Concurrent processing reveals competition between visual representations of faces

Corentin Jacques[CA] and Bruno Rossion

Unité de Neurosciences Cognitives et Laboratoire de Neurophysiologie, Université Catholique de Louvain, 10 Place du Cardinal Mercier, 1348 Louvain-la-Neuve, Belgium

[CA]Corresponding Author: corentin.jacques@psp.ucl.ac.be

Scalp electrophysiological recordings in humans indicate that the processing of faces differs from other categories between 100 and 200 ms after stimulus onset, peaking at the N170. To test the effect of the addition of a second face stimulus in the visual field on the face-related N170, we measured this component in response to a laterally presented face stimulus while subjects were processing either a central face or a control stimulus. As early as 130 ms following the lateralized face stimulus, there was a strong (~40% of signal) and specific reduction of the N170 amplitude when subjects were concurrently processing a central face. This observation suggests that the early stages of the N170 reflect the activation of individual faces having overlapping and competing neural representations in the human occipito-temporal cortex. *NeuroReport* 15:2417–2421 © 2004 Lippincott Williams & Wilkins.

Key words: Competition; Concurrent visual processing; Event-related potentials; Face processing; N170; Neural representation of faces

INTRODUCTION

The processing of individual faces in the human brain is thought to be both efficient and very fast, as a result of evolutionary constraints and/or of our visual expertise with this category [1]. The speed and the time course of facial categorization processes can be inferred from scalp event-related potentials (ERPs) in normal humans, which provide millisecond time resolution of the electrophysiological processes taking place at the system level. Face processing is associated with a large visual potential, the N170, peaking between 140 and 180 ms over bilateral occipito-temporal regions of the scalp. The N170 is the earliest and only consistent larger response for faces as compared to multiple other non-face object categories [2–5]. Evidence from source localization [3,5], intracranial recordings [6] and the combination of EEG and fMRI data [7] suggest that the N170 originates from a network of regions including the middle fusiform gyrus, the inferior occipital cortex, and the inferior, middle and superior temporal gyri. These localizations largely overlap with regions where face-sensitive activations have been described in neuroimaging studies (see [8] for a review) or in single-cell recordings in monkeys [9].

In monkeys, cells that respond more to facial patterns (either whole faces or face parts) than to a wide variety of other complex stimuli have been found in the infero-temporal cortex (IT) and in the superior temporal sulcus (STS). Although a face-selective neuron is sharply tuned to a particular class of stimuli (faces), it is not tuned to a single face, but appears to have a different response (i.e. spike rate) to different members of a set of faces [10,11]. Thus, evidence from single-cells recordings in monkeys suggest that individual faces are coded by overlapping sub-populations of neurons, each neuron's response reflecting a coding element of a distributed facial representation [11].

If the face-N170 recorded from the scalp in humans largely originates from the activity of overlapping populations of neurons encoding different individual faces, we reasoned that it should be massively reduced in amplitude in response to a face stimulus when another face stimulus is being processed concurrently. More precisely, if two faces are presented to the visual system concurrently, the N170 to the second face stimulation should be reduced in amplitude, because the neurons involved in processing the first face would be continuously at work, and would thus be only partially available to encode the second face.

To test this hypothesis, the N170 response to a lateralized face was recorded while subjects were visually processing a centrally presented stimulus. The central stimulus could either be a face or a phase-scrambled face that preserved the low-level visual properties of the face image. In agreement with the hypotheses, the N170 to the lateralized face stimulus was specifically and markedly reduced as early as 125 ms when it was presented concurrently with another central face stimulus, suggesting that individual faces have overlapping and competing neural representations in the human occipito-temporal cortex activated as early as when faces are discriminated from other categories.
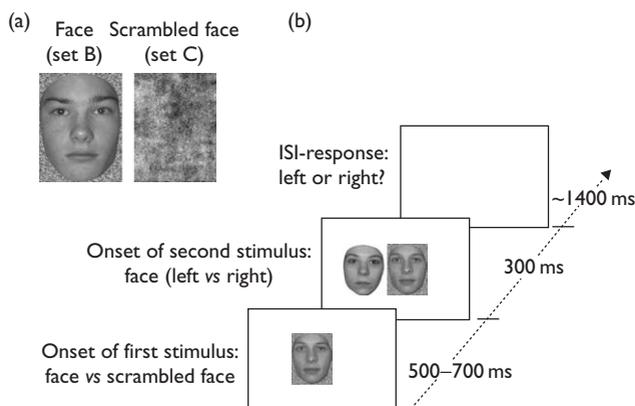
MATERIALS AND METHODS

Subjects: Fourteen paid volunteers (six females, one left-handed, mean age 23.3 ± 4.8 years) participated in this study. One subject was rejected for poor EEG signal-to-noise ratio.

Stimuli: A set of 36 colored photographs of full-front faces (set A) without external features were used. An additional set (B) was composed of the faces from set A embedded in a rectangle-shaped colored white noise (Fig. 1a). A second additional set (C) was built by scrambling the faces from set B using a phase randomization procedure (see [12] for details). The phase randomization procedure yields images that preserve the global low-level properties of the original image (e.g. luminance, contrast, spectral energy), while blurring completely category-related information (Fig. 1a). Stimuli from set B were embedded in colored white noise so that they subtended equal shape, size and contrast against background as the phase scrambled faces (set C). All stimuli subtended $\sim 2.8 \times 3.7°$ of visual angle.

Procedure: Subjects were instructed to fixate the centre of the monitor (viewing distance: 100 cm) during the presentation of 4 blocks of 72 trials. In each trial, a first stimulus (face or phase-scrambled face) appeared centrally for ~ 900 ms. Approximately 600 ms (randomized between 500 and 700 ms) after the onset of the first stimulus, a face stimulus (set A) was presented for 300 ms either on the left (half of the trials) or on the right side of the first stimulus (Fig. 1). The centre of the second stimulus was located 3.1° away from the centre of the screen. The offset of the two stimuli was simultaneous and was followed by a blank screen for ~ 1400 ms (randomized between 1200 and 1600 ms). Half of the 288 trials were made of a face + face sequence, and the other half of a phase-scrambled face + face sequence. In a given face + face sequence, identity of the two faces was always different. There were 72 trials/condition, presented in random order. The subject's task was to press one of two response keys corresponding to the side of the second stimulation, as accurately and as fast as possible. Subjects were instructed to maintain eye gaze fixation to the centre of the screen during the whole trial.

EEG recordings and ERP analyses: EEG was recorded from 58 tin electrodes, mounted in an electrode cap (Quick cap, Neuromedical Supplies, Inc.) adapted from the 10-20 standard montage. Vertical and horizontal eye movements were monitored using four additional electrodes placed above and below the left eye and on the outer canthus of each eye. During EEG recording, all electrodes were referenced to the left earlobe and electrode impedances were kept below 10 kΩ. EEG was digitized at a 1024 Hz sampling rate. After 30 Hz low-pass filtering of the EEG, artifacts were removed using a $[-35; +35 \mu V]$ standard deviation over 200 ms intervals on all electrodes, and blink artifacts were corrected. For each subject, for correct trials only, averaged epochs centered on the second stimulus were computed and baseline corrected using the 200 ms before stimulus onset. Subjects' averages were re-referenced to a common average reference and filtered using a 1 Hz high-pass filter in order to cut-off slow anticipatory waves which may be elicited before the onset of the second stimulus.

Statistical analyses: Mean N170 amplitude measured within 30 ms time windows centered around peak latency (contralateral hemisphere (145–175 ms), and ipsilateral hemisphere (175–205 ms); Fig. 2, Table 1) at two occipito-temporal electrodes[1] (PO7, PO8) were submitted to repeated-measures ANOVAs with preceding stimulus (face *vs* scrambled face), visual field (right *vs* left), and hemisphere (right *vs* left) as factors. In addition, to define precisely the onset time and duration of ERP differences between conditions, point-by-point *t*-tests (df=12) were conducted at $p < 0.05$ between 0 and 800 ms on two ERP difference waveforms for the same pair of occipito-temporal electrodes (Fig. 2) [13].

RESULTS

Behavioral results: Subjects performed the task almost at ceiling in all conditions ($\sim 99\%$ on average) with mean reaction times ranging from 320 to 335 ms. There were no significant differences between conditions for accuracy and reaction times (all $p > 0.1$).

ERP results: The N170 elicited by the first stimulus, either faces or scrambled faces, peaked at the same latency (~ 155 ms). However, faces evoked a much larger N170 than scrambled faces in both hemispheres (faces $-5.57 \mu V$ and $-5.50 \mu V$; scrambled faces $+2.93 \mu V$ and $+6.26 \mu V$ for the left and right hemisphere respectively).

On average, the N170 elicited by the second stimulus (lateralized face) peaked around 160 ms when measured in the contralateral hemisphere, and was delayed by about 20–30 ms (~ 190 ms) when measured in the ipsilateral hemisphere (Fig. 2; Table 1), in agreement with previous observations [14]. Most importantly, the N170 was substantially reduced in amplitude when the face stimulus was presented concurrently with another face stimulus relative to when it was presented in the context of a scrambled face (Fig. 2, Fig. 3; Table 1).

Statistical analyses conducted on mean N170 amplitude confirmed these observations. The main effect of preceding stimulus (F(1,12)=44.91, p=0.000022), reflected the N170 amplitude reduction when the face was presented concurrently with another face. This effect was qualified by a

Fig. I. Stimuli and procedure used in this study. (**a**) Examples of faces from set B, and of phase scrambled faces from set C. (**b**) Time line of the stimulation sequence for a face + face sequence.

---

[1]Identical results were obtained when conducting these analyses on a pool of 5 pairs of occipito-temporal electrodes (see scalp topographies, Fig. 3).
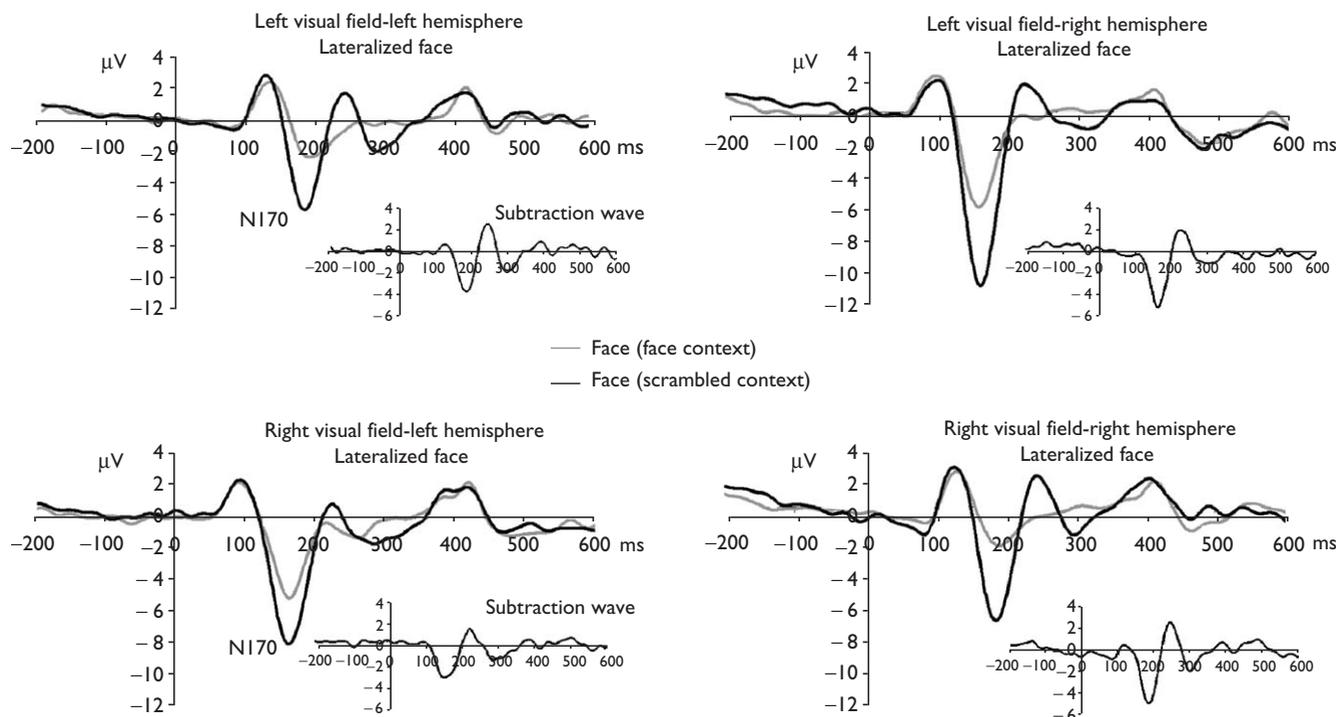
Fig. 2. Grand averaged ERPs elicited by the second stimulus of the trial sequence (lateralized face) for two occipito-temporal electrodes (PO7, PO8). Below, in black: difference waveforms obtained for each visual field and hemisphere separately by subtracting ERPs for faces presented in a face context to ERPs for faces presented in a scrambled face context (black curve minus grey curve).

Table I. Mean latencies and amplitudes of the NI70 recorded in response to the second stimulus.

|		Face context		Scrambled face context	
		Left hemisphere	Right hemisphere	Left hemisphere	Right hemisphere
LVF stimulation	Latencies (ms)	I92	I6I	I85	I63
	Amplitudes (µV)	−2.68	−5.99	−6.23	−II.05
RVF stimulation	Latencies (ms)	I62	I92	I6I	I88
	Amplitudes (µV)	−5.15	−2.54	−8.12	−7.I3

LVF: left visual field; RVF: right visual field.

significant interaction between preceding stimulus and hemisphere (F(1,12)=13.52, p=0.0032), the reduction effect being larger in the right hemisphere (4.83 µV, F(1,12)=54.81, p<0.00001), although it was also significant in the left hemisphere (3.26 µV, F(1,12)=26.69, p=0.00023). Finally, there was also a significant visual field × hemisphere interaction (F(1,12)=12.52, p=0.0041), due to larger N170 in the contralateral hemisphere.

Differential ERP activities (subtraction waves) related to the presentation of a face preceded either by a scrambled face or by another face are depicted in Fig. 2. When presented in the left visual field, the subtraction wave differed significantly from zero at 125 ms after the onset of the second stimulus in the right hemisphere, and about 20 ms later (144 ms) in the left hemisphere. Following right visual field face stimulation, the ERPs differed at 125 ms in the left hemisphere and 30 ms later (155 ms) in the right hemisphere. On average, the effect remained significant for 70 ms and 61 ms in the contralateral and ipsilateral hemi-

sphere respectively. In all conditions, the latency of the maximal difference corresponded to the latency of the N170 (Fig. 2; Table 1). Approximately 20 ms after the end of the first significant difference, a second significant difference between conditions occurred, first in the contralateral hemisphere (∼210 ms), and then in the ipsilateral hemisphere (∼225 ms, Fig. 2). On average, the difference remained significant for about 40 ms. Finally, a third significant difference appeared ∼275 ms after the onset of the second stimulus in all visual fields and hemispheres, and persisted until about 340 ms.

In sum, the N170 to lateralized faces undergoes a major amplitude reduction when a face stimulus is presented while another face is being processed, relative to when a scrambled face is processed. The onset latencies of the difference between conditions, as reflected by ERP differential activities, occurred around 125 ms in the contralateral hemisphere and ∼25 ms later (∼150 ms) in the ipsilateral hemisphere.
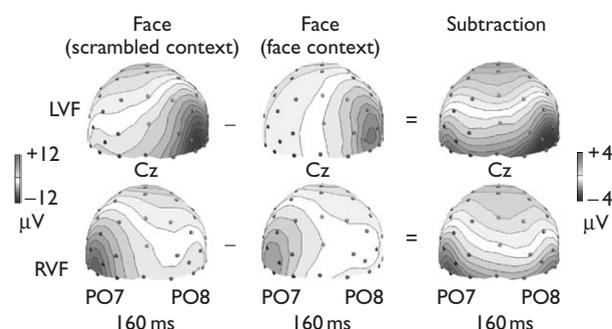
Fig. 3. Scalp topographies depicting the distribution of the Nl70 at 160 ms following the onset of the second stimulus (lateralized face) for the left visual field (LVF - top) and the right visual field (RVF - bottom) stimulation. Left: response to faces presented in a scrambled face context. Middle: response to faces presented in a face context. Right: topography of the differential activity obtained by subtracting the response elicited by faces in a face context to the response evoked by faces in a scrambled-face context.

DISCUSSION

Scalp ERPs were used to examine how different facial representations, putatively coded by overlapping neuronal populations, interact over time. If different faces have overlapping neural representations in the occipito-temporal cortex, we hypothesized that there should be a reduction in signal amplitude as early as the N170 in response to a face, if another face picture was already present in the visual field. In agreement with these hypotheses, we observed that the N170 amplitude to a lateralized face stimulus was largely reduced (40–50% of signal) when it was presented in the context of a central face stimulus. It is worth noting that this large modulation of the N170 contrasts with previous studies using upright face stimuli and in which no or few modulations of the N170 as a result of task manipulations or selective attention were reported [15,16].

The results obtained here suggest that when two individual faces are presented in the visual field, they compete for overlapping neural representations. Sample-by-sample analyses performed on ERP difference waveforms revealed that this competition effect starts around 125 ms, a timing compatible with the activation of high-level visual areas in the ventral stream [17] and with the onset of spiking activity in face-selective cells in the monkey's IT and STS (around 80–130 ms) [9]. Hence, the competition effect found here for concurrently presented faces is likely to arise from regions in the cortex where face-specific activities are found, that is where faces and non-face objects are functionally segregated. Several arguments support these assumptions. First, the onset of the N170 amplitude difference between faces and other objects reported in previous ERP experiments (110–130 ms) [2–5] has a similar latency as the onset of the competition effect between individual faces reported here. Second, the difference waves (Fig. 2, Fig. 3) reflecting the competition effect have a very similar distribution on the scalp as a classical N170 response to faces, suggesting that the locus of the competition effect largely lies in the occipito-temporal regions that participate in generating the N170. Finally, both the N170 amplitude advantage for faces over objects [2–5] and the competition effect between concurrently presented faces are larger in the right hemisphere, in agreement with the prominent role of this hemisphere in face processing [5,8,18].

The results of the present experiment can be related to several single-cell recording studies in monkeys, which reported that a neuron's response to an optimal stimulus is significantly reduced when another stimulus is presented simultaneously in the cell's receptive field [19,20]. However, the mechanisms involved in response suppression at the single-cell level may be different than those implicated in the N170 reduction observed in the present study given that single-cell recordings and ERPs have distinct physiological bases. Single-unit micro-electrodes record spike trains at the level of the neuron's axon, whereas ERPs are large field potentials generated mostly by post-synaptic depolarization generated along the apical dendrites of cortical pyramidal cells oriented perpendicularly to the scalp. Furthermore, in monkeys, competitive suppression between simultaneously presented stimuli is not present from the very start of the neurons' responses and usually takes tens of milliseconds to build up [21]. Thus, given that scalp ERPs reflect a transient afferent response, it is unlikely that one would observe any modulation of the N170 as a result of concurrent face stimulus presentation with simultaneous onset times. Along the same lines, a series of fMRI studies performed on human subjects have reported a reduction of BOLD signal in several extrastriate visual areas of the ventral stream (V2, V4, TEO, TE) as a result of simultaneously presented compared to sequentially presented complex shapes, providing evidence for competitive suppression in the human visual system [22].

What may be the physiological mechanisms involved in generating the scalp ERP competition effect observed here? We believe that there may be at least two mechanisms at work. First, the presentation of the central face is associated with a sustained response in sub-populations of face-cells in occipito-temporal cortex. Therefore, because different individual faces have overlapping neural representations [10,11], neurons that would have been massively recruited by the lateralized face are only partially available, causing a dramatic reduction of signal in response to this lateralized face. A second plausible mechanism for the competition effect observed here at the scalp level, is that neurons recruited by the central face have suppressed the response of surrounding neurons coding for other facial identities, by means of local lateral inhibitory connections [10,19,23]. Such inhibitory mechanisms appear to be at work in human extrastriate cortex [23] and are implicated in the formation of neurons' selectivity to a particular class of stimuli [24].

The concurrent processing paradigm described here may also be used to investigate how visual processing of faces and non-face objects interacts over time. Given that faces and other object categories have at least partially distinct neural representations in the occipito-temporal cortex [1–6], the N170 in response to faces should not be substantially reduced when subjects are engaged in processing non-face objects. However, neuroimaging studies have shown that the development of a visual expertise with non-face objects leads to the recruitment of face-related processes in the occipito-temporal cortex [25]. Thus, if visual expertise with non-face object categories recruits early face-related visual processes, then the N170 in response to faces should be reduced in amplitude in the presence of another object for which the observer is an expert.

CONCLUSION

This study provides evidence that the neural representations of individual faces in the human occipito-temporal

cortex largely overlap and interact in a competitive fashion, as early as when faces are distinguished from other categories, i.e. at 130 ms following stimulus onset. The concurrent processing paradigm during EEG recording may be a powerful tool to investigate the time course of the interactions between high-level object and face visual representations.

REFERENCES

1. Bruce V and Young AW. *In the Eye of the Beholder: the Science of Face Perception*. New York: Oxford University Press. 1998.
2. Bentin S, McCarthy G, Perez E, Puce A and Allison T. Electrophysiological studies of face perception in humans. *J Cogn Neurosci* 1996; **8**:551–565.
3. Itier RJ and Taylor MJ. N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebr Cortex* 2004; **14**:132–142.
4. Rossion B, Gauthier I, Tarr MJ, Despland P, Bruyer R, Linotte S *et al.* The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *Neuroreport* 2000; **11**:69–74.
5. Rossion B, Joyce CA, Cottrell GW and Tarr MJ. Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage* 2003; **20**:1609.
6. Allison T, Puce A, Spencer DD and McCarthy G. Electrophysiological studies of human face perception. I: potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebr Cortex* 1999; **9**:415–430.
7. Henson RN, Goshen-Gottstein Y, Ganel T, Otten LJ, Quayle A and Rugg MD. Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebr Cortex* 2003; **13**:793–805.
8. Haxby JV, Hoffman EA and Gobbini MI. The distributed human neural system for face perception. *Trends Cogn Sci* 2000; **4**:223–233.
9. Perrett DI, Hietanen JK, Oram MW and Benson PJ. Organization and functions of cells responsive to faces in the temporal cortex. *Phil Trans R Soc Lond B Biol Sci* 1992; **335**:23–30.
10. Baylis GC, Rolls ET and Leonard CM. Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. *Brain Res* 1985; **342**:91–102.
11. Young MP and Yamane S. Sparse population coding of faces in the inferotemporal cortex. *Science* 1992; **256**:1327–1331.
12. Sadr J and Sinha P. Object recognition and random image structure evolution. *Cogn Sci* 2004; **28**:259–287.
13. Rugg M, Doyle M and Wells T. Word and non-word repetition within- and across-modality: an event-related potential study. *J Cogn Neurosci* 1995; **7**:209–227.
14. Terasaki O and Okazaki M. Transcallosal conduction time measured by visual hemifield stimulation with face images. *Neuroreport* 2002; **13**:97–99.
15. Goffaux V, Jemel B, Jacques C, Rossion B and Schyns PG. ERP evidence for task modulations on face perceptual processing at different spatial scales. *Cogn Sci* 2003; **27**:313–325.
16. Cauquil AS, Edmonds GE and Taylor MJ. Is the face-sensitive N170 the only ERP not affected by selective attention? *Neuroreport* 2000; **11**:2167–2171.
17. Bullier J. Integrated model of visual processing. *Brain Res Rev* 2001; **36**:96–107.
18. Hillger LA and Koenig O. Separable mechanisms in face processing – evidence from hemispheric-specialization. *J Cogn Neurosci* 1991; **3**:42–58.
19. Miller EK, Gochin PM and Gross CG. Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain Res* 1993; **616**:25–29.
20. Rolls ET and Tovee MJ. The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive field. *Exp Brain Res* 1995; **103**:409–420.
21. Keysers C and Perrett DI. Visual masking and RSVP reveal neural competition. *Trends Cogn Sci* 2002; **6**:120–125.
22. Kastner S, De Weerd P, Pinsk MA, Elizondo MI, Desimone R and Ungerleider LG. Modulation of sensory suppression: implications for receptive field sizes in the human visual cortex. *J Neurophysiol* 2001; **86**:1398–1411.
23. Allison T, Puce A and McCarthy G. Category-sensitive excitatory and inhibitory processes in human extrastriate cortex. *J Neurophysiol* 2002; **88**:2864–2868.
24. Wang Y, Fujita I and Murayama Y. Neuronal mechanisms of selectivity for object features revealed by blocking inhibition in inferotemporal cortex. *Nature Neurosci* 2000; **3**:807–813.
25. Gauthier I, Tarr MJ, Anderson AW, Skudlarski P and Gore JC. Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. *Nature Neurosci* 1999; **2**:568–573.