the Pearson correlation between the mean FACS compositions of each facial expression type. Percentages of matching success for each combination of facial expressions were correlated to their overall appearance similarity using matrix correlations (Hemelrijk 1990a; Hemelrijk 1990b). Matrix correlations were calculated using the MatrixTester software with 5000 permutations. We then used a multidimensional scaling procedure (MDS) to further explore the pattern of errors. The MDS produces a plot based on a confusion matrix and thus represents the perceived dissimilarity between facial expressions. A low percentage of error results in expressions being plotted further apart and therefore indicates a high degree of perceived dissimilarity (Parr and Heintz 2009). The significance of the dimensions (label of the axes of the plots) generated by the MDS is subjective and their meaning can differ between experiments.

Results:

In the first experiment using pictures, *Bai* performed above chance on the eighth session (binomial test, one-tailed, 68.75 %, p = 0.025), *Dru* on the second session (75 %, p = 0.003) and *Sat* on the first session (68.75 %, p = 0.025). Only Bai's performances in the picture task can be examined in more detail because he is the only individual with enough repetitions of the same trials (16) before performing above chance. His performances did not improve significantly with the number of repetitions (r = -0.065, p = 0.949) suggesting that he did not learn the association between stimuli. In the second experiment using videos, *Bai* reached 57.50 % correct (p = 0.215) on the third session and then stopped participating. In contrast, *Dru* performed above chance on the second session (67.5 %, p = 0.019) and *Sat* on the third session (67.5 %, p = 0.019).

Table 1. Matrices showing the average percentage of success and confusion for each expression dyad.

Picture experiment Success	BT	OM	PF	SC
BT		68.18	54.55	63.64
OM	50.00		50.00	63.64
PF	59.09	59.09		72.73
SC	72.73	54.55	50.00	

Picture experiment Confusion	BT	OM	PF	SC
BT		28.00	40.00	32.00
OM	36.67		36.67	26.67
PF	37.50	37.50		25.00
SC	22.22	37.04	40.74	

Video experiment Success	BT	N	OM	SC	Y
BT		50.00	68.75	62.50	56.25
N	62.50		56.25	43.75	81.25
OM	68.75	43.75		87.50	75.00
SC	68.75	43.75	75.00		75.00
Y	31.25	12.50	31.25	31.25	

Video experiment Confusion	BT	N	OM	SC	Y
BT		30.77	19.23	23.08	26.92
N	24.00		28.00	36.00	12.00
OM	25.00	45.00		10.00	20.00
SC	22.73	40.91	18.18		18.18
Y	23.40	29.79	23.40	23.40	

Sample expressions are given in rows and foils in columns (BT: bared-teeth, N: Neutral, OM: half open-mouth threat, SC: open-mouth bared-teeth face or scream face, Y: yawn). Following Parr and Heinz (2009), confusion matrices show the distribution of errors for each foil expression, based on the overall number of errors for each sample (number of errors for dyad A-B divided by total number of error for expression A). In the picture experiment, neutral faces were excluded to obtain a square matrix (they were only used as foils).

Detailed performances for each facial expression dyad can be seen in table 1. Only some of the facial expressions were used as stimuli in both experiments (BT, OM and SC). Although we cannot test it statistically due to small overall numbers, the macaques seem to perform better with videos than with pictures when having to discriminate BT against OM (pictures: 59.1 %; videos: 68.8 %) and OM against SC (pictures: 59.1 %; videos: 81.3 %) but not BT against SC (pictures: 68.18 %; videos: 65.62 %).

In both experiments, the percentage of success did not correlate with overall similarity (Pictures: $\tau_{Kr} = 0.367$, $p = 0.333$, Videos: $\tau_{Kr} = 0.460$, $p = 0.162$), meaning that the macaques' ability to discriminate between facial expressions, presented as pictures, or videos, did not vary according to the appearance similarity between the facial expressions. These results held true even when the analyses were conducted separately for each individual (results not presented here). Results of the MDS can be visualised in figure 1 and the confusion matrices used in the MDS in table 1. Overall, the models prove to be a good fit for the data (Pictures: Young's S-Stress = 0.001, Kruskal's stress I = 0.017, Dispersion accounted for = 0.999; Videos: Young's S-Stress I = 0.081, Kruskal's stress I = 0.145, Dispersion accounted for = 0.979). In the picture experiment, one cluster clearly emerges: BT seems to be confused with SC more than with the other expressions. Dimension 1 might indicate the degree of mouth opening (open as in SC or closed as in BT) and dimension 2 the degree of teeth exposure (exposed as in BT and SC, or covered as in OM). In the video experiment, SC and OM seemed to be confused with N more than with the other expressions. This time, dimension 2 seems to match the degree of teeth exposure. Dimension 1 is more difficult to interpret but could correspond to the context in which the expressions are used rather that their visual appearance (SC is used in response to others' behaviour, while OM and BT are usually used to initiate interactions).

Discussion:

Overall, our findings demonstrate that crested macaques can discriminate between different facial expressions. As with rhesus macaques, most of our subjects performed above chance after only a few sessions, generalising their understanding of the MTS task to a novel context (facial expression rather than individual recognition, Micheletta et al. in review). Because the stimuli represented facial expressions produced by different unfamiliar individuals of the same sex, the macaques' were not basing their choices on the identity or sex of the individuals.

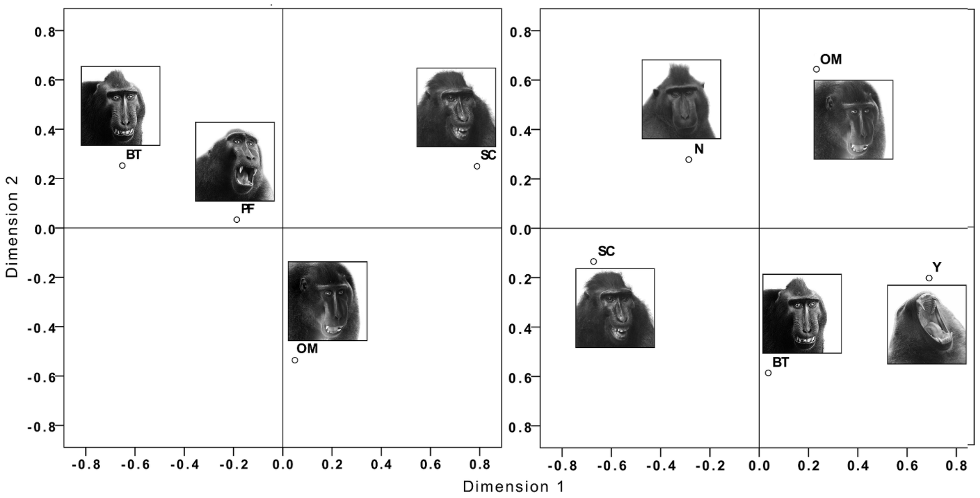

Figure 1. Cluster plot created through multidimensional scaling analyses showing the perceived dissimilarity of facial expressions in both experiments (BT = bared-teeth, N = neutral, OM = half-open mouth, PF = relaxed open-mouth or play-face, SC = open-mouth bared-teeth or scream face, Y = yawn). The left panel is based on performances in the picture experiment. The right panel is based on performances in the video experiment. The further away two facial expressions are, the more dissimilar they were perceived by the macaques according to their performances.

Some facial expressions were clearly more difficult to discriminate than others. This did not seem to be the result of overall similarity of appearance between the different stimuli. As indicated by the matrix correlations, some expressions were easily differentiated even though they had a high number of features in common (e.g. in BT and SC). Either certain features of the expressions caused the macaques to make errors (e.g. mouth opening or teeth exposure), or facial expressions were matched according to their meaning rather than their appearance. In the picture experiment, for example, BT and PF seemed particularly problematic for crested macaques. Both expressions share a characteristic and prominent component movement: teeth exposure (caused by Action Units 9+10 and 12 in maqFACS). In addition to morphological similarity, observational evidence revealed that in tolerant macaques species such as crested macaques, these two facial expressions often blend and merge one into another (Thierry et al 1989; Preuschoft 1995). The convergence between BT and the PF in these species has been proposed to result from the loss of the need to signal subordination, which freed BT from its purely submissive role in favour of a more affiliative one and therefore led to a convergence with the PF (Thierry et al. 1989). These morphological and functional similarities likely make these expressions difficult to discriminate, especially when presented with still images. Interestingly, when tested with videos, the macaques performed poorly when they had to match Y against other expressions. This is quite surprising since Ys are morphologically very distinct from all other expressions. The pattern of error (i.e. equal distribution across all foils) indicates that Ys were not mistaken for other expressions but instead were actively avoided. It is possible that the macaques did not see Y as a facial expression and that this interfered with the MTS rule, or that they avoided Y as it can signal social anxiety (Maestripieri et al 1992). The latter is rather unlikely since the macaques did not systematically avoid threatening facial expressions such as OM and SC.

Crested macaques' performances seemed to improve when using dynamic video displays although we could not test this statistically. Performances on the trials featuring OM appeared easier in with videos than pictures. In reality, facial expressions are highly dynamic, blended and graded signals (Waller and Micheletta 2013). It is possible that the use of videos in MTS tasks represents an improvement in terms of ecological validity. However, the increase in performance could also just be the result of experience, because the macaques all started with pictures before being exposed to videos. Nevertheless, future studies could take advantage of this methodology. Furthermore, the development of a FACS taking into account the morphological specificities of crested macaques will provide an objective and precise tool to further explore the categorisation of facial expressions as well as their function.

Acknowledgements

This work was funded by the Leakey Foundation: Research Grant to B.M.W. and J.M. We thank the staff from Marwell Wildlife: Tim Woodfine, Heidi Mitchell, Will Justice, Shelly Parkes, Kevin Saunders, Andy Double, John Pullen, and all the keepers of the East section for granting us access to the animals and for their assistance and enthusiasm throughout this study. We thank all members of the Macaca Nigra Project (www.macaca-nigra.org) for their support.

References:

Dobson SD (2009a) Socioecological correlates of facial mobility in nonhuman anthropoids. Am J Phys Anthropol 139:413–420.

Dobson SD (2009b) Allometry of facial mobility in anthropoid primates: Implications for the evolution of facial expression. Am J Phys Anthropol 138:70–81.

Gothard K, Erickson C, Amaral D (2004) How do rhesus monkeys (Macaca mulatta) scan faces in a visual paired comparison task? Anim Cogn 7:25–36.

Hemelrijk CK (1990a) A matrix partial correlation test used in investigations of reciprocity and other social interaction patterns at group level. J Theor Biol 143:405–420.

Hemelrijk CK (1990b) Models of, and tests for, reciprocity, unidirectionality and other social interaction patterns at a group level. Anim Behav 39:1013–1029.

Kanazawa S (1996) Recognition of facial expressions in a Japanese monkey (Macaca fuscata) and humans (Homo sapiens). Primates 37:25–38.

Leopold DA, Rhodes G (2010) A comparative view of face perception. J Comp Psychol 124:233–251. doi: 10.1037/a0019460

Maestripieri D, Schino G, Aureli F, Troisi A (1992) A modest proposal: displacement activities as an indicator of emotions in primates. Anim Behav 44:967–979. doi: 10.1016/S0003-3472(05)80592-5

Micheletta J, Engelhardt A, Matthews L, et al (2013) Multicomponent and multimodal lipsmacking in crested macaques (Macaca nigra). Am J Primatol 75:763–773. doi: 10.1002/ajp.22105

Parr LA (2011) The evolution of face processing in primates. Philos Trans R Soc B Biol Sci 366:1764–1777. doi: 10.1098/rstb.2010.0358

Parr LA, Heintz M (2009) Facial expression recognition in rhesus monkeys, Macaca mulatta. Anim Behav 77:1507–1513.

Parr LA, Hopkins WD, De Waal FBM (1998) The Perception of Facial Expressions By Chimpanzees, Pan Troglodytes. Evol Commun 2:1–23. doi: 10.1075/eoc.2.1.02par

Parr LA, Waller BM, Burrows AM, et al (2010) Brief communication: MaqFACS: A muscle-based facial movement coding system for the rhesus macaque. Am J Phys Anthropol 143:625–630. doi: 10.1002/ajpa.21401

Parr LA, Waller BM, Heintz M (2008) Facial expression categorization by chimpanzees using standardized stimuli. Emotion 8:216–231.

Preuschoft S (1995) “Laughter” and “smiling” in macaques: an evolutionary perspective. Faculteit Biologie, University of Utrecht, Utrecht

Santana SE, Dobson SD, Diogo R (2014) Plain faces are more expressive: comparative study of facial colour, mobility and musculature in primates. Biol Lett 10:20140275. doi: 10.1098/rsbl.2014.0275

Thierry B, Demaria C, Preuschoft S, Desportes C (1989) Structural convergence between silent bared-teeth display and relaxed open-mouth display in the Tonkean macaque (Macaca tonkeana). Folia Primatol (Basel) 52:178–184.

Waller BM, Micheletta J (2013) Facial expression in nonhuman animals. Emot Rev 5:54–59. doi: 10.1177/1754073912451503

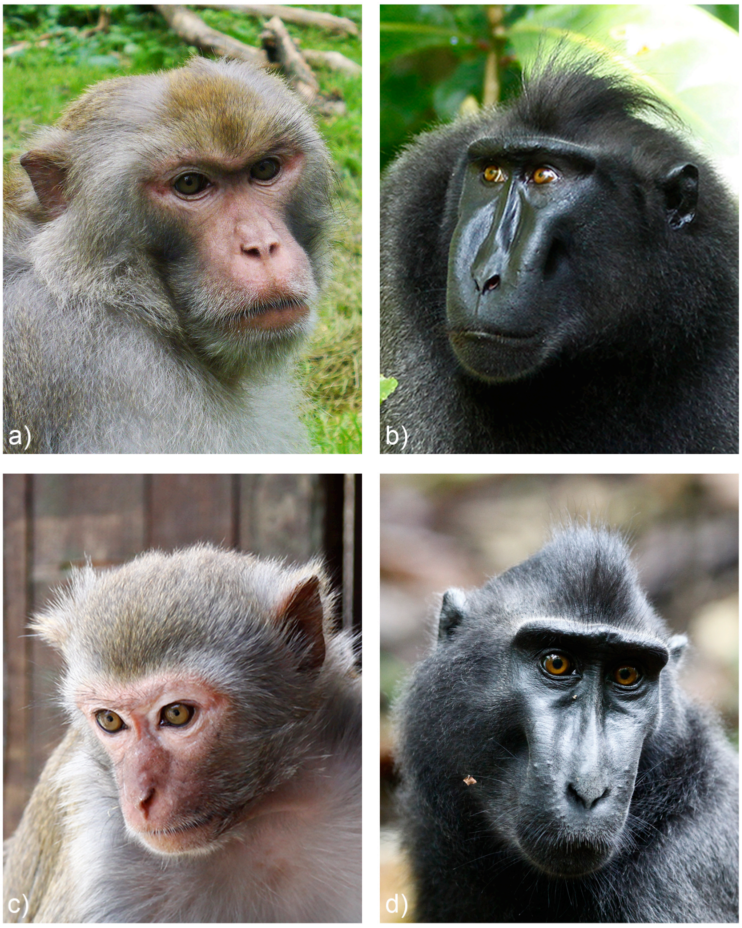

Fig S1. Morphological differences between the face of rhesus and crested macaques: (a) Adult male rhesus macaque, (b) adult male crested macaque, (c) adult female rhesus macaque, (d) adult female crested macaque.

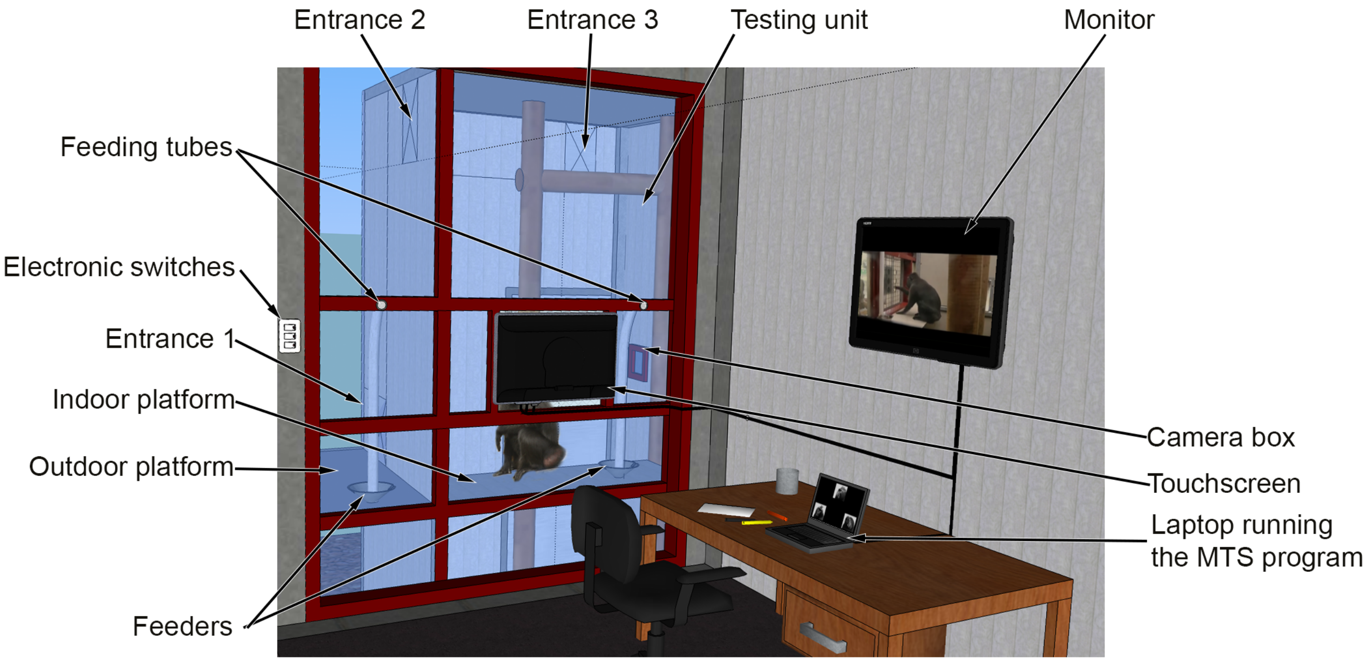

Fig S2. Overview of the Macaque Study Centre. The macaques sit on the outdoor platform until they are granted access to the testing unit. Then, they sit on the indoor platform where they can use the touchscreen. Access in and out of the testing unit is controlled via electronic switches. The match-to-sample task has been programmed using visual basic and runs on the laptop. A camera placed in a protected box is linked to a monitor, allowing the researchers to monitor the behaviour of the macaques. Food rewards are given through plastic tubes, they fall in a bowl within reach of the macaques.

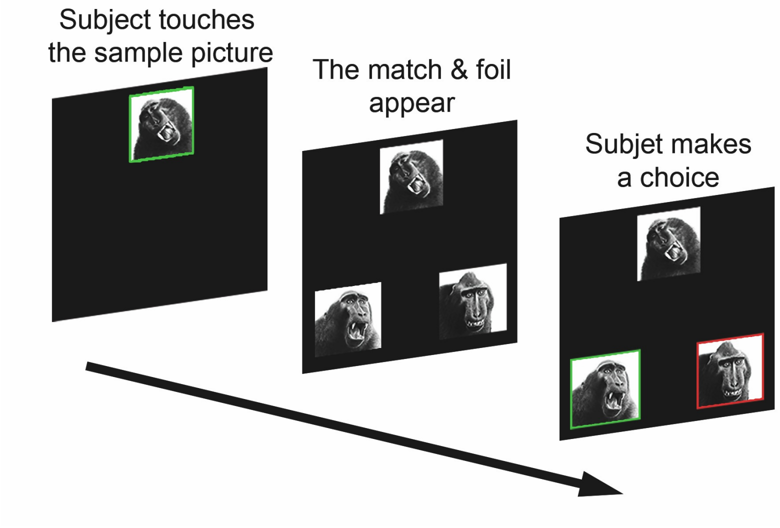

Fig S3. General procedure of the matching-to-sample task. During a trial, the subject touches the sample image 3 times. Then, two additional images appear: the match (which correspond to the sample) and the foil (which does not correspond to the sample). If the subject touches the match, he hears a bell sound and receives a food reward. The inter-trial interval lasts 2 s. If the subject touches the foil, he hears a buzz sound and he does not receive a reward. The inter-trial interval lasts 8 s.

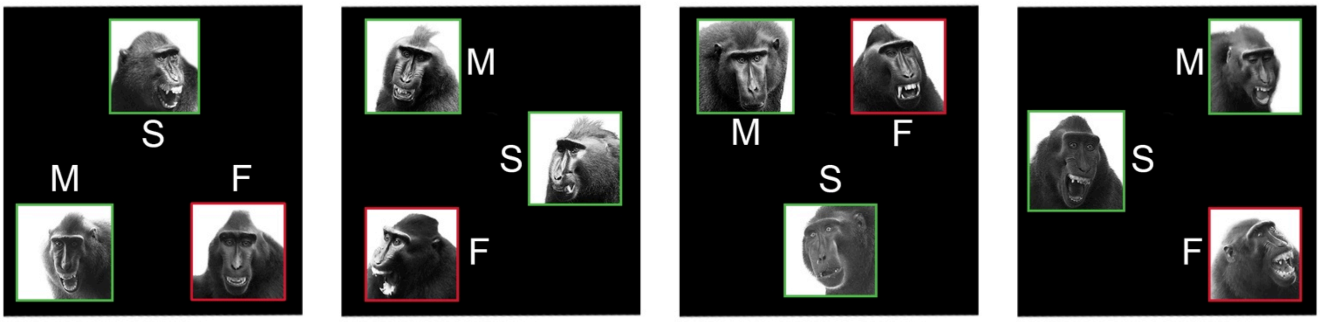

Fig. S4. The possible configurations randomly encountered by the subjects and some possible stimuli (from left to right: SC vs. BT, BT vs. SC, OM vs. PF and PF vs. SC). The sample (S) appears randomly in a central position on the top, bottom, left or right side of the screen. After the macaque's initial response, two comparison images – the match (M) and the foil (F) – are displayed on the screen, on the opposite side of the sample, forming a triangular configuration. The side on which the match and the foil appeared was randomised. Matching stimuli are displayed with a green outline. The foil is displayed with a red outline. Note that neither the colour nor the letters appeared in the real task.