Anim Cogn (2008) 11:485–493
DOI 10.1007/s10071-008-0139-2

ORIGINAL PAPER

Ultra-rapid categorisation in non-human primates

P. Girard · C. Jouffrais · C. H. Kirchner

Received: 8 August 2007 / Revised: 18 December 2007 / Accepted: 17 January 2008 / Published online: 8 February 2008

Abstract The visual system of primates is remarkably efficient for analysing information about objects present in complex natural scenes. Recent work has demonstrated that they perform this at very high speeds. In a choice saccade task, human subjects can initiate a first reliable saccadic eye movement response to a target (the image containing an animal) in only 120 ms after image onset. Such fast responses impose severe time constraints if one considers neuronal responses latencies in high-level ventral areas of the macaque monkey. The question then arises: are non-human primates able to perform the task? Two rhesus macaque monkeys (*Macaca mulatta*) were trained to perform the same forced-choice categorization task as the one used in humans. Both animals performed the task with a high accuracy and generalized to new stimuli that were introduced everyday: accuracy levels were comparable both with new and well-known images (84% vs. 94%). More importantly, reaction times were extremely fast (minimum reaction time 100 ms and median reaction time 152 ms). Given that typical single units onset times in Inferotemporal cortex (IT) are about as long as the shortest behavioural responses measured here, we conclude that visual processing involved in ultra rapid categorisations might be based on rather simple shape cue analysis that can be achieved in areas such as extrastriate cortical area V4. The present paper demonstrates for the first time, that rhesus macaque monkeys (*Macaca mulatta*) are able to match human performance in a forced-choice saccadic categorisation task of animals in natural scenes.

Keywords Reaction time · Saccades · Rhesus macaque monkey · Categorisation · Natural scenes

P. Girard (✉) · C. Jouffrais · C. H. Kirchner
Université de Toulouse, CerCo, UPS, 31062 Toulouse, France
e-mail: pascal.girard@cerco.ups-tlse.fr

P. Girard · C. Jouffrais · C. H. Kirchner
Faculté de Médecine de Rangueil, CNRS, UMR5549,
31062 Toulouse, France

Present Address:
C. Jouffrais
IRIT CNRS, Université Toulouse 3, INPT,
Université Toulouse 1, Université Paul Sabatier,
118 route de Narbonne, 31062 Toulouse cedex 9, France

Introduction

Investigations with macaque monkeys have demonstrated that they can accomplish categorisation tasks with natural images with high levels of performance (Fabre-Thorpe et al. 1998). They can perform ultra-rapid categorisations of targets such as food or animals with high levels of accuracy even if the images are shown in black-and-white and are new to them (Delorme et al. 2000). In such studies, the monkeys responded by touching the screen only if a target object was present in the image. Monkey reaction times in this kind of go/no-go categorisation task are generally very short, and correct responses start to outnumber errors as early as 180 ms. This is much faster than humans who reach minimum reaction times around 250 ms in such a task (Macé et al. 2005).

Have these experiments reached the temporal limits of the primate brain? All the above observations refer to behaviour in a manual go/no-go task. However, it is a well-established fact that eye movements in primates are also initiated with short latencies. Recent work with humans (Evans and Treisman 2005; Grill-Spector and Kanwisher 2005; Codispoti et al. 2006; Kirchner and Thorpe 2006; Serre et al. 2007) has demonstrated that high level processes such as categorisation

of animals in a natural scene can be accomplished very rapidly, either based on behavioural responses or evoked potentials. In the choice saccade task, two photographs of natural scenes are flashed for a short time period left and right of fixation while human participants are asked to make a saccade to the side with the target object. In these conditions, accuracy is around 90%, while correct responses start to outnumber errors at latencies as short as 120 ms (Kirchner and Thorpe 2006). It is interesting to note that saccade reaction times in this forced-choice categorisation task are almost twice as fast as manual responses. This is all the more remarkable as the targets varied considerably in size and type of animal. Target representations thus had to be highly generalised for the visual system to capture the essential features needed to permit ultra-rapid categorisations (Thorpe et al. 1996). Johnson and Olshausen (2003) have proposed that the fastest reaction times might be elicited by low-level features that are shared by the animal category and not by the distractor category. Kirchner and Thorpe (2006) made an analysis of 13 low-level image parameters (first-order and second-order) and showed that both performance accuracy and median reaction times were not changed for a subset of targets and distractors having no difference in these parameters. In the forced-choice task, it then looks much easier to decide which one of two images is more likely to contain an animal than to decide whether an image contains an animal or not (standard go/no-go task). Johnson and Olshausen (2003) and Kincses et al. (2006) support the contradicting view that go/no-go should require less time than a two-choice task. However, a direct comparison of accuracy levels between masked go/no-go and forced-choice categorisations further indicated that the sensitivity is similar whatever the task and that both imply a common initial analysis within the first 40 ms after scene onset (Bacon-Mace et al. 2007). Only with longer masking onset delays did performance improve in the go/no-go task relative to forced-choice conditions indicating different levels of categorisation.

Such ultra rapid categorisation imposes severe timing constraints in the visual system that needs to first visually process the target and to elicit a motor saccadic output. In the monkey, the earliest saccades occur at around 70 ms in response to the onset of a single visual cue presented in the near periphery (Boch et al. 1984). At the other end of the chain from perception to action, it has been shown that it takes 20–25 ms from decision time for performing a saccade until the actual contraction of the eye muscles (Schiller and Kendall 2004). Published studies of visual latencies (Nowak and Bullier 1997; Schmolesky et al. 1998; Thorpe and Fabre-Thorpe 2001) show that cells in primary visual cortex start to discharge at 25–30 ms whereas cells in ventral visual areas such as V4 and inferotemporal cortex (IT) only start around 60–80 ms (Thorpe and Fabre-Thorpe 2001), thus leaving little time for saccadic output.

The question then first arises whether a species without language—the rhesus monkey *(Macaca mulatta)*—is able to perform such ultra rapid and demanding forced-choice saccadic categorisation. Indeed, Grill-Spector and Kanwisher (2005) masking experiments show that humans do not need more processing time to name the category of a target than to detect the target, and they suggest an implicit fast semantic processing inherent to the categorisation. Even more crucial is the speed to which monkeys can achieve the task. Based on former work with manual responses (Macé et al. 2005) and the fact that the selectivity of IT and STP cells is built up fast enough to allow categorisation within the first 100 ms of their response (Oram and Perrett 1992; Freedman et al. 2003, 2006), we made the hypothesis that monkeys should be faster than humans in the forced-choice saccade task. As neuronal latencies in different stages—cortical areas—of the monkey brain are known, the estimate of monkey saccadic reaction times allows one to infer what areas may come into play for such an elaborate task. In the present study, we used a paradigm adapted from the one used in humans in order to perform direct comparisons across species in terms of accuracy levels and (minimum) saccadic reaction times. Monkeys perform this task almost as reliably as humans with ultra-rapid saccadic reaction times starting at 100 ms. This result is reliable across a number of different tasks including generalisation across different groups of distractor images as well as generalisation across different groups of target objects. The present study shows that rhesus macaque monkeys are a highly suitable model to study fast high-level categorisation tasks. Further work should determine which stages of visual processing might underlie ultra-rapid forced-choice categorisations.

Methods

Two adult rhesus monkeys (*Macaca mulatta*; M1, a 12-year-old female and M2 an 8-year-old male), weighing 3 and 6 kg respectively, were involved in the experiments. The animals were not naive to behavioural testing and had previous experience with saccade tasks involving simple geometrical forms. They were born in captivity in Europe. Head fixation devices (CRIST®) were implanted under general anaesthesia (ketamine induction 16 mg/kg IM; Saffan®, 15 mg/kg/h IV with rate adjusted if required) and sterile conditions. A pain reliever (Ketofen, 20 mg/kg) and systemic antibiotics were administrated just before surgery.

The monkeys sat in a primate chair with their head fixed. The behavioural task consisted in forced-choice categorizations as represented in Fig. 1. They were first required to fixate a central fixation spot on the screen (Iiyama vision master pro 510, CRT with 75 Hz frame-rate) during a random period of 500–700 ms. This period was followed by

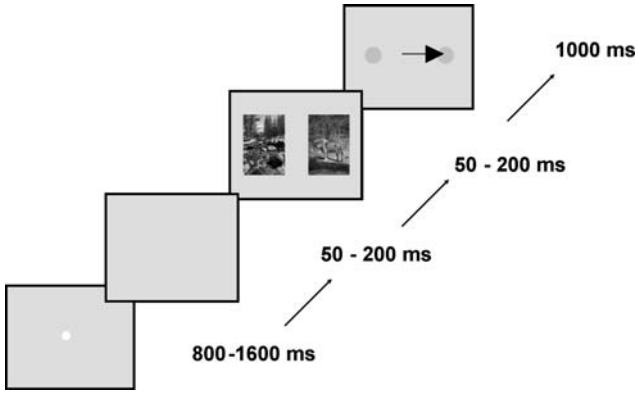

Fig. 1 Forced-choice categorisation task. After a pseudo-random fixation period, a 50–200 ms blank screen (gap period) preceded the simultaneous presentation of two natural scenes in the left and right hemifields (presentation duration: 50–160 ms). The images were followed by two *grey* fixation disks indicating the saccade landing positions

a gap lasting 50–200 ms (to avoid expectancy) with no fixation spot or stimulus on the screen but the animals maintained fixation. Trials were aborted if the animal broke fixation during the gap period. After the gap, while the animals maintained fixation, two stimuli were displayed, one in each lower hemifield (or on the horizontal meridian). The images were always centered at 5° of eccentricity (7° for larger images, see below). The target stimulus always contained an animal, while the other stimulus, the distractor, never contained animals. The side (left/right) of the target presentation was randomized. The monkeys had to make a saccade to the target as soon as the stimuli appeared. Preliminary data in humans let us think that presentation time is an important parameter that could accelerate reaction times. Therefore presentation times of the images vary from session to session (50 or 160 for M1 and 50, 100, or 160 ms for M2), and when the duration of the stimuli was shorter than reaction time, two 0.5° discs were immediately displayed at the center of each image in order to provide an endpoint for the saccade (this was done dynamically, as a function of when the saccade was made on a given trial). Correct saccades were rewarded by a drop of water, incorrect ones by a low tone. Monkeys were weighed before each session and additional water was given at the end of the session if necessary. CORTEX software (courtesy of NIMH) controlled behavior, stimulus presentation and data acquisition. All animal procedures complied with guidelines of the European Ethics committee on Use and Care of Animals.

We used 1,841 pictures (929 animals, 912 distractors). These stimuli were photographs of natural scenes taken from the COREL© database. The target stimuli included a large variety of animals of different classes, orders and species. Animals were photographed in naturalistic scenes in full body view or close up view (including the head). Distractors were made of a large variety of outdoor sceneries containing landscapes and/or manmade items in far or close-up views. Stimuli were displayed on a uniform grey background (luminance 14 cd/m^2). Targets and distractors were 256 grey-levels bitmaps of size 5° × 6° or 10° × 12° (in different sessions).

In a typical daily session, two kinds of stimuli pairs were presented to the monkeys. The first kind consisted of pairs of images where the monkey had neither seen the target nor the distractor before. Such pairs were termed "new pairs". The second kind of pairs consisted in stimuli that had been presented (both the target and the distractor) at least once in a previous session. Such pairs were coined "familiar pairs". In the second half of the experiment, the pairing of familiar targets and distractors was randomly shuffled before each session, so that a given familiar target never appeared with the same familiar distractor in a subsequent session. Results concerning these shuffled pairs were pooled together with non-shuffled ones because, for both monkeys, performance accuracy were very similar and excellent for both kinds of pairs.[1] In each session, 10 new pairs were randomly intermixed with 40 familiar pairs. Each of the 50 pairs could appear between 5 and 20 times in a session. At the end of the experiment, the monkeys had seen many different new pairs (M1: 619, M2: 778).

Finally, a simple control task was designed with only one stimulus appearing on the screen (with the same presentation and gap times as the categorisation task). This task aimed at determining the saccadic reaction times in absence of categorisation. The stimuli were simple patches of textures. This control task was run on days different from those of the main task.

Response recording and detection

Eye position was recorded by an infrared camera (ISCAN, 120 Hz) and stored on a PC. Precise timing of the picture display was assessed with a photodiode. Saccade latencies were computed offline in Matlab as indicated in Kirchner and Thorpe (2006). The derivative (speed) of the eye trace was calculated. A threshold was defined as the maximum speed of the horizontal eye trace over the fixation period (corresponding to the noise of the signal). When the derivative crossed the threshold, after stimulus onset, the abscissa of this point was taken as the reaction time, referred below as RT or latency. We then checked that the eye position signal was not coming down to fixation level for at least the two following points. An example of the computation of the latency for a single trial is represented on Fig. 2.

[1] For M1, accuracy is not different between non-shuffled and shuffled conditions (80.3% vs. 79.4% correct, $\chi^2 = 2.06$, $df = 1$, $P = 0.15$). For M2 the difference is significant but accuracy is high (89.9% vs. 95.4%, $\chi^2 = 392$, $df = 1$, $P < 0.001$).

This procedure proved to be correct when checked visually over a large number of trials. Furthermore, each eye trace leading to a latency inferior to 111 ms was checked visually.

To determine a value for the minimum saccadic reaction time (minimum RT), we divided the saccade latency distribution into 10 ms bins (e.g., the 120 ms bin contained latencies from 115 to 124 ms) and we searched for the first bin that contained significantly more correct than erroneous responses (chi square test, $P < 0.05$). To validate minimum RTs, performances for each of the five consecutive bins (after the minimum RT bin) were required to be above chance.

Results

Performance accuracy

Both monkeys achieved the task with a high level of performance. The overall percentage of correct responses were 79.3 and 90% for M1 and M2, respectively. These values correspond to a large number of saccades (18,472 for M1 and 44,829 for M2).

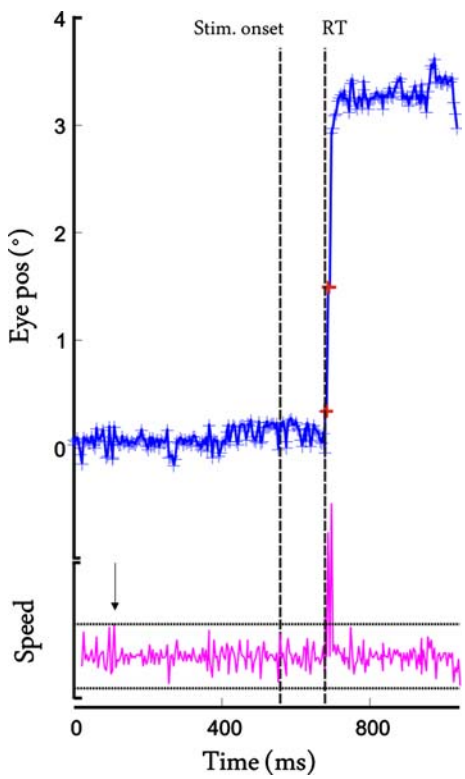

Fig. 2 Example of a single trial with computation of the RT. A threshold (*dotted horizontal lines*) is computed from the maximum (*arrow*) of the derivative (*pink trace*) of the eye trace during fixation. RT is the abscissa at which the derivative crossed the threshold after stimulus onset. We then check that the eye position signal (*blue trace*, 0 ms corresponding to fixation of the central spot) was not coming down to fixation level for at least the two following points (*red crosses*). RT for this trial is 120 ms

One possible pitfall of this kind of study is that animals do not understand the task and rely on rote learning of exemplars. To discard such an explanation, it is necessary to examine the generalisation of the performance to new images. The accuracy on the new pairs was high. If we consider the performance on the first session when the pair was new, the monkeys achieved respectively 73% ($n = 2{,}009$) and 76% ($n = 5{,}906$) of correct responses. This good level of performance relies on several presentations of each new pair in the first session in which it appeared. However, it is well known that macaque monkeys are able to learn and retain a large number of arbitrary stimuli (Ringo et al. 1986; Fagot and Cook 2006) and hence could learn within a few trials the positive stimulus of a given pair although it is intermixed with other pairs. Is then the performance of the animal above chance when a given target is seen for the very first time? Both monkeys achieved 76.6% ($n = 619$) and 77.6% ($n = 778$) of correct responses on the first trial of new pairs, which is clearly above chance level ($\chi^2 = 94$, $P < 0.001$ for M1; $\chi^2 = 128$, $P < 0.001$ for M2, $df = 1$).

The mean percentage of correct responses is significantly higher for the images that became "familiar" (M1 80%, $n = 2{,}009$ for 121 familiar pairs, M2 92% $n = 4{,}520$ for 577 familiar pairs; χ^2 test = 27.51 for M1 and 463 for M2, $df = 1$, $P < 0.001$ for each monkey).

Saccade latencies

Distributions

Figure 3 shows the distributions of saccade latencies for both monkeys. Distributions are split between new and familiar pairs. Distributions of RT were unimodal and sharp (cf. Fig. 3). The overall (new and familiar images pooled together) median RT for correct trials was remarkably fast at 152 ms. Median RTs were 144 ms for M1 (interquartile = 27 ms) and 153 ms for M2 (interquartile = 21 ms). We discarded very few outliers ($n = 6$ below 76 ms and $n = 78$ above 300 ms) out of a total of 63,301 saccadic RTs.

Minimum RT across the whole set of images for both monkeys was 100 ms. Even for the first bin of the RTs distribution (latencies above 94 ms and inferior to 105 ms, $n = 215$), accuracy scores were high (M1 87%, $n = 61$ and M2 69%, $n = 154$, respectively).

RT and familiarity

Familiarity with images led to significantly faster correct RTs (new pairs vs familiar pairs, Mann–Whitney, $U = 1.26 \times 10^8$, n1 = 5,976, n2 = 49,025, $P < 0.0001$) but

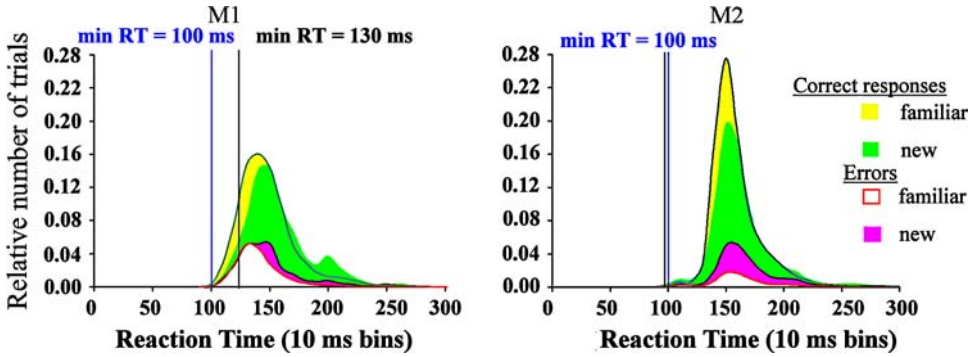

Fig. 3 Distribution of saccadic reaction times for monkey M1 (*left*) and M2 (*right*). The ordinate axis represents the number of trials for each bin divided by the total number of trials for each category of images (familiar or new). Minimum reaction times are represented by *vertical bars* (*blue* for familiar images, *black* for new images). Correct responses are represented in yellow (familiar images) or green (new images). Errors are represented in white (familiar images) or violet (new images).

to only a moderate shift in the median of the RTs distribution. Median correct RT to familiar images are 144 and 153 ms for M1 and M2 respectively whereas median correct RT to new images are 151 and 157 ms for M1 and M2, respectively.

The minimum RT across new pairs (i.e. the trials of the first sessions when these new pairs appeared) was, respectively, 130 and 100 ms for M1 and M2 (100 ms for both monkeys pooled together). A striking effect was seen on the minimum RT of monkey M1, which shortened to 100 ms from new to familiar targets (Fig. 3).

RT and image parameters

Previous work from the laboratory has shown that latencies did not depend on descriptive statistics of low-level features in the images (Kirchner and Thorpe 2006). We computed for each correct saccade the difference in mean luminance or in RMS contrast between the target and the distractor images. Linear regressions between latencies and these low-level features did not indicate any correlation ($R^2 = 0.002$ and 0.0004 for mean luminance and RMS, respectively). Latencies were also not correlated with the luminance and the RMS contrast of the target alone ($R^2 = 0.005$ and 0.0002, respectively). Nevertheless, it could be suggested that an influence of such low level features might only be apparent for the fastest RTs. However, no correlation was found even when we restricted the regression to latencies between 90 and 105 ms, which were in the range of the minimum RT ($R^2 = 0.01$ and 0.004, respectively).

Saccadic reaction times were dependent on the duration of the stimuli. Figure 4 shows that median RTs were faster for both monkeys when image presentation time increased (M1: Mann–Whitney, $U = 78 \times 10^6$, n1 = 3,964, n2 = 10,686, $P < 0.0001$; M2: Kruskall–Wallis, $H = 5{,}769$, n1 = 12,297, n2 = 9,577, n3 = 18,477, $P < 0.0001$).

Comparison with human data

Accuracy for both monkeys was in the range of that of human subjects (see Kirchner and Thorpe 2006 for comparison). The accuracy of the macaques (85 and 69% correct, see above) was in the range of humans (64–100%) for the first bin of the distribution including the minimum RT. However, the overall median RT (152 ms) for monkeys was much lower than human median RT (228 ms, interquartile 77 ms). A Mann Whitney test on saccadic reaction times (correct trials) indicates that monkeys were faster than humans (54,942 correct saccade trials for monkeys and 8,998 for humans, values coming from Kirchner and Thorpe (2006); $U = 5.32 \times 10^7$, $P < 0.0001$); this still holds when we restrict the analysis to a subset of 420 stimuli pairs that have been presented to both humans and macaques (3,001 correct saccade trials for monkeys and 8,998 for humans, $U = 3.89 \times 10^6$, $P < 0.0001$). Only one outlier human subject (#5) in Kirchner and Thorpe (2006) had a median reaction time approaching those of monkeys. If we compare each monkey to this subject, monkeys median RT are not significantly shorter than that of the human subject

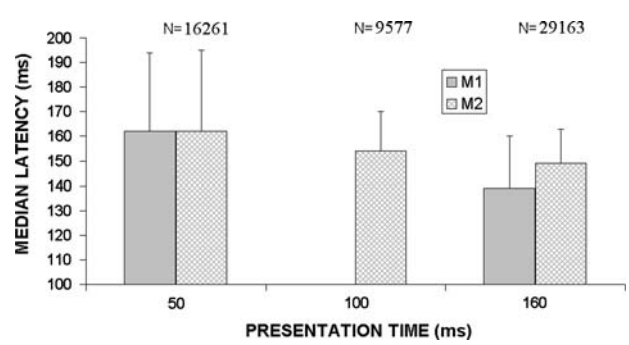

Fig. 4 RTs (median latency in ms) for correct trials in both monkeys (M1 and M2) as a function of presentation time. M1 was tested only with 2 presentation times, see "Methods". *Error bars* are interquartile

(Mann Whitney, M1 versus fastest human: $U = 1.63 \times 10^5$, n1 = 575, n2 = 582, $P = 0.40$; M2 versus fastest human: $U = 6.88 \times 10^5$, n1 = 2,426, n2 = 582, $P = 0.33$).

However, the minimum RTs of monkeys are 30 ms shorter than the minimum RT of the human subject #5. Furthermore, none of the human subjects had a minimum RT below 120 ms. For monkeys, a minimum RT of 100 ms is also found when the analysis is restricted to common images between monkeys and humans on which humans achieved a minimum reaction time of 120 ms (Kirchner and Thorpe 2006), $\chi^2 = 11.94$, $P < 0.001$, $df = 1$. Interestingly, among the 42 images that elicited the fastest reaction times (130–150 ms) in two of the fastest human subjects (Kirchner and Thorpe 2006, their Fig. 6), eight were also found to elicit reaction times shorter than 105 ms in the monkeys (Fig. 5). Twelve other images elicited reaction times below 130 ms and the twenty-two remaining images elicited longer RT in the monkey compared to humans (i.e. above 130 ms). The eight images in the top row of Fig. 5 elicit a high score for both monkeys and humans (each above 88% correct responses in both species). However, there is no overall good match between both species for the preferred images. In the subset of 420 images used in humans and monkeys, 270 and 217 gave rise to scores above 90%, respectively, but 150 reached that percentage in both species.

One could claim that the overall faster reaction times observed in monkeys are a consequence of using longer presentation times than in humans. However, two facts argue against this interpretation. First, the median RTs obtained with a 50-ms duration of presentation (cf. Fig. 4, comparable with presentation times in the study with humans) are much shorter than the respective value in humans (162 for each monkey vs. 228 ms for humans; Mann–Whitney $U = 2.34 \times 10^7$, n1 = 16,261, n2 = 8,998 $P < 0.0001$). Second, the minimum RT obtained with a 50 ms duration of presentation is still 100 ms in the monkey, which is identical to the minimum RT obtained with all saccades (see above).

Finally, we compared RTs in the forced-choice categorisation task with those obtained in the control task. Four hundred and eighty-three control saccades have been recorded in both monkeys as a response to a single texture patch. In this task, monkeys achieved a minimum reaction time of 80 ms. This leads to a 20-ms difference in minimum reaction time between the control task and the forced-choice task. The corresponding difference in human subjects, however, was larger (47.5 ms on average).

Discussion

The present paper demonstrates for the first time, that rhesus macaque monkeys (*Macaca mulatta*) are able to match human performance in a forced-choice saccadic categorisation task of animals in natural scenes. They achieve an overall level of accuracy that is in the range seen with humans. Excellent performances were obtained in both monkeys and did not depend upon the examined low-level image statistics. Both animals quickly generalised to new stimuli. Both animals were born and bred in captivity and had little opportunity to encounter living exemplars of the animal types being displayed. However, it should be noted that, apart from being paired with another rhesus macaque

Fig. 5 Examples of images used in the present study. These images triggered responses with reaction times between 95 and 104 ms (corresponding to the bin with the minimum RT). They were also used in the previous study in man (Kirchner and Thorpe 2006). Images in the upper row elicited short RT in human subjects represented on Fig. 6 of Kirchner and Thorpe (2006)

monkey, both animals had access to daily TV shows as part of an enrichment program. Those TV shows were not controlled or recorded but could contain wildlife shows. Our results complement former work (Yoshikubo 1985; Schrier and Brady 1987; Roberts and Mazmanian 1988; Fabre-Thorpe et al. 1998; Delorme et al. 2000; Macé et al. 2005) that provided convincing evidence that monkeys categorise pictures of animals and generalise to new exemplars. Our data go along these lines since, for both animals, the scores on the first presentation of new targets were excellent. Our results even go one step further since we assessed the fastest categorisation process by using a choice saccade task.

Both minimum and median reaction times of the monkeys were faster than humans' RTs. This is not a simple consequence of choosing young animals. The median lifespan of rhesus macaque monkeys is 25 years (Bodkin et al. 2003). If we extrapolate to humans by a factor of 3, the monkeys of the present study were in the range of age of human subjects in Kirchner and Thorpe (2006). Furthermore, the fastest monkey was comparatively older than the human subjects.

Previous studies (Fabre-Thorpe et al. 2001; Kirchner and Thorpe 2006) showed that human minimum RTs are not reduced by image familiarity. It was argued that humans may already perform at a ceiling speed, presumably because they are already 'overtrained' on detecting animals through exposure to books, films and television. Interestingly, one monkey did show an effect of familiarity that consisted of a shift of reaction times towards smaller values. This observation could give a hint that familiarity processes could build up very rapidly not only in accuracy but also in speed.

Although the median reaction times that we recorded are very fast, we cannot claim that they include express saccades. For both monkeys, distributions are clearly not bimodal. This is reminiscent of the work of Schiller and collaborators that shows that saccadic distributions are not bimodal in natural scenes though short latencies are observed (Schiller et al. 2004). This is also the case in the studies in which fewer express saccades are observed when the target competes with one (Weber and Fischer 1994) or several distractors (McPeek and Schiller 1994, their Fig. 9).

Comparison of manual and saccadic responses

Former work (Fabre-Thorpe et al. 1998; Delorme et al. 2000; Macé et al. 2005) has demonstrated that manual responses (on a touch screen) of macaque monkeys are about two-third shorter than those of humans in the same experimental conditions (minimum RTs were 180 and 290 ms, respectively). When the behavioural responses are saccades, the two-third difference in minimum RTs does not hold anymore if we compare across individuals, since some humans were as fast as 120 ms. However the two-third difference holds if we globally compare monkeys to humans minimum RTs (100 vs. 150 ms). It also holds for the comparison of median RTs of both species. Although manual and saccadic responses were not obtained from the same subjects, the constant proportion between categorisation reaction times suggests that similar processes come into play in both species and tasks (Fig. 6). A second argument reinforcing this conclusion comes from minimum RTs derived from control (simple detection) tasks. The comparison of a simple image detection with a categorisation process—both of them being assessed through saccadic responses—shows an average 47.5 ms difference in humans (Kirchner and Thorpe 2006). Present data show that this difference is 20 ms shorter in monkeys. Very comparable differences are also found with manual responses (data on monkeys:20 ms and humans: 40 ms, in Macé et al. 2005).

Where in the brain does categorisation take place?

These timing differences (delay) between categorisation and detection tasks may reflect the size difference of the human and monkey brains. In that case, the delay may cor-

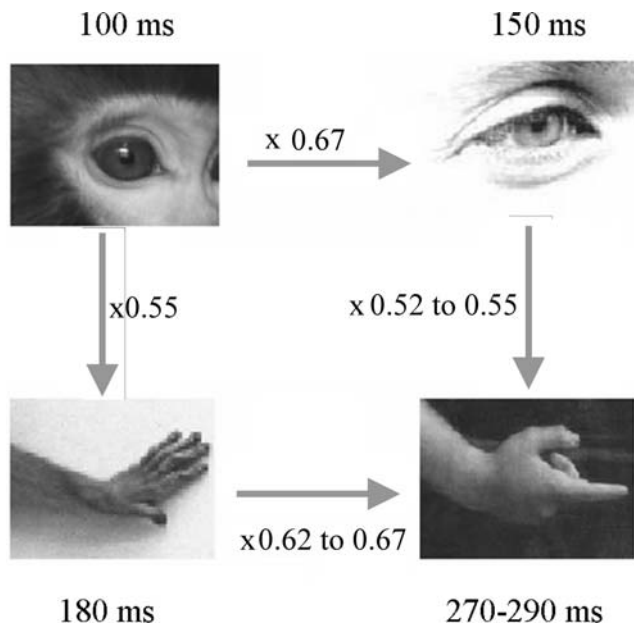

Fig. 6 Constant time factors between minimum RTs of macaque monkey (*left*) and human (*right*) when eye (*top*) or hand (*bottom*) movements are considered. Time factors between species or between effectors are indicated close to the arrows that indicate the computation (for instance 0.67 in the *top row* is the time factor computed by minimum saccadic RT in the monkey/minimum saccadic RT in man). Minimum RTs for forced-choice saccades are mean minimum RTs of subjects. Minimum RTs for manual responses are minimum RTs of the population responses. Manual responses for monkeys are taken from a go/no-go task (Fabre-Thorpe et al. 1998); manual responses for humans are taken from the same go/no-go task (290 ms) or a forced choice manual task (270 ms) (Bacon-Mace et al. 2007)

respond to expansion of cortical territories and/or supplementary areas in the human brain (Orban et al. 2004; Grefkes and Fink 2005; Van Essen 2005) or mean axonal length differences. For instance the approximate proportion of brain areas devoted to vision is two-third between macaques and humans (Orban et al. 2004; Van Essen 2005) and matches the two-third difference in minimum reaction times. Whatever the brain scaling differences between species (and the variations between individuals), both comparisons demonstrate a constant time factor of 0.5–0.55 between saccadic and manual responses (100/180 for monkeys and 150/290 for humans; Fig. 6). Hence the gain in speed conduction of the respective circuits leading to a saccade or a manual response is identical in both species.

It remains to specify which brain areas are involved in such fast and complex processing. Diagnostic elements, i.e. body parts like the eyes, are sufficient for categorisation and may trigger the fastest reaction times (Johnson and Olshausen 2003; Guyonneau et al. 2006; Delorme 2000; Ullman et al. 2002). Where this detailed part-based processing occurs remains unclear. Macaque monkeys constitute a valid model to probe the brain at its fastest highly cognitive abilities (Girard et al. 2002). Reversible inactivation along the ventro-temporal pathway could help to reveal those areas that might be involved in ultra-fast object detection in complex natural scenes.

Acknowledgments Maxime Rosito for eye movement analysis; Franck Lefèvre et Sebastien Aragones for animal care; Michèle Fabre-Thorpe and Simon Thorpe for helpful discussion. All animal procedures complied with guidelines of the European Ethics committee on Use and Care of Animals and complied with French laws.

References

Bacon-Mace N, Kirchner H, Fabre-Thorpe M, Thorpe SJ (2007) Effects of task requirements on rapid natural scene processing: From common sensory encoding to distinct decisional mechanisms. J Exp Psychol Hum Percept Perform 33:1013–1026

Boch R, Fischer B, Ramsperger E (1984) Express-saccades of the monkey: reaction times versus intensity, size, duration, and eccentricity of their targets. Exp Brain Res 55:223–231

Bodkin NL, Alexander TM, Ortmeyer HK, Johnson E, Hansen BC (2003) Mortality and morbidity in laboratory-maintained Rhesus monkeys and effects of long-term dietary restriction. J Gerontol A Biol Sci Med Sci 58:212–219

Codispoti M, Ferrari V, De Cesarei A, Cardinale R (2006) Implicit and explicit categorization of natural scenes. Prog Brain Res 156:53–65

Delorme A (2000) Traitement visuel rapide de scènes naturelles chez le singe, l'homme et la machine: une vision qui va de l'avant ... In. Paul Sabatier - Toulouse 3, Toulouse, p 314

Delorme A, Richard G, Fabre-Thorpe M (2000) Ultra-rapid categorisation of natural scenes does not rely on colour cues: a study in monkeys and humans. Vision Res 40:2187–2200

Evans KK, Treisman A (2005) Perception of objects in natural scenes: is it really attention free? J Exp Psychol Hum Percept Perform 31:1476–1492

Fabre-Thorpe M, Richard G, Thorpe SJ (1998) Rapid categorization of natural images by rhesus monkeys. Neuroreport 9:303–308

Fabre-Thorpe M, Delorme A, Marlot C, Thorpe S (2001) A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. J Cogn Neurosci 13:171–180

Fagot J, Cook RG (2006) Evidence for large long-term memory capacities in baboons and pigeons and its implications for learning and the evolution of cognition. Proc Natl Acad Sci U S A 103:17564–17567

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2003) A comparison of primate prefrontal and iTemporal cortices during visual categorization. J Neurosci 23:5235–5246

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2006) Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. Cereb Cortex 16:1631–1644

Girard P, Lomber SG, Bullier J (2002) Shape discrimination deficits during reversible deactivation of area V4 in the macaque monkey. Cereb Cortex 12:1146–1156

Grefkes C, Fink GR (2005) The functional organization of the intraparietal sulcus in humans and monkeys. J Anat 207:3–17

Grill-Spector K, Kanwisher N (2005) Visual recognition. Psychol Sci 16:152–160

Guyonneau R, Kirchner H, Thorpe SJ (2006) Animals roll around the clock: the rotation invariance of ultrarapid visual processing. J Vis 6:1008–1017

Johnson JS, Olshausen BA (2003) Timecourse of neural signatures of object recognition. J Vis 3:499–512

Kincses ZT, Chadaide Z, Varga ET, Antal A, Paulus W (2006) Task-related temporal and topographical changes of cortical activity during ultra-rapid visual categorization. Brain Res 1112:191–200

Kirchner H, Thorpe SJ (2006) Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. Vision Res 46:1762–1776

Macé MJ, Richard G, Delorme A, Fabre-Thorpe M (2005) Rapid categorization of natural scenes in monkeys: target predictability and processing speed. Neuroreport 16:349–354

McPeek RM, Schiller PH (1994) The effects of visual scene composition on the latency of saccadic eye movements of the rhesus monkey. Vision Res 34:2293–2305

Nowak LG, Bullier J (1997) The timing of information transfer in the visual system. In: Rockland KS, Kaas JH, Peters A (eds) Extrastriate visual cortex in primates, vol 12. Plenum Press, New York, pp 205–241

Oram MW, Perrett DI (1992) Time course of neural responses discriminating different views of the face and head. J Neurophysiol 68:70–84

Orban GA, Van Essen D, Vanduffel W (2004) Comparative mapping of higher visual areas in monkeys and humans. Trends Cogn Sci 8:315–324

Ringo JL, Lewine JD, Doty RW (1986) Comparable performance by man and macaque on memory for pictures. Neuropsychologia 24:711–717

Roberts WA, Mazmanian DS (1988) Concept learning at different levels of abstraction by pigeons, monkeys, and people. J Exp Psychol Anim Behav Process 14:247–260

Schiller PH, Kendall J (2004) Temporal factors in target selection with saccadic eye movements. Exp Brain Res 154:154–159

Schiller PH, Slocum WM, Carvey C, Tolias AS (2004) Are express saccades generated under natural viewing conditions? Eur J Neurosci 20:2467–2473

Schmolesky MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, Leventhal AG (1998) Signal timing across the macaque visual system. J Neurophysiol 79:3272–3278

Schrier AM, Brady PM (1987) Categorization of natural stimuli by monkeys (*Macaca mulatta*): effects of stimulus set size and modification of exemplars. J Exp Psychol Anim Behav Process 13:136–143

Serre T, Oliva A, Poggio T (2007) A feedforward architecture accounts for rapid categorization. Proc Natl Acad Sci U S A 104:6424–6429

Thorpe SJ, Fabre-Thorpe M (2001) Neuroscience. Seeking categories in the brain. Science 291:260–263

Thorpe SJ, Fize D, Marlot C (1996) Speed of processing in the human visual system. Nature 381:520–522

Ullman S, Vidal-Naquet M, Sali E (2002) Visual features of intermediate complexity and their use in classification. Nat Neurosci 5:682–687

Van Essen DC (2005) Surface-Based Comparisons of Macaque and Human Cortical Organization. In: Dehaene S, Duhamel JR, Hauser MD, Rizzolatti G (eds) From monkey brain to human brain. A Fyssen Foundation Symposium, pp 3–20

Weber H, Fischer B (1994) Differential effects of non-target stimuli on the occurrence of express saccades in man. Vision Res 34:1883–1891

Yoshikubo S (1985) Species discrimination and concept formation by rhesus Monkeys. Primates 26:285–299