Speed of processing in the human visual system

Simon Thorpe, Denis Fize & Catherine Marlot

Centre de Recherche Cerveau & Cognition, UMR 5549, 31062 Toulouse, France

How long does it take for the human visual system to process a complex natural image? Subjectively, recognition of familiar objects and scenes appears to be virtually instantaneous, but measuring this processing time experimentally has proved difficult. Behavioural measures such as reaction times can be used[1], but these include not only visual processing but also the time required for response execution. However, event-related potentials (ERPs) can sometimes reveal signs of neural processing well before the motor output[2]. Here we use a go/no-go categorization task in which subjects have to decide whether a previously unseen photograph, flashed on for just 20 ms, contains an animal. ERP analysis revealed a frontal negativity specific to no-go trials that develops roughly 150 ms after stimulus onset. We conclude that the visual processing needed to perform this highly demanding task can be achieved in under 150 ms.

Neurophysiological measurements of the latencies of selective visual responses can be used to provide estimates of visual processing time[3]. For example, it is known that higher-order visual areas such as the primate superior temporal sulcus contain neurons that can respond selectively to faces with latencies of ~ 100 ms[4–6]. In humans, face-selective evoked potentials have been demonstrated using both surface ERP recordings[7,8] and implanted intracerebral electrodes[9–11]. Such potentials typically peak at ~ 200 ms after stimulus onset, but may start as early as 140 ms. It is unclear, however, whether such latencies are typical of visual processing in general. One problem is that face processing may involve highly specialized and optimized neural pathways, and although there have been a few reports of early differential responses to other stimuli, including words[10,12,13] and line drawings[14], no previous ERP studies have attempted to measure processing times for more natural scenes. A second problem is that the existence of a short latency differential response does not imply that visual processing has been completed—responses to faces, for example, could correspond to an early processing stage such as 'structural encoding'. One can tackle this problem by using a task that requires the subject to make some sort of categorical judgment about the stimulus, an approach that has been used in a variety of studies using not only photographs of faces[14–17], but also both line drawings[18,19] and photographs[20] of everyday objects. However, differential effects in such tasks occur at considerably longer latencies, typically involving the N400 component of the ERP.

The present study used a task that provides a very serious challenge to the processing capacities of the human visual system. Subjects performed a go/no-go categorization in which they had to decide on the basis of a 20-ms presentation whether an image contained an animal or not. Earlier studies using rapid sequential visual presentation (RSVP) had shown that subjects can detect photographs of animals in a string of images at high presentation rates[21], but this is the first time such a scene categorization task has been performed using ERPs. We used a set of over 4,000 commercially available colour photographs, of which roughly half were used as targets and included a wide range of animals in their natural environments (mammals, birds, reptiles, fish); the remainder were distractors that included pictures of forests, mountains and lakes, as well as buildings, flowers and fruit. As in a number of other studies[8,9,11,16,19], each stimulus was only ever seen once, thus eliminating the possibility of stimulus-specific learning effects.

Despite the very high demands made on the visual system by such a task (the subjects had no *a priori* information about the type of animal to look for, its position or size, or even the number of animals present), performance was remarkably good. The average proportion of correct responses was 94%, with one of the fifteen subjects achieving 98% correct responses. The median reaction times on 'go' trials was 445 ms, although this value varied considerably between subjects, from a minimum of 382 ms to as much as 567 ms (Fig. 1). This remarkable level of performance was possible despite the very brief presentations, which effectively rule out the use of eye movements during image processing.

Whereas the behavioural reaction times put an upper limit on the time required for visual processing, the analysis of event-related potentials provided a much stronger constraint. By comparing average brain potentials generated on correct 'go' trials with those generated on correct 'no-go' trials, we were able to demonstrate that the two potentials diverge very sharply at ~ 150 ms after stimulus onset. The effect was particularly clear at frontal recording sites, and was characterized by a nearly linear increase in the voltage difference over the following 50 ms or so, the potential being more negative on no-go trials (Fig. 2). All 15 subjects showed the effect (Fig. 3), and although the onset latency varied somewhat between subjects, the differences were very minor compared with the very large differences in behavioural reaction times. Furthermore, there was no correlation whatsoever between behavioural reaction time and the onset latency for the differential response. This makes it unlikely that the differential

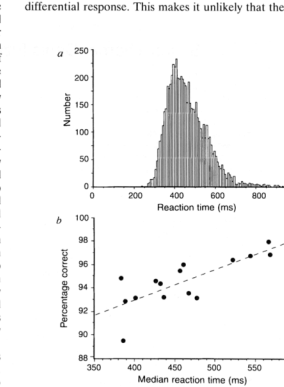

FIG. 1 Behavioural performance measures. *a*, Distribution of reaction times on target trials for the 15 subjects participating in the study (7 males and 8 females, aged between 22 and 45 years of age), each of whom performed at least 700 trials (range, 700–2,000). Subjects pressed a button to start a sequence of trials and, after an interval of 1 to 2 seconds, an image was presented at the centre of the screen. A small fixation cross was present before and after stimulus presentation. Subjects were instructed to release the button if they saw an animal ('go' trials), and to keep their finger on the button otherwise ('no-go' trials). Target and distractor trials were presented at random, with roughly equal probability in blocks of 100 trials. All the images were 384 by 256 pixels in size and were presented for 20 ms (two frames at 100 Hz) using a Cambridge Vision Research VSG 2/2 graphics board mounted in a PC compatible computer. *b*, Accuracy as a function of median reaction time for each of the 15 subjects. The dashed line plots a linear regression ($r = 0.623$) and indicates the presence of a speed–accuracy trade-off.

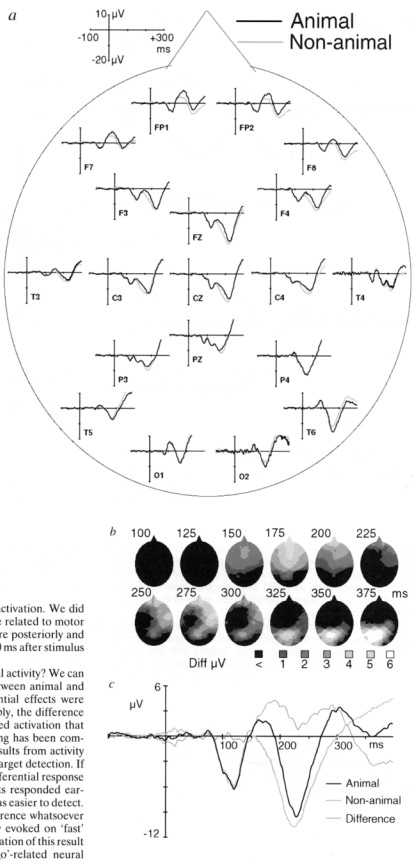

FIG. 2 Evoked potential data for one subject. *a*, Event-related potentials plotted for a 400-ms period starting 100 ms before stimulus onset. Solid lines plot the average response on correct target trials; grey lines plot averages for correct distractor trials. Data were obtained from one subject during 4 separate sessions (a total of 1,580 trials). The two averages overlap until ~150 ms, at which point a clear difference emerges between the potentials on animal and non-animal trials which was clearest at frontal recording sites. Recordings were made using a 20-electrode Electrocap bonnet in the 10–20 configuration connected to a Neuroscan SynAmps system sampling at 1,000 Hz. Signals were digitally filtered using a low pass filter with a cut-off frequency of 100 Hz and a notch filter to remove 50 Hz mains interference. Potentials on individual trails were baseline corrected on the basis of the 100 ms preceding stimulus onset, and any trials contaminated by eye movement artefacts were excluded from the analysis. Data analysis was performed using the SCAN software suite. *b*, Two-dimensional plots of the difference between correct 'go' and 'no-go' trials averaged over 25-ms time slices, starting at 100 ms. No difference can be seen in the first two slices, but from 150 ms onwards a clear difference appears which affects all the frontal electrodes. At around 325 ms, there are clear signs of a lateralized difference at more posterior sites which probably reflects motor activity in this right-handed subject. *c*, Event-related potentials on target (thick black lines) and distractor (grey) trials calculated by averaging the responses for all seven frontal electrodes (FP1, FP2, F3, F4, F7, F8 and FZ) for this one subject. The difference curve (thin black line) demonstrates the sharp onset of the differential response which is more negative on no-go trials. Using the statistical procedure proposed by Rugg *et al.* for determining the onset of the differential effect[28] (at least 15 consecutive *t*-test values exceeding the 0.05 level of significance) the first significant effect was seen for electrode F8 at 152 ms ($d.f. = 1,578$). All seven electrodes reached significance by 157 ms ($2.20 < t < 3.74$), and the level of significance increased monotonically to reach a peak at 186 ms (mean *t*-score for the seven electrodes, 6.72).

effect seen at frontal sites is related to motor activation. We did see lateralized differential activity that could be related to motor preparation (Fig. 2*b*) but this was localized more posteriorly and did not occur until considerably later (300 to 400 ms after stimulus onset).

What could be producing the early differential activity? We can rule out some simple systematic difference between animal and non-animal images because the same differential effects were seen for a very wide range of images. Presumably, the difference must be related to some sort of decision-related activation that occurs only once the necessary visual processing has been completed. One possibility is that the difference results from activity that is specific to 'go' trials, perhaps related to target detection. If target detection was indeed critical, then the differential response should start earlier on trials where the subjects responded earlier—that is, on trials where the target animal was easier to detect. However, Fig. 3*c* shows that there was no difference whatsoever between the latency of the differential activity evoked on 'fast' trials and 'slow' trials. The most plausible explanation of this result is that the difference is not generated by 'go'-related neural activity, but rather by neural activity that is specifically generated on 'no-go' trials. This could reflect a role for frontal areas in inhibiting inappropriate behavioural responses, an interpretation supported by a number of earlier studies that reported frontal

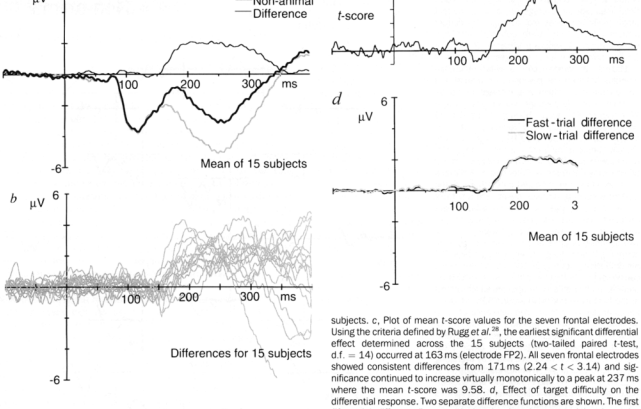

FIG. 3 Event-related potentials for 15 subjects. *a*, As Fig. 2*c*, but averaged over all 15 subjects. *b*, Average difference curves for the seven frontal electrodes plotted separately for each of the 15 subjects. Note that all subjects show a similar difference function, more negative on 'no-go' trials, and that the onset of the differential response is relatively constant across subjects. *c*, Plot of mean *t*-score values for the seven frontal electrodes. Using the criteria defined by Rugg *et al.*[28], the earliest significant differential effect determined across the 15 subjects (two-tailed paired *t*-test, d.f. = 14) occurred at 163 ms (electrode FP2). All seven frontal electrodes showed consistent differences from 171 ms ($2.24 < t < 3.14$) and significance continued to increase virtually monotonically to a peak at 237 ms where the mean *t*-score was 9.58. *d*, Effect of target difficulty on the differential response. Two separate difference functions are shown. The first ('fast-trial difference') was calculated using those go trials where the subjects reaction time was faster than the median, whereas the second ('slow-trial difference') used those trials where the subject was slower. The fact that the two curves overlap virtually perfectly indicates that the differential response is probably the result of activity specific to 'no-go' trials.

activity specific to 'no-go' trials at around the same latency, but where the visual processing requirements were much less demanding than in this study[22,23]. It is clear, however, that additional work using techniques such as fMRI and/or source analysis will be needed to determine the precise structures involved in generating the differential response.

The presence of 'no-go' specific activity at frontal recording sites at 150 ms implies that a great deal of visual processing must have been completed before this time. Indeed, although activity related to target detection could be explained relatively easily (the presence of an eye or feathers would be enough to decide that an animal is present), 'no-go' specific activity implies that the visual system has already performed enough processing to conclude that no animal is present anywhere in the image. It therefore seems clear that the very rapid processing seen previously in the case of faces also occurs in the case of much more complex scene analysis. Quite how the human visual system achieves such a phenomenal amount of computation in such a short time is clearly a challenge for current theories of object vision[24–26], but given the large number of processing stages involved in primate visual system, it seems likely that much of this processing must be based on essentially feed-forward mechanisms[3,4,27]. □

Received 21 November 1995; accepted 22 March 1996.

1. Potter, M. C. & Faulconer, B. A. *Nature* **253,** 437–438 (1975).
2. Rugg, M. D. & Coles, M. G. H. *Electrophysiology of Mind: Event-related Brain Potentials and Cognition* (Oxford University Press, Oxford, 1995).
3. Thorpe, S. J. & Imbert, M. in *Connectionism in Perspective* (eds Pfeifer, R., Schreter, Z., Fogelman-Soulié, F. & Steels, L.) 63–92 (Elsevier, Amsterdam, 1989).
4. Oram, M. W. & Perrett, D. I. *J. Neurophysiol.* **68,** 70–84 (1992).
5. Perrett, D. I., Rolls, E. T. & Caan, W. *Exp. Brain Res.* **47,** 329–342 (1982).
6. Rolls, E. T. & Tovee, M. *J. Proc. R. Soc. Lond. B* **257,** 9–15 (1994).
7. Jeffreys, D. A. & Tukmachi, E. S. *Exp. Brain Res.* **91,** 340–350 (1992).
8. Seeck, M. & Grüsser, O. *J. Exp. Brain Res.* **92,** 338–349 (1992).
9. Halgren, E., Baudena, P., Heit, G., Clarke, M., Marinkovic, K. & Clark, M. *J. Physiol., Paris* **88,** 1–50 (1994)
10. Allison, T., McCarthy, G., Nobre, A., Puce, A. & Belger, A. *Cereb. Cortex* **4,** 544–554 (1994).
11. Allison, T. *et al. J. Neurophysiol.* **71,** 821–825 (1994).
12. Nobre, A. C., Allison, T. & McCarthy, G. *Nature* **372,** 260–263 (1994).
13. Rudell, A. P. *EEG. Clin. Neurophysiol.* **83,** 77–82 (1992).
14. Barrett, S. E., Rugg, M. D. & Perrett, D. I. *Neuropsychologia* **26,** 105–117 (1988).
15. Barrett, S. E. & Rugg, M. D. *Neuropsychologia* **27,** 913–922 (1989).
16. Holcomb, P. J. & MacPherson, W. B. *Brain Cogn* **24,** 259–276 (1994).
17. Bobes, M. A., Valdes Sosa, M. & Olivares, E. *Brain Cogn.* **26,** 1–22 (1994).
18. Barrett, S. E. & Rugg, M. D. *Brain Lang.* **38,** 424–437 (1990).
19. Barrett, S. E. & Rugg, M. D. *Brain Cogn* **14,** 201–212 (1990).
20. Pratarelli, M. E. *Brain Cogn* **24,** 137–157 (1994).
21. Intraub, H. *J. Exp. Psychol. Hum. Perc. Perf.* **7,** 604–610 (1981).
22. Gemba, H. & Sasaki, K. *Neurosci. Lett.* **101,** 263–268 (1989).
23. Sasaki, K., Gemba, H., Nambu, A. & Matsuzaki, R. *Neurosci. Res.* **18,** 249–252 (1993).
24. Vetter, T., Hurlbert, A. & Poggio, T. *Cereb. Cortex* **5,** 261–269 (1995).
25. Bülthoff, H. H., Edelman, S. Y. & Tarr, M. *Cereb. Cortex* **5,** 247–260 (1995).
26. Hummel, J. E. & Biederman, I. *Psychol. Rev.* **99,** 480–517 (1992).
27. Celebrini, S., Thorpe, S., Trotter, Y. & Imbert, M. *Vis. Neurosci.* **10,** 811–825 (1993).
28. Rugg, M. D., Doyle, M. C. & Wells, T. *J. cog. Neurosci.* **7,** 209–227 (1995).

ACKNOWLEDGEMENTS. This work was supported by the Human Capital and Mobility program, the CNRS and the Conseil Regional Midi-Pyrenees.

CORRESPONDENCE and requests for materials should be addressed to S.T. (e-mail: thorpe@cerco.ups-tlse.fr).