

# Spotting animals in natural scenes: efficiency of humans and monkeys at very low contrasts

Marc J.-M. Macé · Arnaud Delorme ·  
Ghislaine Richard · Michèle Fabre-Thorpe

Received: 13 March 2009 / Revised: 23 September 2009 / Accepted: 20 October 2009 / Published online: 18 November 2009  
© Springer-Verlag 2009

**Abstract** The ability of monkeys to categorize objects in visual stimuli such as natural scenes might rely on sets of low-level visual cues without any underlying conceptual abilities. Using a go/no-go rapid animal/non-animal categorization task with briefly flashed achromatic natural scenes, we show that both human and monkey performance is very robust to large variations of stimulus luminance and contrast. When mean luminance was increased or decreased by 25–50%, accuracy and speed impairments were small. The largest impairment was found at the highest luminance value with monkeys being mainly impaired in accuracy (drop of 6% correct vs. <1.5% in humans), whereas humans were mainly impaired in reaction time (20 ms increase in median reaction time vs. 4 ms in monkeys). Contrast reductions induced a large deterioration of image definition, but performance was again remarkably robust. Subjects scored well above chance level, even when the contrast was only 12% of the original photographs ( $\approx 81\%$  correct in monkeys;  $\approx 79\%$  correct in humans). Accuracy decreased with contrast reduction but only reached chance level -in both species- for the most extreme condition, when only 3% of the original contrast remained. A progressive reaction time

increase was observed that reached 72 ms in monkeys and 66 ms in humans. These results demonstrate the remarkable robustness of the primate visual system in processing objects in natural scenes with large random variations in luminance and contrast. They illustrate the similarity with which performance is impaired in monkeys and humans with such stimulus manipulations. They finally show that in an animal categorization task, the performance of both monkeys and humans is largely independent of cues relying on global luminance or the fine definition of stimuli.

**Keywords** Natural scenes · Categorization · Macaques · Humans · Low contrast · Fast visual processing

## Introduction

The ability of monkeys to categorize objects in photographs of natural scenes has now been shown for a variety of object categories from subordinate to superordinate levels, and including kingfishers, birds, fish, trees, primates, animals, food-objects, etc. (Roberts and Mazmanian 1988; Fabre-Thorpe et al. 1998; Vogels 1999; Martin-Malivel and Fagot 2001). In the animal kingdom, categorization abilities are not restricted to monkeys. In fact, the first pioneering study on image categorization in animals (Herrnstein and Loveland 1964) showed that pigeons can peck selectively in response to photographs that contain representation of humans. Since then, numerous studies have shown the remarkable abilities of pigeons or birds in general at building object classes. However, remarkable categorization abilities can be displayed without real conceptual abilities. Pigeons have been shown to take into consideration multiple stimulus features for complex categorizations

---

M. J.-M. Macé · A. Delorme · G. Richard · M. Fabre-Thorpe (✉)  
Université de Toulouse, UPS,  
Centre de recherche Cerveau et Cognition,  
133 route de Narbonne, Toulouse, France  
e-mail: michele.fabre-thorpe@cerco.ups-tlse.fr

M. J.-M. Macé · A. Delorme · G. Richard · M. Fabre-Thorpe  
CNRS, CerCo, Toulouse, France

### Present Address:

M. J.-M. Macé  
Université de Toulouse, UPS, CNRS,  
IRIT, 118 route de Narbonne, Toulouse, France

(Von Fersen and Lea 1990; Huber and Lenz 1993). For example, although pigeons are able to categorize female human faces separately from male faces (Troje et al. 1999), they mainly rely on surface information and average luminance intensity of faces. Moreover, in conditions where luminance intensity is normalized, pigeons use color, luminance gradients, and shading (Huber et al. 2000).

Surprisingly, the features used by monkeys performing categorization tasks have not been studied in much detail. Monkeys have been reported to rely on color for categorizing images (D'Amato and Van Sant 1988; Roberts and Mazmanian 1988). However, there is a controversy concerning the role played by color: using an animal/non-animal and a food/non-food fast categorization task, we have shown that removing color cues has little effect on performance in either monkeys or humans. A very mild accuracy drop can be observed in some subjects, and a 10–15-ms reaction time increase is observed with food-objects as targets (Delorme et al. 2000). The reason for this discrepancy might be found in the temporal constraint that we use in our fast go/no-go visual categorization task. In our protocol, photographs of natural scenes were only displayed for 28 ms, and monkeys had 1,000 ms to trigger a go response to a target. Although color has often been reported to be a main cue in scene recognition and memorization (Gegenfurtner and Rieger 2000; Wichmann et al. 2002; Spence et al. 2006), it might only be used in late stages of scene recognition (Delorme et al. 2000; Yao and Einhauser 2008). The minor role played by color in object categorization was also reported by Vogels (1999) in rhesus monkeys categorizing trees and fish. He showed that color, simple form, or texture features could not account for monkeys' performance, but that performance dropped with image scrambling. In the absence of color, monkeys could rely—as pigeons do—on other low-level cues or a combination of low-level features including contrast or/and global luminance. Indeed, like all living beings, monkeys are faced in their lives with natural conditions that are far from optimal for object recognition (dusk, dawn, foggy conditions). The present study measured the performance of monkeys in an animal/non-animal categorization task using natural scenes in which color was removed and that were presented with large variations of luminance and large reductions in contrast.

Another question to consider is the phylogenetic position of monkeys that are closer to humans than they are to birds. Although the evolution of monkeys and humans separated about 30 millions years ago, they might still share a common neural substrate for general conceptual representations that could have been already present in their common ancestor. If so, efficient low-level cues might also be used to some extent by the human visual system. Indeed, the

visual systems of monkeys and humans are very similar. Object recognition involves the ventral visual pathway and is reported to rely mainly upon parvocellular visual information. Thus, categorization performance might be impaired when low contrasts are below parvocellular threshold, although performance appears to be very robust in humans (Macé et al. 2005a).

Even if their visual systems are similar, human and monkeys might not rely on the same (conceptual?) representation of object categories. When humans and monkeys are compared in the same “animal/non-animal” or “food/non-food” fast go/no-go visual categorization tasks, monkeys are slightly less accurate than humans (around 90 vs. 94% correct), but they are much faster. The earliest correct responses are observed at a latency of about 180 ms in monkeys and at 270–280 ms in humans (Fabre-Thorpe et al. 1998; Delorme et al. 2000; Macé et al. 2005a). This difference in minimal input–output processing times might reflect differences in neural processes underlying visual categorization in the two species. Indeed, Martin–Malivel and collaborators, (Martin–Malivel et al. 2006) compared the information used by baboons and humans when categorizing human faces from baboon faces. They showed that, unlike humans, baboons' performance was similar to that of an ideal observer responding solely on the basis of pixel similarities between probe and training images. However, faces are a very special stimulus category. In studies involving pigeons and monkeys, the face stimuli are typically very stereotyped, and they usually occupy most of the picture area. Moreover, in humans, faces might be processed differently than other object categories (Bentin et al. 1996; Farah et al. 1998). The characteristics of faces might not apply to other object classes such as animals for which exemplars can be very dissimilar in shape, size, and texture especially when photographed in their natural environment.

In the present study, we compared monkey and human performance in an animal/non-animal fast categorization task with modified gray-level natural photographs. The task was performed on randomly mixed sequences of photographs at different mean luminance shifts or different contrast reductions resulting in degraded image definition. Such conditions of degraded vision mimic those faced by humans and monkeys under the extreme lighting conditions caused by fog, dazzling sun, or shadowy light, as well as at dusk and dawn. Similar performance impairments in both species would suggest that categorization in humans and monkeys might rely on similar neural mechanisms. Alternatively, if monkeys show stronger impairments than humans in one or both luminance-reduction and contrast-reduction conditions, it might suggest that their categorization performance is more dependent on such image characteristics.

## Experiment 1: variations of mean luminance

### Methods

A group of 18 human subjects and 2 rhesus monkeys were involved in this experiment.

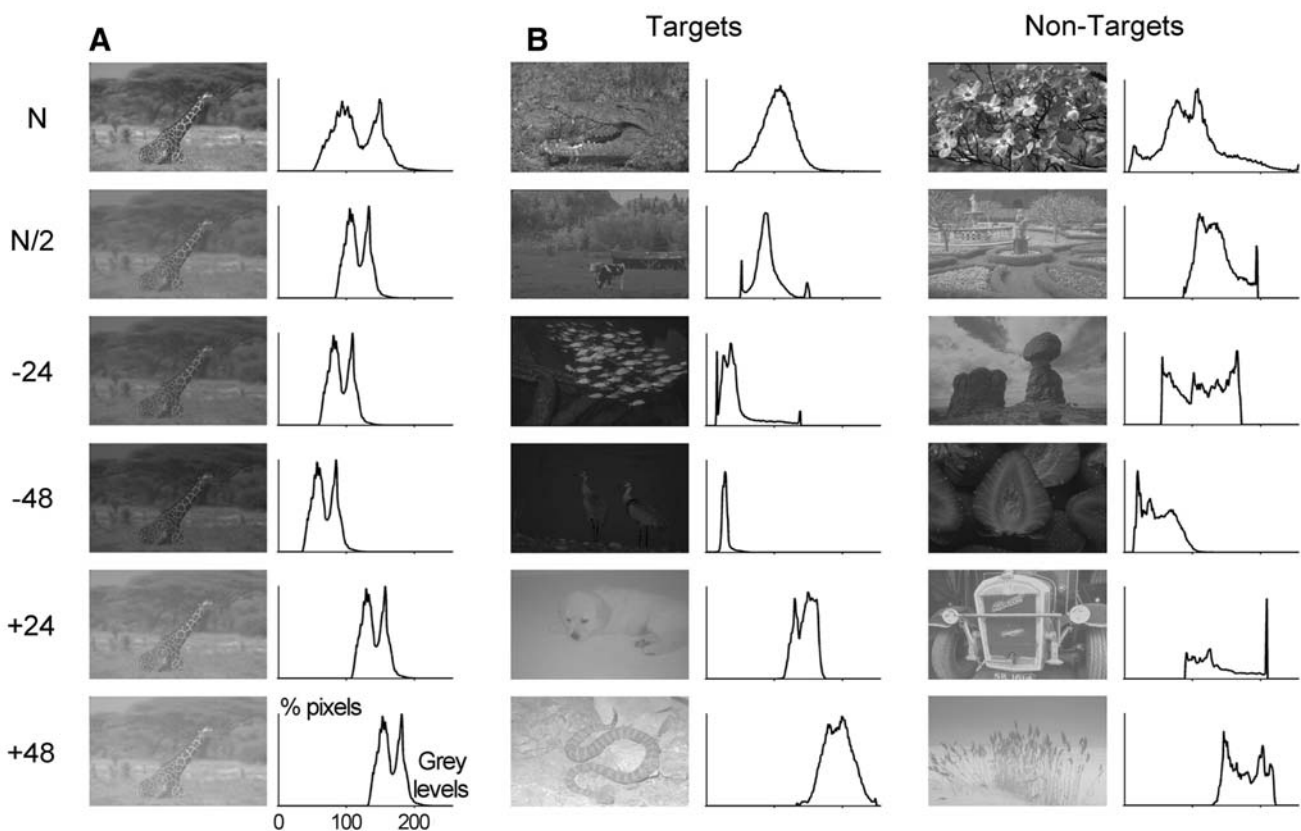
**Human subjects:** 18 human subjects (9 women, 9 men, mean age 27, ranging from 19 to 49) volunteered in this study. As the monkeys were experts in the task, we will consider more closely the results of the 8 expert subjects among our group of 18 subjects, although the performance of experts and naive subjects showed very similar pattern of results. All subjects gave their informed written consent and had normal or corrected to normal vision.

**Monkey subjects:** Two rhesus macaques (male Ry and female Eu) were trained to perform a rapid go/no-go visual animal superordinate categorization task. These monkeys had already been tested in different experiments that showed their ability to categorize familiar photographs, to generalize to new photographs (Fabre-Thorpe et al. 1998) and to perform as well with color or achromatic photo-

graphs (Delorme et al. 2000). In the present study, the monkeys were respectively 7 and 6 years old. All procedures conformed to French and European standards concerning the use of animals in experimentation.

**Stimuli:** We selected a pool of 600 images (300 targets and 300 non-targets) that monkeys had categorized many times during previous training sessions and for which performance had stabilized. The images were thus very familiar to the monkeys. They had been taken from a large commercial CD-ROM library (Corel Stock Photo Libraries) and chosen to be as varied as possible (examples shown in Fig. 1). Targets included fish, birds, mammals, and reptiles presented in their natural environments; non-targets included landscapes, trees, flowers, and non-biological man-made objects, monuments, cars, etc. Both targets and non-targets included close-ups and general views.

These 24-bit (16 millions colors) photographs (384 by 256 pixels) occupied a visual angle of about  $20^\circ \times 13.5^\circ$  for monkeys and  $14.5^\circ \times 10^\circ$  for humans. Most of the images (about three quarters) were horizontal photographs. For the present experiment, the color photographs were



**Fig. 1** Stimuli used in experiment 1. **a** A given target-image example is shown in all tested conditions of luminance from  $N$ : original achromatic photograph to conditions  $N/2$ ,  $-24$ ,  $-48$ ,  $+24$ ,  $+48$ , together with the corresponding distribution of pixel luminance in the range 0–255 ( $Y$ -axis has a variable range to normalize distributions' size).

The distributions of pixel luminance illustrate how the reduction in contrast ( $N/2$ ) allows shifts of luminance without saturation of extreme pixel luminance. **b** Picture examples of target and non-target images are shown in all six testing conditions next to their distribution of pixel luminance

converted to 256 gray-level photographs using the Corel Photo CD tool based on a weighed average of the red, green, and blue channels. During testing, images were flashed for three frames at a refresh rate of 90 Hz resulting in a central picture presentation of around 30 ms on the screen. In this set of original images that was used as the control condition in the present study, the mean luminance was around 106 on the 0–255 gray-level scale, corresponding to a mean value of  $2.27 \text{ Cd/m}^2$  (ranging between  $0.1\text{--}16.1 \text{ Cd/m}^2$ ). To evaluate the distribution of local contrasts, the Michelson contrast was calculated for each pixel of each image relatively to the mean of the eight surrounding pixels. This analysis showed that half of the local contrasts were below 7% and 90% were below 30%. The luminance can also be expressed in trolands that considers the retinal illuminance by taking into account the surface of the pupil. Subjects were tested in semi-darkness, so that their pupil diameter can be roughly estimated (Pong and Fuchs 2000; Winn et al. 1994) at around 7 mm for macaques and 8 mm for humans. Thus, for the set of stimuli, the mean luminance expressed in trolands is somewhat higher for humans than for monkeys (respectively 114 and 87 Td at maximal pupil size).

In the present experiment, there were six different conditions of achromatic stimulus presentation (Fig. 1). In the control condition ( $N$ ), photographs were shown with their original contrast. In a second control condition ( $N/2$ ), the original contrast of the photographs was divided by two, while keeping mean image luminance constant. Finally, four other conditions were built from these  $N/2$  images by shifting the gray-level distribution of pixels toward higher or lower values along the 0–255 scale of gray levels ( $\pm 24$  or  $\pm 48$  luminance values). These four types of stimuli will be referred to as (+24,  $-24$ , +48, and  $-48$ ). These luminance modifications corresponded to an increase or a decrease of about 25 and 50% of the stimulus luminance. The use of  $N/2$  stimuli for luminance shifts instead of  $N$  stimuli avoided a biased saturation of the most extreme gray levels in the +48 and  $-48$  conditions. A total of 3,600 stimuli were thus obtained from the 600 selected targets and non-targets images.

**Task and protocol:** Monkeys and humans were presented with a random succession of different natural scenes, half of which were targets. Subjects (human or monkey) started stimulus presentation by placing one hand on a response pad equipped with infrared LEDs and photodiodes with sub-millisecond precision. When an image that contained an animal was flashed, the subjects had to quickly lift their hand and touch the screen (go response), otherwise they had to keep their hand on the pad (no-go response). Subjects were given a maximum of 1,000 ms to respond, after which any response was considered a no-go response. Monkeys were

restrained in a primate chair during testing (Crist Instruments, GA USA) and were free to move their heads.

Monkeys and humans performed the task in semi-darkness, and they were sat in front of a tactile screen at about 30–35 cm for the monkeys and 40–50 cm for humans. A small fixation cross appeared in the center of the screen, and pictures were flashed around the fixation point on a black background for only 28 ms. The image was not masked, but the short presentation time prevented any exploratory eye movements. The response pad used to start the sequence of images and to record the motor reaction time was located below the screen at waist level. Two successive images were separated by a random 1.6–2-s intertrial period. Correct (go or no-go) responses were rewarded by a beep noise associated, for monkeys, with a drop in fruit juice. Incorrect decisions were followed by a 3-s display of the incorrectly classified stimuli delaying the next trial and allowing time for ocular exploration. The monkeys worked daily for as long as they wanted (1–3 h), 5 days a week. They were provided ad libitum water at the end of each testing session and during weekends.

Before testing, the human subjects performed a short training session with 50 photographs that were distinct from the 600 photographs used in the present experiment and were presented in the  $N$  condition. Each human subject was tested on 3 conditions for each of the 600 stimuli (a total of 1,800 trials in 18 blocks of 100 trials). All conditions were equiprobable, and in each condition, the presentation of targets and non-targets was also equiprobable. Thus, over the group of 18 human subjects, each of the 600 stimuli was seen nine times in each of the six testing conditions.

When they started this experiment, the two rhesus monkeys were familiar with the photographs and had stabilized their performance in the categorization task with their color and achromatic versions. The testing was run in 12 sessions of 300 images to see each of the 3,600 stimuli. The 12 sessions were built at random with the following requirements: in a given series of 300 images, each photograph must be seen only once (thus in one given testing condition) and as for humans, all the testing conditions were equiprobable. A daily session was taken into consideration when the monkey had performed at least 650 trials. The 50-first trials were warm-up trials to allow performance stabilization when monkeys just started performing the task after being brought from the animal facility. Different photographs were used, and the trials were not taken into consideration, after which the sequence of 300 stimuli started repetitively until the monkey stopped working. Monkeys' performance was analyzed on the 1st and 2nd pass of the sequence starting from trial no. 51–651.



**Behavioral analysis:** The performance of both the monkeys and humans was analyzed in terms of accuracy and speed. A response was considered as a go response when the subject lifted his hand from the response pad to touch the screen. This go response was considered as correct on a target and incorrect on a non-target when produced within 1,000 ms after stimulus onset. The reaction time (RT) considered in the present study was measured between stimulus onset and hand lift. Although the movement time measured between hand lift and contact with the screen could have been of potential interest, it was remarkably stable across all conditions and will not be considered here.

Accuracy was evaluated by the percentage of correct (go and no-go) responses. Significant accuracy differences between two conditions were demonstrated using a  $\chi^2$  test. Paired Wilcoxon tests are also applied to assess the reliability of an effect between two conditions over a group of subjects.

Speed of performance was assessed by the RT distribution on correct go responses using 10-ms time bins. All comparisons were made using a Mann–Whitney test as some RT distributions did not pass normality test. Over a group of subjects, median RTs are compared for two different conditions by a paired Wilcoxon test.

## Results

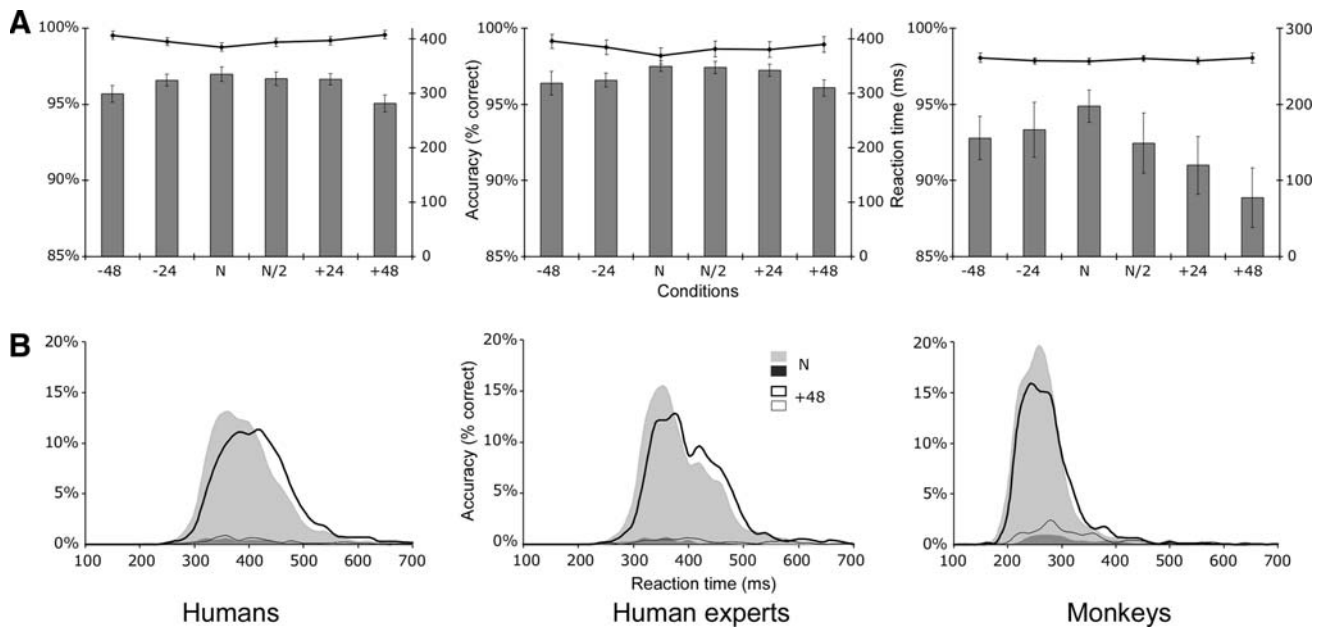
Monkey and human performance was analyzed in terms of accuracy and speed (RT), a summary of the data can be found in Table 1.

**Accuracy performance** (Fig. 2a): Each of the 18 human subjects was tested on 300 trials for each of the six conditions. For the 8 experts among them, overall accuracy was 97.5% correct in the normal  $N$  condition, 97.4% in condition  $N/2$  (compared to  $N$ :  $\chi^2_1 = 0.07$ ,  $p = 0.80$ ), and 97.3% in condition  $+24$  (compared to  $N$  and to  $N/2$ :  $\chi^2_1 = 0.56$ ,  $p = 0.45$  and  $\chi^2_1 = 0.25$ ,  $p = 0.62$ ). Accuracy was 96.6% correct in condition  $-24$ , a very limited drop of 0.8–0.9% that reached significance (compared to  $N$  and to  $N/2$ :  $\chi^2_1 = 6.11$ ,  $p = 0.013$  and  $\chi^2_1 = 5.05$ ,  $p = 0.025$ ). Stimuli presented with maximal luminance shift ( $-48$  and  $+48$ ) induced a limited drop in performance: 96.4% ( $-48$ ) and 96.1% ( $+48$ ), a 1–1.3% difference that was significant relatively to the  $N/2$  condition (respectively  $\chi^2_1 = 7.45$ ,  $p = 0.006$  and  $\chi^2_1 = 11.34$ ,  $p = 0.0008$ ). Note that although robust and significant, this drop in accuracy remained very limited (<1.5% correct): subjects still scored over 96% correct for all the luminance conditions. The same pattern of results was observed with the whole group of 18 subjects with an average accuracy above 95% correct and a drop in accuracy that reached significance only with maximal luminance shifts.

In monkeys, accuracy was 94.9% in the  $N$  condition, and a drop was observed with  $N/2$  stimuli (92.5%,  $\chi^2_1 = 20.10$ ,  $p < 0.001$ ). Accuracy remained stable compared to  $N/2$  (93.3 and 92.8%:  $\chi^2_1 = 2.95$ ,  $p = 0.086$  and  $\chi^2_1 = 0.40$ ,  $p = 0.53$ ) when luminance decreased in conditions  $-24$  and  $-48$ , but a further accuracy drop was observed when luminance was increased with 91.0% correct in condition  $+24$  and 88.8% in condition  $+48$  (compared to  $N/2$ , respectively:  $\chi^2_1 = 6.23$ ,  $p = 0.012$  and  $\chi^2_1 = 31.17$ ,  $p < 0.001$ ).

**Table 1** Overall performance is shown (bold values) for the group of two monkeys and the group of 8 human experts in all 6 conditions tested in experiment 1. Individual performance for monkeys Eu and Ry are also given. SD is indicated in brackets

		-48	-24	$N$	$N/2$	+24	+48
<b>Global accuracy (%)</b>	Monkeys	<b>92.8 (2.0)</b>	<b>93.3 (2.5)</b>	<b>94.9 (1.5)</b>	<b>92.5 (2.8)</b>	<b>91.0 (2.7)</b>	<b>88.9 (2.8)</b>
	Eu	88.8 (2.8)	89.4 (4.9)	92.1 (2.1)	88.5 (4.6)	86.4 (4.7)	84.0 (4.1)
	Ry	96.8 (2.3)	97.3 (2.0)	97.7 (1.4)	96.4 (2.2)	95.6 (1.5)	93.8 (2.3)
	Humans	<b>96.4 (2.2)</b>	<b>96.6 (1.3)</b>	<b>97.5 (1)</b>	<b>97.4 (1.2)</b>	<b>97.3 (1.1)</b>	<b>96.1 (1.5)</b>
<b>Accuracy on targets</b>	Monkeys	<b>95.8 (2.8)</b>	<b>96.3 (1.9)</b>	<b>96.7 (2.1)</b>	<b>95.0 (2.4)</b>	<b>94.8 (2.9)</b>	<b>93.3 (1.8)</b>
	Eu	93.2 (4.3)	94.7 (3.8)	95.3 (2.6)	92.8 (4.1)	92.0 (4.4)	89.3 (3.2)
	Ry	98.5 (2.4)	98.0 (2.1)	98.0 (2.1)	97.2 (2.7)	97.5 (2.3)	97.2 (2.5)
	Humans	<b>97.1 (2.7)</b>	<b>98.9 (1.3)</b>	<b>99.2 (1.6)</b>	<b>99.5 (0.7)</b>	<b>99.6 (0.3)</b>	<b>98.2 (2.4)</b>
<b>Accuracy on non-targets</b>	Monkeys	<b>89.8 (2.0)</b>	<b>90.3 (4.5)</b>	<b>93.1 (2.3)</b>	<b>89.9 (4.9)</b>	<b>87.3 (4.3)</b>	<b>84.5 (4.5)</b>
	Eu	84.3 (2.9)	84.2 (8.2)	88.8 (3.5)	84.2 (8.1)	80.8 (7.2)	78.7 (6.2)
	Ry	95.2 (3.0)	96.5 (2.9)	97.3 (2.2)	95.7 (2.9)	93.7 (3.3)	90.3 (3.6)
	Humans	<b>95.7 (2.4)</b>	<b>94.3 (2.6)</b>	<b>95.8 (2.2)</b>	<b>95.3 (1.9)</b>	<b>94.9 (2.5)</b>	<b>94 (2.5)</b>
<b>Median RT (ms)</b>	Monkeys	<b>261 (9)</b>	<b>257 (6)</b>	<b>257 (6)</b>	<b>260 (5)</b>	<b>257 (6)</b>	<b>261 (10)</b>
	Eu	288 (5)	282 (8)	278 (9)	284 (7)	283 (5)	287 (9)
	Ry	242 (11)	233 (11)	232 (10)	234 (10)	235 (10)	238 (12)
	Humans	<b>398 (37)</b>	<b>386 (38)</b>	<b>377 (42)</b>	<b>383 (40)</b>	<b>383 (41)</b>	<b>393 (40)</b>
<b>Mean RT</b>	Monkeys	<b>270 (7)</b>	<b>264 (5)</b>	<b>264 (9)</b>	<b>267 (7)</b>	<b>269 (10)</b>	<b>272 (11)</b>
	Eu	300 (12)	293 (9)	293 (15)	296 (11)	300 (11)	304 (13)
	Ry	242 (11)	237 (9)	236 (9)	239 (9)	239 (13)	243 (13)
	Humans	<b>408 (37)</b>	<b>394 (38)</b>	<b>383 (42)</b>	<b>391 (38)</b>	<b>393 (40)</b>	<b>403 (41)</b>



**Fig. 2** Performance recorded in different luminance conditions. **a** Global mean accuracy performance in % of correct go and no-go response (*gray histogram*) and median reaction times in ms (*top curve and black diamond* on the left for the whole group of humans, in the *middle* for the human experts and on the *right* for the two monkeys. *Error bars* correspond to the standard error of the mean. Note that, for reaction time, the *Y-axis* is different for humans and monkeys as monkeys are

much faster than humans. **b** Overall reaction time distributions of go responses. The number of correct and incorrect go responses is expressed by 20-ms bins in percentage of the total number of go responses. The maximal performance impairment was observed between the two illustrated conditions: condition *N* (*shaded histogram*) and condition *+48* (*empty histograms*); *top curves*: correct go responses; *bottom curves*: incorrect go responses

Thus, with large shifts of luminance (*+48* and *-48*), performance was always affected, and the impairment was maximal with the *+48* shift in both species. In all other conditions, accuracy was robust in humans, although there was an accuracy drop of 1–6% of correct responses in monkeys.

**Speed of response** (Fig. 2b): In humans, the median reaction time computed on the whole RT distribution was 377 ms in condition *N*. Relatively, to the *N* stimuli, a significant increase in RT (Mann–Whitney *U* ranging  $5.35 \times 10^5$ – $6.50 \times 10^5$ ,  $n_1 = 1,190$ ,  $n_2$  ranging 1,165–1,195, all  $p < 0.0003$ ) was observed in all other conditions (median RT: *N/2* = 383 ms; *-24* = 386 ms; *+24* = 383 ms; *-48* = 398 ms and *+48* = 393 ms). Limited to 10 ms in the 3 intermediate conditions of perturbation, the RT increase reached about 20 ms (16–21 ms) in the two extreme luminance conditions. The RT increase was robust within the group of subjects (paired Wilcoxon: *Z* ranging 3.37–3.72, all  $p < 0.0008$  for all comparisons with the *N* condition).

The monkeys were much faster than humans, confirming previous studies (Fabre-Thorpe et al. 1998; Delorme et al. 2000; Macé et al. 2005b; Girard et al. 2008). Median RT was very stable (range 257–261 ms) for all conditions. However, there was a slight shift toward longer latencies in the two extreme luminance conditions *-48* and *+48* (median RT 261 ms). This increase was statistically significant

for condition *+48* (paired Wilcoxon:  $Z = 2.94$ ,  $p = 0.003$ ) but did not reach significance with the *-48* stimuli (paired Wilcoxon:  $Z = 1.88$ ,  $p = 0.06$ ). Note that the observed increase for the two extreme conditions compare to *N* was limited to a few milliseconds (see mean RT in Table 1). The RT distributions in conditions *N* and *+48* are shown in Fig. 2b and illustrate the maximal shift of response latencies.

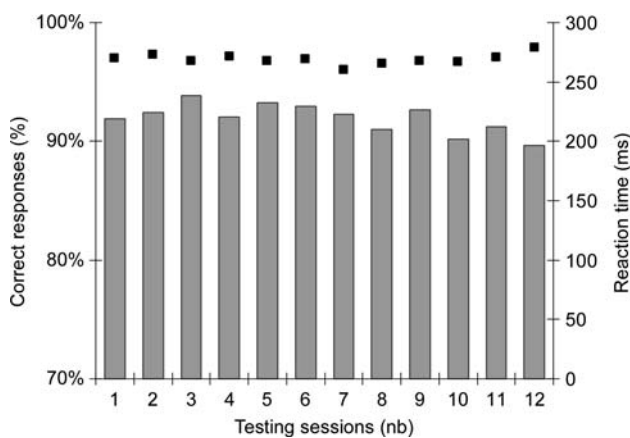
To summarize the results, for both humans and monkeys, performance was impaired in the two extreme conditions of luminance, and the impairments were most pronounced for the highest luminance condition. Whereas human performance was mainly affected in terms of response speed, monkey performance was affected in terms of accuracy.

**Evolution of performance across sessions: is there an effect of training in monkeys?** The monkeys saw the stimuli many times and were always rewarded for correct go responses. There is a large literature showing that task performance of monkeys is affected by reward size, delay to reward, and satiation. In a recent paper (Minamimoto et al. 2009), the error rate was shown to increase from about 10 to 40% with reward size decreasing from 8 drops to 1 drop of juice. But, monkeys had to infer the reward size from a cue that had no pertinence for task performance. In other

words, if monkeys can associate new stimulus conditions with a lack of reward, they will tend to perform badly, and their ability on the new conditions will be largely underestimated. On the other hand, when rewarding the new conditions, the results could reflect an effect due to training. Indeed, humans (Brady et al. 2008) and baboons (Fagot and Cook 2006) have been shown to have such a huge memory capacity that learning could interfere with our results. To determine the potential effect of learning, we computed the accuracy and RT for each of the 12 testing sessions. As illustrated in Fig. 3, the performance level was very stable from the first to the last repetition of the stimuli. This stability is seen both for accuracy, which fluctuates around a mean at 92% correct, and for mean RT that fluctuates around 269 ms. Note that in each session (1 to 12) all six conditions of presentation are almost equally represented. Thus, the effects reported for variations of luminance are not compounded by learning.

## Discussion

Both monkeys and humans showed very robust performance despite large variations of global luminance. Very little deficit is seen when subjects (both monkeys and humans) have to categorize stimuli that randomly appear with large variations of luminance. In the most challenging situation (+48), the performance is only slightly impaired with either an accuracy drop for monkeys of about 6.5% of correct responses or a RT increase of about 20 ms for human experts. The effects of luminance variation on stimulus contrast were not symmetric when luminance is increased or decreased. Expressed in pixel luminance, it is



**Fig. 3** Evolution of the mean accuracy performance in % of correct go and no-go response (*gray histogram*) and mean reaction times in ms (*top black squares*) throughout all testing sessions performed by the two monkeys in experiment 1. SD are not indicated as each data point represents the average of only two individuals with a large difference in their level of performance

easier to discriminate two pixels with an absolute luminance of 10 and 20 than a respective luminance of 210 and 220. This is because the ratio between the two values is different. Thus, it is not really surprising to find that the condition with the highest luminance is the most difficult condition in experiment 1. What is surprising are the remarkably small effects induced by such large luminance variations. Another interesting result is that the deficits observed in the +48 condition appeared more as an accuracy decrease in monkeys and a RT increase in humans. In fact, it is often true that when facing difficult situations, the performance speed–accuracy trade-off is affected in terms of speed in humans that favor high accuracy and in terms of accuracy in monkeys that tend to keep performing at the same speed with a corresponding decrease of accuracy.

Thus, the ability of humans and monkeys to detect the presence of an animal in achromatic natural scenes remains very high even when shown at half the original contrast and presented randomly to the subject with large variations of luminance.

In experiment 2, we further explored how monkey and human performance would be affected by large reductions in contrasts.

## Experiment 2: variations of contrast

Previous studies showed that the performance of human subjects is robust to variations of contrast (Macé et al. 2005a). The present experiment was run on the two same rhesus monkeys to compare the impact of contrast variations on categorization. Object recognition is believed to rely on the processing of visual information along the ventral cortical pathway, and mainly upon parvocellular information. Indeed, high acuity details of object shape and color information are carried by the parvocellular system, but parvocellular retinal cells are not activated at contrast levels below 10% (Kaplan and Shapley 1986, see Macé et al. 2005a for references). Thus, at very low contrast levels, object vision probably relies on visual information carried by the “magnocellular stream.” The magnocellular system is fast and still activated at low contrast levels, but the information it carries is both achromatic and of poor spatial resolution. Visual representations built from magnocellular information can only be very coarse, and although they might provide the only visual information available in natural conditions such as dusk or dawn, the potential role of the magnocellular stream in object recognition has been largely ignored. The performance levels reached by monkeys with stimuli at low contrasts thus has implications for the potential role of magnocellular information in object categorization.

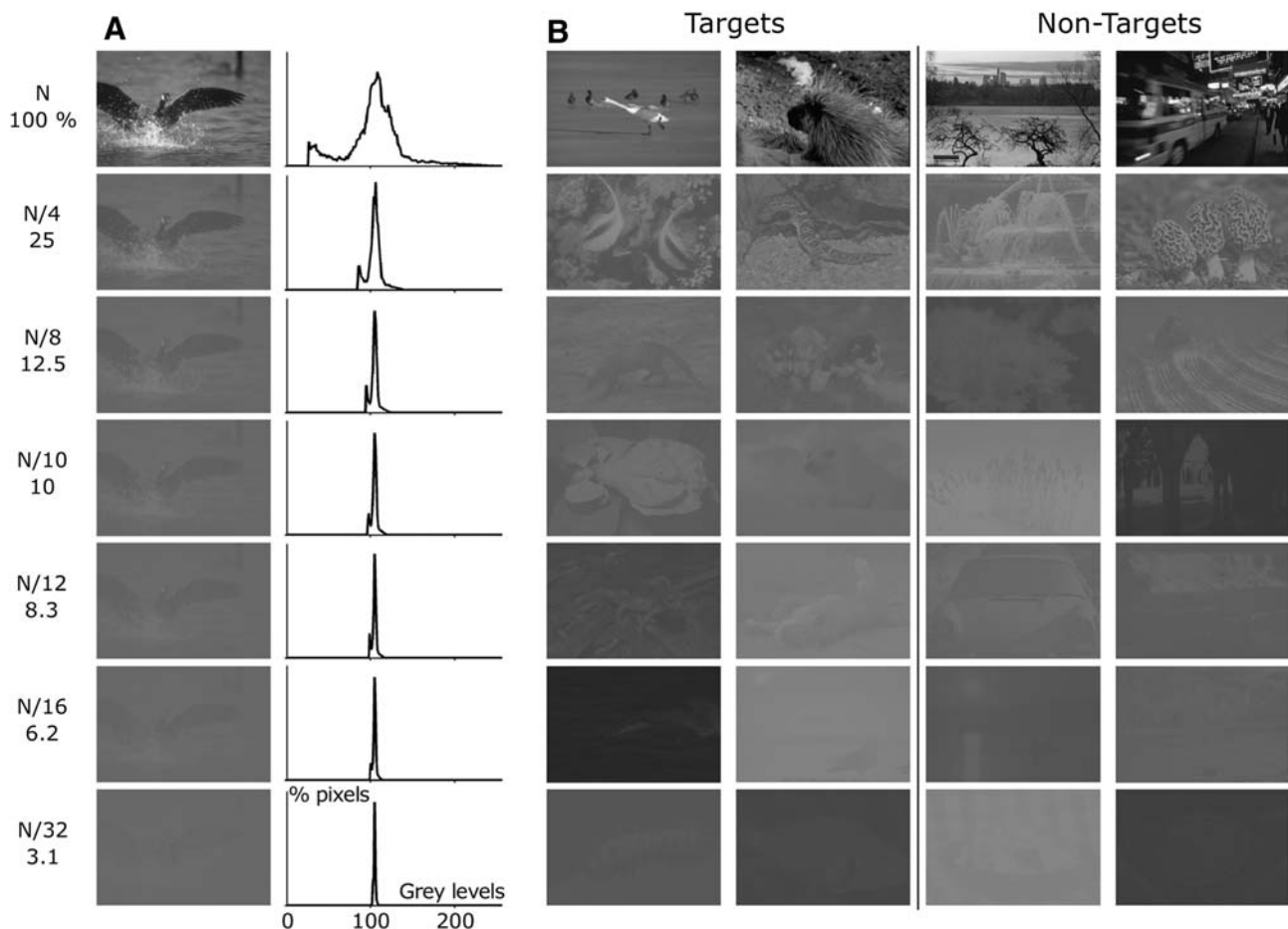
## Methods

**Monkey subjects:** The two rhesus monkeys (male Ry and female Eu) tested in experiment 1 also participated in experiment 2.

**Stimuli:** The stimuli were constructed from the 600  $N$  stimuli used in experiment 1. Following the protocol used to build the  $N/2$  stimuli from the original gray-level ( $N$ ) images, we divided the contrasts around the mean luminance of each stimulus to create  $N/4$ ,  $N/8$ ,  $N/10$ ,  $N/12$ ,  $N/16$ , and  $N/32$  stimuli. As 100% contrast was attributed to the original  $N$  condition, the residual contrast in the constructed stimuli was thus 25% at  $N/4$  and decreased down to 3% in condition  $N/32$  (Fig. 4). This residual contrast is a strong overestimation of the overall local contrasts in the test photographs. Contrast studies have used regular sine wave gratings or checkerboard patterns, but this type of artificial stimulus is very different from a natural image. Local contrasts of natural scenes hardly ever reflect the optimal 100%

Michelson contrast as maximum and minimum pixel values are virtually never placed next to each other (for a detailed analysis of these stimuli, see (Macé et al. 2005a)). The monkeys were tested with the eight conditions from  $N$  to  $N/32$  (4,800 stimuli). Due to the different sensitivity to contrast of the two main visual pathways (parvocellular vs. magnocellular), the contrast was reduced with smaller steps around 10% residual contrast that corresponds to the parvocellular threshold. Reducing contrast induces a strong alteration in the stimulus discriminability. The  $N$  stimuli are well defined when 256 different gray-level values are available. But, contrast reduction narrows the distribution of pixel luminance values (Fig. 4), and  $N/8$  stimuli are built with a maximum of 32 consecutive gray levels, a value that decreases for further contrast reductions (25 for  $N/10$ , 16 for  $N/16$  and 8 for  $N/32$ ). The texture of the stimuli that subjects had to categorize could thus be very coarse in certain conditions.

**Tasks and protocol:** The task and protocol were similar to experiment 1.



**Fig. 4** Stimuli used in experiment 2. **a** A given target-image example is shown in all further tested conditions ( $N$ ,  $N/4$ ,  $N/8$ ,  $N/10$ ,  $N/12$ ,  $N/16$ , and  $N/32$ ) together with the corresponding distribution of pixel luminance in the range 0–255 ( $Y$ -axis has variable ranges to normalize

distributions' size). The figure illustrates the reduction in contrast around the mean luminance value of the image. **b** Picture examples of target and non-target images are shown in all conditions of contrast reductions



The two rhesus monkeys performed experiment 2 after experiment 1. Each of the 600 images was seen 50 times in total but was not presented the same number of times at each contrast: 14 times ( $N$ ), 10 times ( $N/2$ ,  $N/4$ ), 6 times ( $N/8$ ), 4 times ( $N/10$ ), and twice for the last three extreme conditions. The unbalanced number of trials per condition was used in order to ensure that monkeys were able to maintain a percentage of correct response around 80–85% correct. As in experiment 1, an erroneous response was followed by a 3s display of the incorrectly classified stimuli. This delayed the next trial (thus the next possibility of reward), and monkeys could become agitated and lose concentration when the proportion of correct trials was too low. The sessions were arranged, so that any given image was only seen once and in one condition of contrast over a period of 2 days. Thus, the present results were recorded over a period of 50 daily sessions for a total of 30,000 trials per monkey.

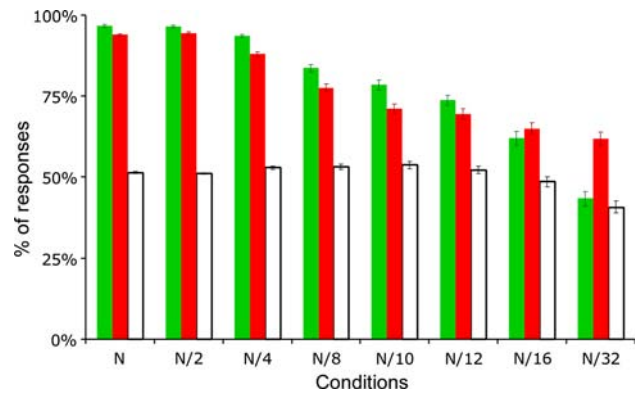
As in experiment 1, a daily session was only taken in consideration when the monkey had performed at least 650 trials. All the other sessions were discarded as performance tends to be highly variable when monkeys are not motivated enough. The 50-first trials were warming-up trials with different photographs to allow performance stabilization and were not taken into consideration, after which the sequence of the 300 test stimuli started repetitively until the monkey stopped working. Monkeys' performance was analyzed on the 1st and 2nd run of the sequence starting from trial no 51–651.

## Results

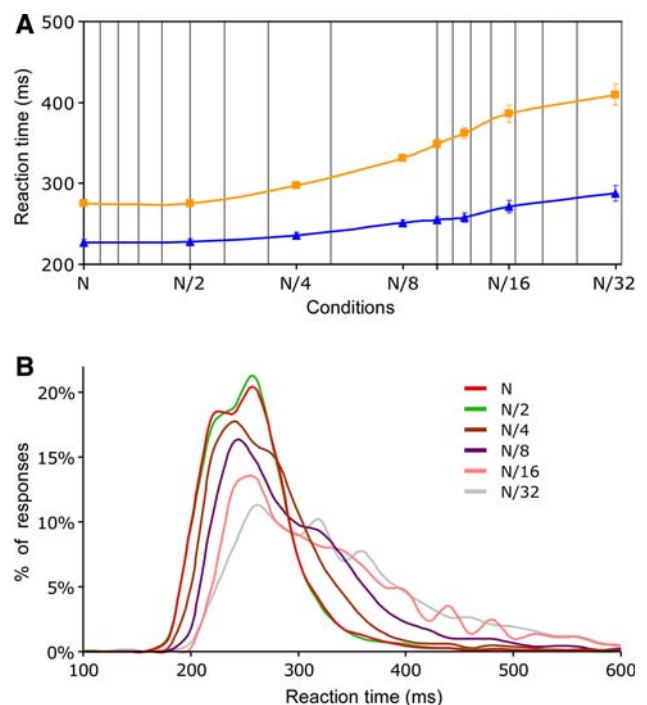
As in experiment 1, the performance of the two monkeys was analyzed in terms of accuracy and speed of response (RT); for individual scores see Table 2.

**Accuracy performance** (Fig. 5 and Table 2): In the  $N$  condition, monkeys scored 95.3% correct. Contrary to experiment 1, no drop was observed with  $N/2$  stimuli (95.4% correct). As could be expected, further reductions in contrast induced a progressive global accuracy drop, but chance level was only reached in the most extreme contrast conditions ( $N/4 = 90.8\%$ ,  $N/8 = 80.6\%$ ,  $N/10 = 74.8\%$ ,  $N/12 = 71.6\%$ ,  $N/16 = 63.4\%$ ,  $N/32 = 52.5\%$ ). This accuracy drop was always statistically significant from one condition of contrast to the next ( $\chi^2_1$  ranging 69.0–4,537.1, all  $p < 0.001$ ). For intermediate conditions ( $N/8$ – $N/12$ ), accuracy remained at a good level (80.6, 74.8, and 71.6% correct) even though the visual system had to process images in which the mean local contrast values were virtually all below 10% (Macé et al. 2005a).

Monkeys were slightly more accurate at responding to targets than at ignoring non-targets (about 96.7 vs. 94.2%)



**Fig. 5** For all contrast reductions from  $N$  to  $N/32$ , the mean percentage of correct go responses on targets (in green) and correct no-go responses on non-targets (in red) is computed for the two monkeys. For each condition, the empty histogram indicates the percentage of total (correct and incorrect) go responses that keeps around 50% except for the extreme condition of contrast reduction



**Fig. 6** Reaction times in selected contrast conditions tested in experiment 2. **a** Median RT for each of the two monkeys (Eu orange squares and curve and Ry blue triangles and curve). The standard errors of the mean are indicated (often too small to be visible). **b**. The distribution of correct go responses toward targets is shown for six of the eight conditions tested:  $N$  (red),  $N/2$  (green),  $N/4$  (brown),  $N/8$  (violet),  $N/16$  (pink), and  $N/32$  (gray). Note the progressive shift of the earliest go responses toward longer latencies together with the flattening of the curves

in condition  $N$  and  $N/2$ . For monkey Ry, this advantage for targets was maintained for all contrast reductions except  $N/32$ . For monkey Eu, the bias toward more accurate responses to targets was inverted from  $N/10$  to  $N/32$ .

**Table 2** Overall performance is shown (bold values) for the group of two monkeys and the group of 16 human experts in all contrast conditions tested in experiment 2. Individual performance for monkeys Eu and Ry are also given. SD is indicated in brackets

	<i>N</i>	<i>N/2</i>	<i>N/4</i>	<i>N/8</i>	<i>N/10</i>	<i>N/12</i>	<i>N/16</i>	<i>N/32</i>
<b>Global accuracy (%)</b>								
Monkeys	<b>95.3 (1.8)</b>	<b>95.4 (2.0)</b>	<b>90.8 (2.8)</b>	<b>80.6 (5.4)</b>	<b>74.8 (6.9)</b>	<b>71.6 (8.2)</b>	<b>63.4 (8.6)</b>	<b>52.5 (8.6)</b>
Eu	92.6 (2.9)	92.6 (3.8)	85.9 (3.9)	72.4 (7.9)	66.9 (10.0)	63.8 (11.2)	57.1 (11.4)	49.3 (8.4)
Ry	98.0 (1.5)	98.2 (1.6)	95.7 (3.4)	88.8 (6.1)	82.7 (8.5)	79.4 (10.2)	69.8 (9.9)	55.8 (12.1)
Humans	<b>94.9 (3.9)</b>		<b>88.3 (6.1)</b>	<b>79.2 (8.1)</b>	<b>71.3 (9.1)</b>	<b>64.5 (9.1)</b>	<b>59.3 (4.1)</b>	<b>50.4 (5.1)</b>
<b>Accuracy on targets</b>								
Monkeys	<b>96.7 (2.5)</b>	<b>96.5 (2.5)</b>	<b>93.6 (3.0)</b>	<b>83.6 (8.0)</b>	<b>78.4 (10.7)</b>	<b>73.7 (11.2)</b>	<b>61.9 (15.7)</b>	<b>43.3 (15.6)</b>
Eu	95.0 (4.5)	94.4 (4.6)	89.6 (5.1)	71.7 (15.1)	64.9 (16.6)	59.0 (18.6)	45.2 (22.7)	31.0 (17.4)
Ry	98.4 (1.8)	98.5 (2.1)	97.6 (2.7)	95.5 (6.1)	91.9 (11.1)	88.3 (13.5)	78.7 (18.5)	55.7 (23.6)
Humans	<b>99.5 (1.2)</b>		<b>95.8 (3.6)</b>	<b>86.5 (10.9)</b>	<b>76.8 (15.1)</b>	<b>67.0 (19.3)</b>	<b>54.2 (14.8)</b>	<b>31.0 (20.0)</b>
<b>Accuracy on non-targets</b>								
Monkeys	<b>94.0 (2.2)</b>	<b>94.4 (2.9)</b>	<b>88.0 (5.1)</b>	<b>77.6 (7.7)</b>	<b>71.2 (10.2)</b>	<b>69.5 (11.6)</b>	<b>64.9 (12.5)</b>	<b>61.8 (14.5)</b>
Eu	90.3 (3.9)	90.8 (5.3)	82.2 (6.8)	73.1 (10.4)	68.9 (12.3)	68.5 (17.8)	69.0 (18.3)	67.5 (19.1)
Ry	97.7 (1.9)	97.9 (2.2)	93.8 (5.4)	82.0 (10.1)	73.4 (12.5)	70.5 (15.5)	60.8 (13.2)	56.0 (19.6)
Humans	<b>90.3 (7.4)</b>		<b>80.7 (11.5)</b>	<b>71.8 (14.8)</b>	<b>65.7 (15.7)</b>	<b>62.0 (12.2)</b>	<b>64.3 (17.0)</b>	<b>69.8 (21.5)</b>
<b>Response percentage</b>								
Monkeys	<b>51.3 (1.6)</b>	<b>51.1 (1.8)</b>	<b>52.8 (3.1)</b>	<b>53.0 (5.7)</b>	<b>53.6 (7.8)</b>	<b>52.1 (7.9)</b>	<b>48.5 (11.3)</b>	<b>40.8 (12.3)</b>
Eu	52.4 (3.0)	51.8 (3.2)	53.7 (4.6)	49.3 (10.3)	48.0 (10.7)	45.3 (14.3)	38.1 (17.2)	31.8 (16.2)
Ry	50.3 (1.1)	50.3 (1.5)	51.9 (2.6)	56.8 (5.7)	59.3 (8.2)	58.9 (10.4)	58.9 (12.7)	49.8 (18.0)
Humans	<b>54.6 (6.2)</b>		<b>57.6 (7.3)</b>	<b>57.3 (11.2)</b>	<b>55.6 (12.3)</b>	<b>52.5 (13.8)</b>	<b>44.9 (15.7)</b>	<b>30.6 (20.2)</b>
<b>Median RT (ms)</b>								
Monkeys	<b>252 (7)</b>	<b>252 (8)</b>	<b>265 (10)</b>	<b>280 (16)</b>	<b>289 (19)</b>	<b>295 (29)</b>	<b>309 (37)</b>	<b>324 (40)</b>
Eu	275 (9)	275 (11)	297 (15)	331 (22)	349 (38)	362 (45)	386 (73)	410 (92)
Ry	227 (13)	228 (13)	236 (12)	251 (15)	255 (15)	258 (22)	271 (28)	288 (34)
Humans	<b>407 (46)</b>		<b>411 (46)</b>	<b>441 (47)</b>	<b>445 (49)</b>	<b>463 (48)</b>	<b>464 (54)</b>	<b>461 (99)</b>
<b>Mean RT</b>								
Monkeys	<b>260 (8)</b>	<b>259 (9)</b>	<b>276 (11)</b>	<b>300 (15)</b>	<b>313 (22)</b>	<b>318 (25)</b>	<b>332 (30)</b>	<b>349 (40)</b>
Eu	289 (11)	289 (16)	314 (18)	353 (27)	378 (38)	385 (43)	409 (67)	436 (90)
Ry	231 (12)	231 (13)	242 (14)	260 (17)	268 (18)	274 (31)	288 (29)	301 (36)
Humans	<b>414 (45)</b>		<b>421 (44)</b>	<b>456 (48)</b>	<b>454 (53)</b>	<b>480 (52)</b>	<b>478 (49)</b>	<b>482 (98)</b>

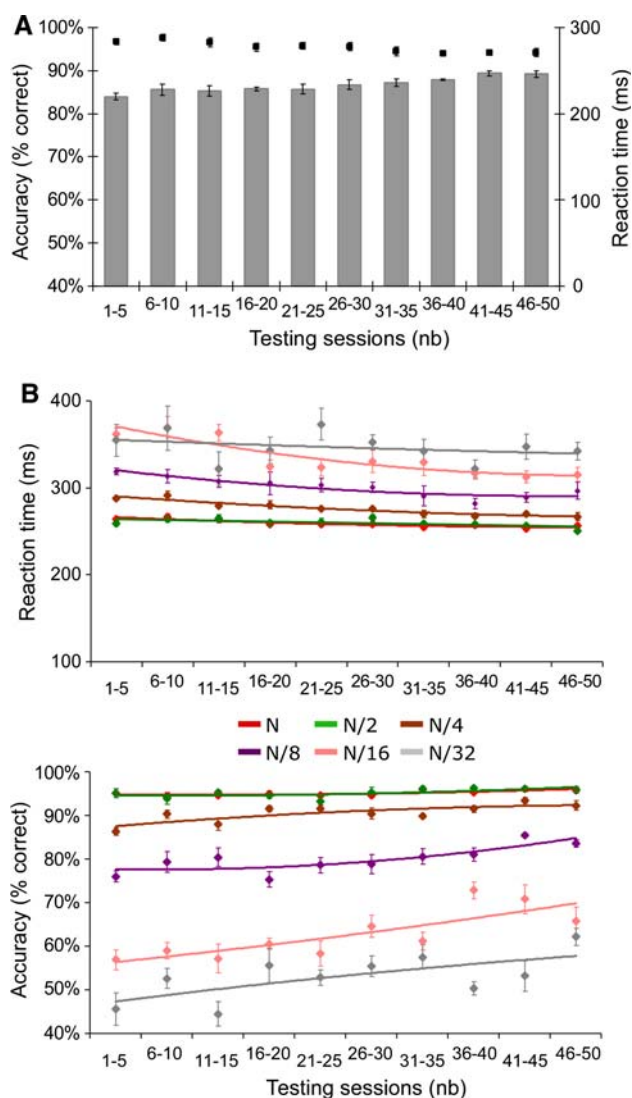
Interestingly, although performance of monkey Eu decreased from 95% correct to 31% correct on targets, it only dropped from 90 to 68% correct on non-targets. Since targets and non-targets were equiprobable, go responses should be recorded on 50% of the cases, and we checked for biases. Overall, monkey Eu responded on about 50% of the trials, but the percentage of go responses dropped abruptly for the three extreme conditions of contrast (*N/12*–*N/32*). The monkey thus appeared to be coding for the animal go-category and to respond only when enough information had been accumulated to decide (with a high probability of being correct) that the photograph contained an animal.

**Speed of response** (Fig. 6): For the *N* and *N/2* conditions, reaction time was unaffected (median RT: *N* = 252 ms, *N/2* = 252 ms). Further reductions in contrast induced a progressive RT increase (median RT: *N/4* = 265, *N/8* = 280, *N/10* = 289, *N/12* = 295, *N/16* = 309, *N/32* = 324 ms). This RT increase reached 72 ms between conditions *N* and *N/32*. The two monkeys displayed similar performance impairments (Fig. 6a and Table 2).

The RT distributions for the different contrast conditions are illustrated in Fig. 6b. Whereas RT distributions in *N* and *N/2* conditions superimpose well, further contrast reductions induced a clear shift and a spreading of the distributions toward longer response latencies. All responses were affected, including the earliest ones. Calculated when the first 5 percentile of correct go responses had been produced,

the response latency showed a progressive increase with contrast reduction from 200 (*N*) to 230 ms (*N/32*). This result suggests that the minimal amount of information used to trigger the earliest responses is available later and later as contrast is reduced.

**Evolution of performance across sessions:** As in Experiment 1, the monkeys had to process the same stimuli many times in different contrast conditions, and since all correct decisions were systematically rewarded, an effect of training could confound the data. To determine the potential effect of learning, we computed the global accuracy and global mean RT by group of 5 sessions for the 50 successive sessions (Fig. 7a). The results showed a progressive increase in accuracy (from 84% correct to 89.2%) and a progressive decrease in median RT (from 283 ms to 271 ms). To further analyze this performance improvement, we considered separately the different contrast conditions (Fig. 7b), and we compared statistically the scores (global accuracy and mean RT) reached during the first 10 sessions (groups 1 & 2) with those reached in the 10 last sessions (groups 9 & 10) using Mann–Whitney tests. Accuracy was significantly improved for all conditions except *N* and *N/2* (Mann–Whitney *U* ranging 2–20, *n*<sub>1</sub> = 10, *n*<sub>2</sub> = 10, all *p* < 0.02); mean RTs were significantly shortened in all conditions except for *N/32* (Mann–Whitney *U* ranging 0–19, *n*<sub>1</sub> = 10, *n*<sub>2</sub> = 10, all *p* < 0.02). Of course, the effect on mean RT with original (*N*) images was very limited



**Fig. 7** Evolution of monkeys' performance in experiment 2. **a.** Evolution of the mean accuracy performance in % of correct go and no-go response (*gray histogram*) and mean reaction times in ms (*top black squares*) throughout all testing sessions (grouped by five) performed by the two monkeys. *Error bars* indicate the standard error of the mean for both accuracy and reaction times over the group of five testing sessions. **b.** *Top:* evolution of the mean reaction times shown in **a**, split across contrast conditions. *Bottom:* evolution of the accuracy in % of correct go and no-go responses presented in **a**, split across contrast conditions. The *diamonds* represent the actual data, and the curves are polynomial fits to the data. *Error bars* indicate the standard error of the mean for both accuracy and reaction times

(10 ms), but these results show that the introduction of degraded stimuli had disturbed the monkey in the task performance even when processing the original familiar photographs. The results also show some differential effects of training across conditions. The accuracy increase was maximal and reached 10–11.5% correct for intermediate conditions  $N/10$ ,  $N/12$ ,  $N/16$  (Mann–Whitney  $U$  ranging 2–16.5,  $n_1 = 10$ ,  $n_2 = 10$ ,  $p$  ranging 0.011–0.0003). The effect on response speed was maximal for condition  $N/16$  with a RT

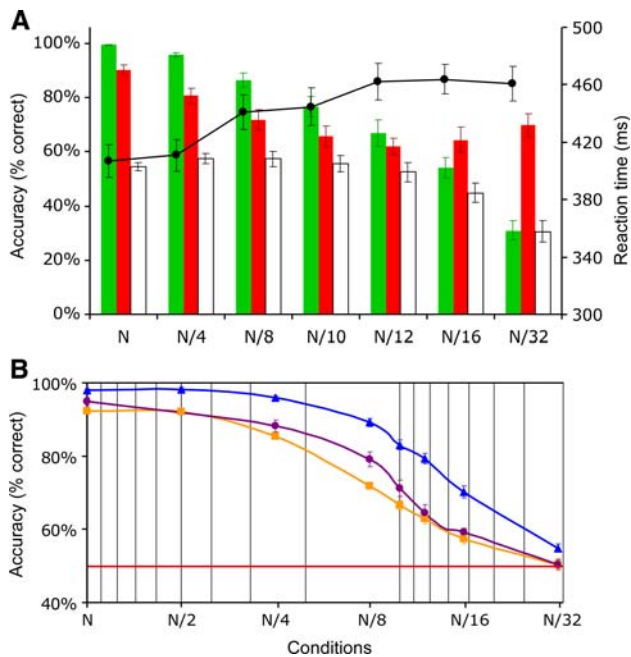
decrease of 52 ms (Mann–Whitney  $U = 2$ ,  $n_1 = 10$ ,  $n_2 = 10$ ,  $p = 0.0003$ ).

Finally, the analysis shows that the effect of contrast reductions follows the same pattern at the beginning and at the end of the 50 testing sessions, so that the relative effect linked to contrast reduction is not affected by learning.

*Comparison between humans and monkeys:* In a preceding study (Macé et al. 2005a), a group of 24 human subjects was tested on 1728 images that included the 600 images used here for monkeys. Humans showed accuracy drops and increased RT that were very similar to those seen here for the monkeys. Here we re-analyzed the human performance restricting the study to the same 600 images tested in the two monkeys for better comparison. Expertise can also be a bias as the monkeys were extensively trained on the task, but human subjects rapidly reached their maximal scores in this categorization task (Fabre-Thorpe et al. 2001). Therefore, for better comparison with monkeys, we also restricted the analysis to the human subjects that could be considered as expert in the task, because they had participated in at least one of our preceding studies using the animal categorization task (16 subjects out of 24).

*Accuracy performance* (Fig. 8): In the  $N$  condition, human experts scored 94.9% correct. Condition  $N/2$  was not tested in this 2005 study, but further reductions in contrast induced, as in monkeys, a progressive global accuracy drop ( $N/4 = 88.3\%$ ,  $N/8 = 79.2\%$ ,  $N/10 = 71.3\%$ ,  $N/12 = 64.5\%$ ,  $N/16 = 59.3$ ). For humans, as for monkeys, chance level was only reached for the most extreme condition of contrast ( $N/32 = 50.4\%$ ). All the contrast reductions induced a statistically significant drop in accuracy relative to the preceding contrast condition ( $\chi^2_1$  ranging 13.7–60.0, all  $p < 0.001$ ). Although we took care to compare the monkeys' performance with those of human experts, it has to be noted that very similar results were obtained when performance was analyzed on the whole group of 24 subjects that included the 8 naïve subjects ( $N = 92.6\%$ ,  $N/4 = 86.3\%$ ,  $N/8 = 76.1\%$ ,  $N/10 = 70.6\%$ ,  $N/12 = 62.9\%$ ,  $N/16 = 59.2\%$ ,  $N/32 = 50.6\%$ ). As for monkeys, accuracy remained quite robust for intermediate conditions ( $N/8$ – $N/12$ ) despite the fact that all the mean local contrast values in the stimuli are below 10%.

The human group showed a bias toward higher accuracy on targets from condition  $N$  until  $N/12$ , the bias was then inverted in the last two conditions (Fig. 8a). Performance decreased from 99% correct to 31% correct on targets, it only dropped from 90 to 70% correct on non-targets. Conversely, for these two extreme conditions of contrast, go responses (that should be recorded in 50% of the cases) dropped below 50% to reach just 29% at  $N/32$ . These performance scores are remarkably similar to those displayed by monkey Eu (Fig. 8b).



**Fig. 8** Comparison between human and monkey performance in the same conditions of contrast reduction. **a** For contrast reduction from  $N$  to  $N/32$ , the accuracy computed on the group of 16 expert human subjects is illustrated as in Fig. 4 for monkeys: mean percentage of correct go responses on targets (in green), mean percentage of correct no-go responses on non-targets (in red). For each condition, the empty histogram indicates the percentage of total (correct and incorrect) go responses that keeps around 50–60% except for the two extreme conditions of contrast reduction. Black dots and curve illustrate the median RT of correct go response for the different task conditions. Error bars indicate the standard error of the mean for both accuracy and reaction times. **b** Mean global accuracy computed over the group of 16 expert human subjects (violet dots and curve) is illustrated in seven out of eight contrast conditions tested from  $N$  to  $N/32$  together with the accuracy performance of each of the two monkeys (Eu orange squares and curve; Ry blue triangles and curve). Error bars indicate the standard error of the mean. The red line at 50% indicates chance level. Note the similarity between monkey and human performance

**Speed of response:** With  $N$  stimuli, median reaction time was 407 ms. Further reductions in contrast induced a progressive RT increase (mean RT:  $N/4 = 411$ ,  $N/8 = 441$ ,  $N/10 = 445$ ,  $N/12 = 463$ ,  $N/16 = 464$ ,  $N/32 = 461$  ms). This RT increase reached a plateau at  $N/12$ , with a maximum increase of about 60 ms between  $N$  and the three extreme conditions of contrast (Fig. 8a).

## Discussion

The performance of both macaques and humans is also very robust to reductions of contrast. Although a progressive drop in accuracy is observed, the initial contrast of the photograph has to be divided by 32 for performance to drop to chance level in both species. In such altered images, subjects have to base their responses on a very limited amount of information. In the original  $N$  condition, only 3% of the

images used the full range of 256 gray levels; although 90% of the stimuli used over 200 gray levels, (Macé et al. 2005a). When contrast is decreased, the sharpness of the stimulus drops dramatically as the number of possible gray levels becomes very limited; no more than 25 gray levels with  $N/10$  stimuli when subjects still score 70–75% correct and no more than 16 gray levels in the  $N/16$  condition in which both humans and monkeys score around 60% correct. Moreover, one has to keep in mind that photographs are flashed for only 28 ms.

The visual system of primates appears well adapted to the visual world they have to deal with. It works very efficiently in challenging natural conditions. Indeed, in everyday life, visual conditions are often far from optimal; at dusk or dawn for example, luminance and contrast can be very low, and surrounding objects are blurred in foggy situations. These results are in agreement with other data reporting impressive performance in categorization of digits and letters with reduced contrast (Strasburger et al. 1991; Strasburger and Rentschler 1996); at least when the task is performed using central vision. But such categorization can be easily influenced by rote learning as there is a small number of digits and letters. In the animal task used here, the wide range of unpredictable target forms and sizes could have induced a dramatic drop in performance with contrast reduction. The high level of categorization performance in monkeys with very degraded stimuli is seen even at the beginning of the testing sessions. Along the 50 sessions, the monkeys certainly learnt to deal with the most degraded images ( $N/4$ ,  $N/8$ ;  $N/16$ ), but the patterns of results remained the same between the different conditions.

The robust categorization performance obtained in monkeys and humans can be related to the impressive response invariance of high-level cortical areas of the ventral visual stream that respond selectively to complex objects (faces, animals...). Indeed, such invariance has been reported using natural photographs of faces and neuronal response in monkeys (Rolls and Baylis 1986), or using line drawings of complex objects and faces and fMRI activations in humans (Avidan et al. 2002). The large response invariance of populations of neurons selective to objects or object categories could underlie the robustness observed here in terms of behavioral performance. Reaction times were marginally increased in Experiment 1 with luminance variations. With contrast reductions, the RT increases were much more pronounced and were often associated with drops in accuracy. This may be related to the fact that in the striate cortex, reduction in contrast induces both a decrease in firing rate and a delay of response onset (Albrecht et al. 2002). In contrast, luminance-reduction does not have much effect on the shape of the response, in agreement with the relatively mild behavioral effects observed in Experiment 1 with luminance variations.



Although monkeys were tested with very familiar images and humans were tested with images they had not seen before, the second striking observation is the extreme similarity between the performance impairments induced by stimulus manipulation in humans and monkeys. This is true even when considering the performance improvement observed during the 50 sessions of testing in monkeys. In experiment 2, contrast reduction induced similar impairments in both species. The mean accuracy calculated for each contrast condition over the group of human subjects follows exactly the performance of monkey Eu (Fig. 8b). This is true for global accuracy, accuracy on targets and non-targets, and even for the rate of go responses produced in the task. In such a go/no-go task, monkeys and humans are probably coding for “animal photographs” and might produce their go responses when the accumulation of information reaches a threshold within one second, the no-go response being the default response. Monkey Ry was extremely good at the task and the shape of its accuracy scores as a function of contrast reduction follows the curves calculated for humans and monkey Eu, but with higher performance. A small discrepancy is observed with the  $N/2$  condition. Tested with all other stimuli with further contrast reduction, monkeys did not show any performance impairment. On the other hand, a mild drop in accuracy was observed in monkeys (mainly with monkey Eu with a 3.6% drop in global accuracy) in Experiment 1 when condition  $N/2$  was tested with various luminance shifts. The monkey could have learnt to deal with  $N/2$  stimuli, but the evolution of performance along the 12 sessions of testing did not show any evidence of learning (Fig. 3). Thus, the explanation might lay in the conditions intermixed with condition  $N/2$ . When intermixed with very degraded stimuli conditions (Experiment 2), monkeys could increase their attentional state and condition  $N$  and  $N/2$ , the two easiest conditions, might benefit from this higher attention.

Whereas experiment 1 showed that neither monkeys nor humans rely on global luminance to perform the task, the results of experiment 2 suggest that they do not rely on the fine definition of the stimuli. Indeed, both monkeys and humans can do the task with a 70–75% global accuracy with  $N/10$  stimuli, a condition in which the number of gray levels used to define the image is around 20–26 and the local contrasts so low that they can only activate the magnocellular visual pathway (Macé et al. 2005a). Magnocellular visual information is coarse (low spatial frequency) and achromatic but can be conveyed faster than parvocellular information. Moreover, the magnocellular pathways respond well to transient stimuli, a characteristic that fits well with the flashed stimuli used in the present experiment.

A decision to respond in the fast animal/non-animal superordinate categorization task used here could be based

on coarse stimulus information. We have already shown that low-level cues are not sufficient on their own, because they would otherwise trigger responses at shorter latencies, as shown in human and monkey in our previous studies (Delorme et al. 2004; Macé et al. 2005b). Thus, superordinate categorization may at least require access to rudimentary representations that could be further refined by the additional processing of parvocellular visual messages conveying color and high spatial frequency information (Macé et al. 2005a; Macé et al. 2009). This idea has already been put forward by other authors (Sherman 1985; Bullier 2001) who consider that the parvocellular system could provide higher acuity capacity to a coarse magnocellularly driven form of vision. In fact, when visual categorization is concerned, such a hypothesis is also supported by the fact that categorization of animals at the basic level (dogs, birds) requires more processing time than superordinate animal categorization (Macé et al. 2009), a result at odds with the well-known advantage of basic categories documented by Rosch and collaborators (Rosch et al. 1976) but in agreement with the coarse to fine hypothesis in visual processing (Schyns and Oliva 1994; Bar 2004).

Overall, the results reported here show the efficiency of the primate visual system at categorizing natural photographs in challenging situations such as luminance variations and very low contrasts. Monkeys and humans were compared in an “animal/non-animal” rapid visual categorization task using the same set-up, the same protocol, and the same degraded stimuli. Data showed that performance in both humans and macaques is extremely robust to variation of global luminance and to deterioration of fine stimulus texture. It is clear that their performance is largely independent of the cues studied here. Performance impairments might have been larger with combined degradation of both luminance and contrast, but this seems unlikely as luminance and contrast are not completely independent factors. Thus, neither color, global luminance nor stimulus contrast is particularly critical for the ability of humans and monkeys to detect the presence of an animal in a scene. Monkeys and humans might not rely on such low-level features but rather on animal features of intermediate complexity. The study also shows strong similarities between the performance impairment displayed by humans and monkeys, so that they could both rely on coarse visual information to perform the animal superordinate categorization task. Such coarse transient (magnocellular?) representations might be particularly important, because our task imposes severe time constraints. The ability of monkeys to recognize objects -and animals in particular- in degraded conditions of vision (such as low luminance and contrast) is no doubt of vital importance for their survival. The similar effects of such experimental manipulation that we have found in both humans and monkeys suggest that this ability has been preserved during primate evolution.

**Acknowledgments** This work was supported by the CNRS, the University Paul Sabatier (Toulouse III), the French government and the “Programme Cognitique”. We would like to thank S. Thorpe for helpful comments and corrections.

## References

- Albrecht DG, Geisler WS, Frazor RA, Crane AM (2002) Visual cortex neurons of monkeys and cats: temporal dynamics of the contrast response function. *J Neurophysiol* 88:888–913
- Avidan G, Harel M, Hendler T, Ben-Bashat D, Zohary E, Malach R (2002) Contrast sensitivity in human visual areas and its relationship to object recognition. *J Neurophysiol* 87:3102–3116
- Bar M (2004) Visual objects in context. *Nat Rev Neurosci* 5:617–629
- Bentin S, Allison T, Puce A, Perez E, McCarthy G (1996) Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience* 8:551–565
- Brady TF, Konkle T, Alvarez GA, Oliva A (2008) Visual long-term memory has a massive storage capacity for object details. *Proc Natl Acad Sci U S A* 105:14325–14329
- Bullier J (2001) Integrated model of visual processing. *Brain Res Brain Res Rev* 36:96–107
- D’Amato MR, Van Sant P (1988) The person concept in monkeys (*Cebus apella*). *J Exp Psychol Anim Behav Process* 14:43–55
- Delorme A, Richard G, Fabre-Thorpe M (2000) Ultra-rapid categorization of natural scenes does not rely on colour cues: a study in monkeys and humans. *Vision Res* 40:2187–2200
- Delorme A, Rousselet GA, Macé MJ, Fabre-Thorpe M (2004) Interaction of top-down and bottom-up processing in the fast visual analysis of natural scenes. *Brain Res Cogn Brain Res* 19:103–113
- 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
- Farah MJ, Wilson KD, Drain M, Tanaka JN (1998) What is “special” about face perception? *Psychol Rev* 105:482–498
- Gegenfurtner KR, Rieger J (2000) Sensory and cognitive contributions of color to the recognition of natural scenes. *Curr Biol* 10:805–808
- Girard P, Jouffrais C, Kirchner H (2008) Ultra-rapid categorisation in non-human primates. *Anim Cogn* 11:485–493
- Herrnstein RJ, Loveland DH (1964) Complex Visual Concept in the Pigeon. *Science* 146:549–551
- Huber L, Lenz R (1993) A test of the linear feature model of polymorphous concept discrimination with pigeons. *Q J Exp Psychol* 46B:1–18
- Huber L, Troje NF, Loidolt M, Aust U, Grass D (2000) Natural categorization through multiple feature learning in pigeons. *Q J Exp Psychol B* 53:341–357
- Kaplan E, Shapley RM (1986) The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proc Natl Acad Sci U S A* 83:2755–2757
- Macé MJ, Thorpe SJ, Fabre-Thorpe M (2005a) Rapid categorization of achromatic natural scenes: how robust at very low contrasts? *Eur J Neurosci* 21:2007–2018
- Macé MJ, Richard G, Delorme A, Fabre-Thorpe M (2005b) Rapid categorization of natural scenes in monkeys: target predictability and processing speed. *Neuroreport* 16:349–354
- Macé MJ, Joubert OR, Nespoulous J-L, Fabre-Thorpe M (2009) The time-course of visual categorizations: you spot the animal faster than the bird. *PLoS One* 4(6):e5927 <http://dx.plos.org/10.1371/journal.pone.0005927>
- Martin-Malivel J, Fagot J (2001) Cross-modal integration and conceptual categorization in baboons. *Behav Brain Res* 122:209–213
- Martin-Malivel J, Mangini MC, Fagot J, Biederman I (2006) Do humans and baboons use the same information when categorizing human and baboon faces? *Psychol Sci* 17:599–607
- Minamimoto T, La Camera G, Richmond BJ (2009) Measuring and Modeling the Interaction Among Reward Size, Delay to Reward, and Satiation Level on Motivation in Monkeys. *J Neurophysiol* 101:437–447
- Pong M, Fuchs AF (2000) Characteristics of the pupillary light reflex in the macaque monkey: metrics. *J Neurophysiol* 84:953–963
- 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
- Rolls ET, Baylis GC (1986) Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Exp Brain Res* 65:38–48
- Rosch E, Mervis CB, Gray WD, Johnson DM, Boyes-Braem P (1976) Basic objects in natural categories. *Cogn Psychol* 8:382–439
- Schyns PG, Oliva A (1994) From blobs to boundary edges—evidence for time-scale-dependent and spatial-scale-dependent scene recognition. *Psychological Science* 5:195–200
- Sherman SM (1985) Functional organization of the W-, X- and Y-cell pathways in the cat: a review and hypothesis. In: Sprague JM, Epstein AN (eds) *Progress in psychobiology and physiological psychology*, vol 11. Academic Press, New York, pp 233–314
- Spence I, Wong P, Rusan M, Rastegar N (2006) How color enhances visual memory for natural scenes. *Psychol Sci* 17:1–6
- Strasburger H, Rentschler I (1996) Contrast-dependent dissociation of visual recognition and detection fields. *Eur J Neurosci* 8:1787–1791
- Strasburger H, Harvey LO Jr, Rentschler I (1991) Contrast thresholds for identification of numeric characters in direct and eccentric view. *Percept Psychophys* 49:495–508
- Troje NF, Huber L, Loidolt M, Aust U, Fieder M (1999) Categorical learning in pigeons: the role of texture and shape in complex static stimuli. *Vision Res* 39:353–366
- Vogels R (1999) Categorization of complex visual images by rhesus monkeys. Part 1: behavioural study. *Eur J Neurosci* 11:1223–1238
- Von Fersen L, Lea SEG (1990) Category discrimination by pigeons using five polymorphous features. *J Exp Anal Behav* 54:69–84
- Wichmann FA, Sharpe LT, Gegenfurtner KR (2002) The contributions of color to recognition memory for natural scenes. *J Exp Psychol Learn Mem Cogn* 28:509–520
- Winn B, Whitaker D, Elliott DB, Phillips NJ (1994) Factors affecting light-adapted pupil size in normal human subjects. *Invest Ophthalmol Vis Sci* 35:1132–1137
- Yao AY, Einhauser W (2008) Color aids late but not early stages of rapid natural scene recognition. *J Vis* 8(12):11–13