LETTERS

Microstimulation of inferotemporal cortex influences face categorization

Seyed-Reza Afraz[1,2], Roozbeh Kiani[1,2] & Hossein Esteky[1–3]

The inferior temporal cortex (IT) of primates is thought to be the final visual area in the ventral stream of cortical areas responsible for object recognition[1,2]. Consistent with this hypothesis, single IT neurons respond selectively to highly complex visual stimuli such as faces[3–6]. However, a direct causal link between the activity of face-selective neurons and face perception has not been demonstrated. In the present study of macaque monkeys, we artificially activated small clusters of IT neurons by means of electrical microstimulation while the monkeys performed a categorization task, judging whether noisy visual images belonged to 'face' or 'non-face' categories. Here we show that microstimulation of face-selective sites, but not other sites, strongly biased the monkeys' decisions towards the face category. The magnitude of the effect depended upon the degree of face selectivity of the stimulation site, the size of the stimulated cluster of face-selective neurons, and the exact timing of microstimulation. Our results establish a causal relationship between the activity of face-selective neurons and face perception.

We trained two adult macaque monkeys to perform a face/non-face categorization task upon viewing single images from one or the other category that were systematically degraded by varying amounts of visual signal. We chose the various signal levels to create a range of difficulties spanning a psychophysical threshold: categorization was easy in some trials and difficult in others (Fig. 1a). In each trial, the monkey was presented briefly (54 ms) with a face or a non-face image degraded by noise. Subsequently, the monkey was required to make a saccadic eye movement to one of two targets to indicate whether the image was a face or a non-face. Each correct response was rewarded by a drop of juice. For pure noise stimuli (Fig. 1a, '0% visual signal'), the monkey was rewarded randomly with a probability of 0.5.

Our central experimental question was whether electrical microstimulation of clusters of face-selective IT neurons would bias the monkeys' choices towards the face category. Because of its relatively precise temporal and spatial characteristics, microstimulation is a particularly powerful tool for establishing causal relationships between physiologically characterized neurons and behavioural performance[7–10]. Even weak microstimulation pulses excite many neurons simultaneously[11–13]; successful use of extracellular microstimulation therefore relies on structural regularities within the cortex, such as the presence of cortical columns[14,15]. Face-selective neurons are found in relatively large clusters in IT[16–18], making them an optimal target for microstimulation.

In each experimental session, we assessed the face selectivity of multiunit clusters of neurons at regular intervals (minimum steps of 150 μm) through a single electrode penetration in IT cortex. At each recording site, selectivity was determined by presenting a large number of face and non-face images while the monkey passively fixated a small fixation point on the monitor screen. Face/non-face stimulus selectivity of multiunit responses was quantified with the d' index (see Methods). A d' value of zero indicates indistinguishable responses to faces and non-faces. Increasingly positive d' values indicate progressively better selectivity for faces.

After recording from several sites within a track (mean number of recorded sites in each track was 4), the electrode was positioned in between the recorded sites and neural response selectivity was determined again. Altogether, we assessed stimulus selectivity at 348 recording sites in 86 electrode penetrations in two monkeys (46 and 40 in monkeys FR and KH, respectively). We conducted microstimulation experiments at 31 face-selective sites and 55 non-selective sites, while the monkey performed the object categorization task. Selectivity for faces was defined as having a d' value >1.

Microstimulation consisted of bipolar current pulses of 50 μA delivered at 200 Hz (refs 19, 20). The stimulation pulses were biphasic, with the cathodal pulse leading. Each pulse was 0.2 ms in duration with 0.1 ms between the cathodal and anodal phase. Each experiment contained three microstimulation conditions differing in the exact time of stimulation delivery as well as an un-stimulated control condition (Fig. 1b). Stimulating pulses were delivered for 50 ms in one of three time periods following onset of the visual stimulus: 0–50 ms, 50–100 ms or 100–150 ms. The first period was

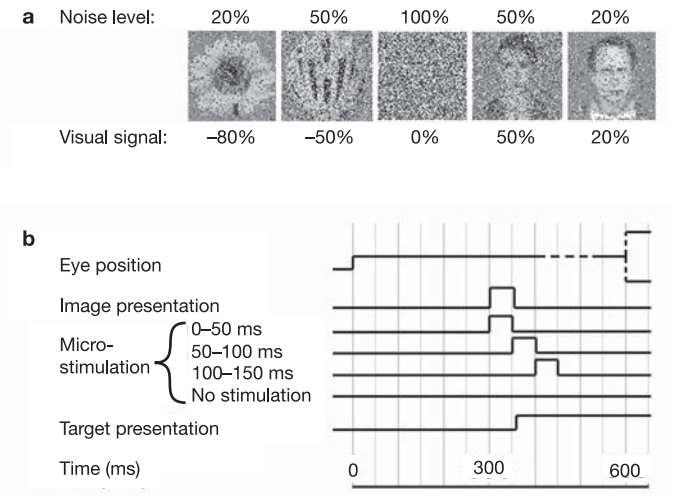

Figure 1 | Visual stimuli and event timing. In each experimental session, the neural stimulus selectivity of several neighbouring cortical sites was first determined in a fixation task using luminance-matched face and non-face greyscale images. Then, in the second part of the experiment (**a**), face and non-face images with varying amounts of noise were used in a face categorization task. **b**, Timing of events in each categorization trial. One of the four possible microstimulation conditions shown was applied randomly in each trial.

[1]School of Cognitive Sciences, Institute for Studies in Theoretical Physics and Mathematics, Tehran, 19395, Iran. [2]Research Center for Brain and Cognitive Sciences, [3]Neuroscience Research Center, Shaheed Beheshti University of Medical Sciences, Tehran, 19385, Iran.

before the earliest visual responses normally observed in IT, and the last two periods correspond to the earliest and later IT responses, respectively[21–23]. The three stimulation conditions and the control trials were randomly interleaved in each experiment.

To reveal the impact of microstimulation on behaviour, the monkey's performance in the categorization task was plotted as the proportion of 'face' choices as a function of the visual stimulus signal for face and non-face images (Fig. 2). We used positive visual signal values for faces and negative values for non-faces to create a continuum. Logistic regression analysis was used to determine whether the three microstimulation time conditions caused a significant shift in the psychometric functions compared to the non-stimulated condition.

Figure 2 illustrates results obtained in two typical microstimulation experiments. In both experiments, microstimulation during the 50–100 ms interval shifted the monkeys' choices significantly towards the face category (Fig. 2a, logistic regression, $P < 0.001$; Fig. 2b, $P < 0.01$). Microstimulation during the 100–150 ms interval biased choices significantly in the experiment of Fig. 2a ($P < 0.001$), but not in the other experimental session depicted in Fig. 2b ($P = 0.283$). (See Supplementary Fig. 1 for more examples of psychometric shifts.) Microstimulation resulted in significant shift of the psychometric function in favour of face choices in at least one of the stimulation conditions for 19 of 31 face selective sites (61%; 9 in right hemisphere and 10 in left hemisphere) and one non-face site with $d' = 0.94$ (Fig. 3). No significant shift in favour of non-face choices was ever observed.

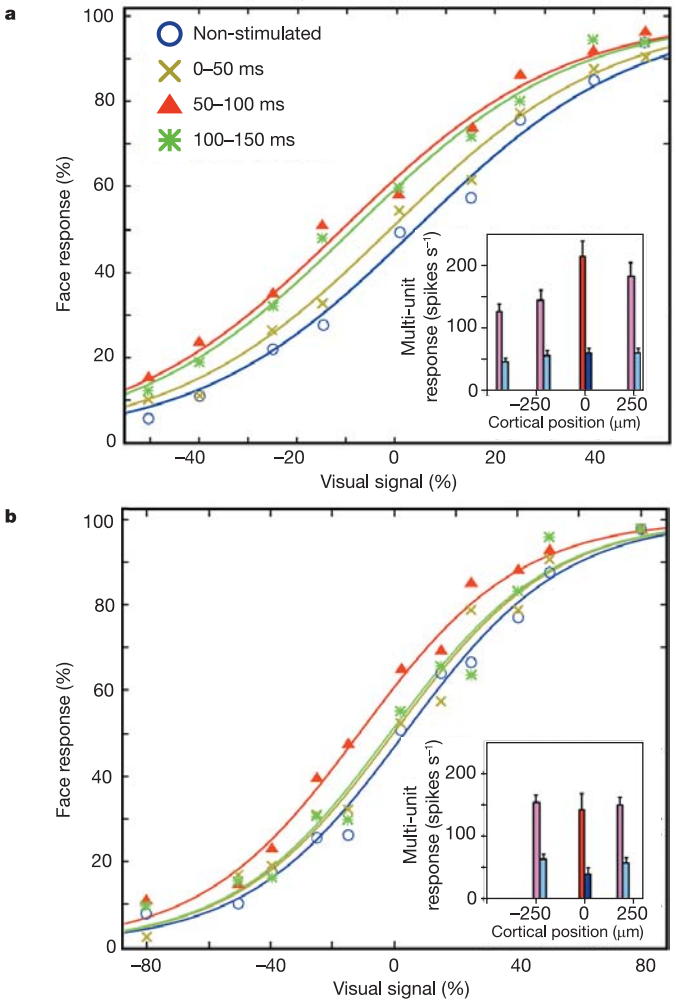

Figure 2 | Effect of microstimulation of two representative face-selective neural clusters in IT cortex. a, Monkey KH; **b**, monkey FR. Data points show the proportion of face choices for different levels of visual signal in the images for different microstimulation conditions. The curves are logistic regression fits to the data points. The insets show averaged multiunit responses of the corresponding stimulated sites and their neighbouring sites. The inset abscissa shows the cortical position of the electrode tip along the recording track relative to the stimulated site (zero on the abscissa). The inset ordinate shows the averaged multi-unit neural responses. Responses to face and non-face stimuli are represented by red and blue bars, respectively. Colours are highlighted for the stimulated site. Error bars, s.e.m.

The impact of microstimulation on perceptual decisions increased as a function of the neural face selectivity of the stimulated site. The scatter plots of Fig. 3a show the correlation between the degree of face selectivity of the stimulated sites and the shift of the psychometric function in different microstimulation conditions. The strongest correlation was found for microstimulation at 50–100 ms after image onset ($r = 0.643$, $P < 0.0001$). Microstimulation at 100–150 ms also showed a significant correlation ($r = 0.539$, $P < 0.0001$). No significant correlation was observed for 0–50 ms ($r = 0.147$, $P = 0.18$).

It has previously been shown that following psychometric function shift the amount of total reward gained by the animal drops in proportion to the amount of shift[7]. Consistently, we observed significant correlations between reward loss and shift value for 50–100 ms and 100–150 ms conditions ($r = 0.42$, $P < 0.001$ and $r = 0.47$, $P < 0.001$, respectively) but not for the 0–50 ms condition ($r = -0.02$, $P = 0.83$).

The impact of microstimulation on perceptual decisions was much larger when current was injected into larger clusters of face-selective neurons. Recall that we measured stimulus selectivity at recording sites adjacent to the stimulation site as well as at the stimulation site itself. When adjacent sites exhibit selectivity similar to that of the recorded site, we may infer that the cluster of physiologically homogeneous neurons is larger, at least along the dimension of our electrode track.

Figure 4a summarizes the effect of stimulating clusters of different sizes. Stimulation effects were substantially more pronounced for larger clusters of face-selective neurons (Fig. 4a; black bars) as compared to smaller clusters (Fig. 4a; grey bars). On average, there was no significant shift in the psychometric function following microstimulation of cortical clusters lacking face selectivity (Fig. 4b). A two-way analysis of variance showed a significant effect of both neighbourhood ($F(1,87) = 11.248$, $P = 0.001$) and stimulation timing ($F(2,87) = 6.092$, $P = 0.003$) on the averaged values of the psychometric function shifts across all face-selective sites (sites with $d' > 1$). No such significant effect was observed in non-selective sites. The averaged d' of neighbouring sites was correlated with the effect of microstimulation: the correlation coefficients for 0–50 ms, 50–100 ms and 100–150 ms stimulation conditions are $r = 0.12$, $P = 0.31$; $r = 0.49$, $P < 0.001$; and $r = 0.44$, $P < 0.001$, respectively.

To prevent the monkeys memorizing specific examples, we used a large image bank, thereby making it unlikely that the monkeys could memorize the specific images. To further examine the unlikely event of whether the monkeys simply memorized all the images in the image set, a behavioural experiment was conducted after the completion of the training in each monkey. In these experiments, 40 novel images (20 faces and 20 non-face objects) were intermixed with 40 familiar face and non-face images (randomly chosen from the learned image bank). The stimuli were presented to the monkey without any visual noise. The face/non-face discrimination performance of the monkeys was measured in several behavioural sessions. In each session, we used a new set of novel images. The monkeys' performance for novel stimuli was as good as it was for familiar stimuli (both above 95%) from the very beginning of the behavioural sessions. Furthermore, to prevent the monkeys using a general rule other than face/non-face categorization, such as detection of living versus non-living objects, we included non-face animate object

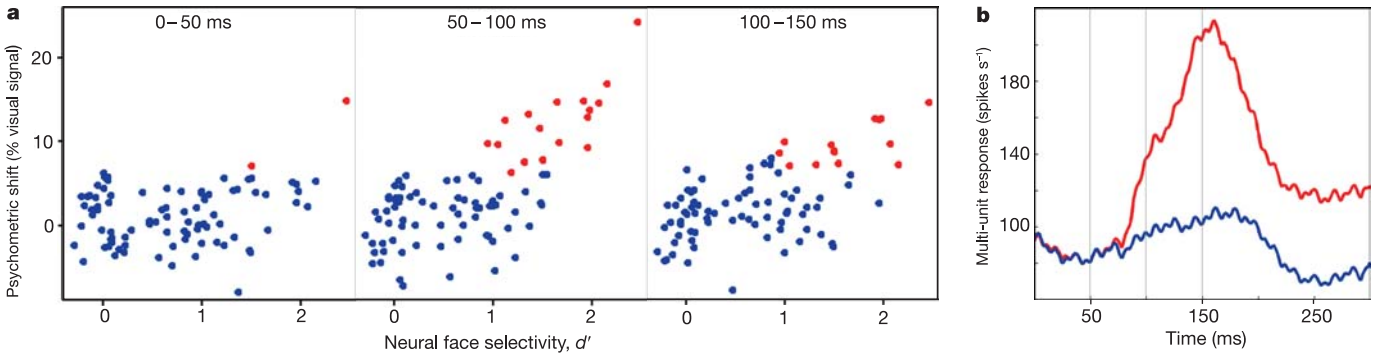

Figure 3 | Correlation between face selectivity of stimulated sites and the behavioural impact of microstimulation. **a**, Positive values on the *y*-axis represent psychometric function shifts in favour of face choices. Large d' values indicate higher selectivity for faces. Red data points indicate statistically significant shifts of the psychometric function. The correlation is significant for microstimulation at 50–100 ms and 100–150 ms conditions, but not for 0–50 ms (see text for details). **b**, Histogram of multiunit neural responses (smoothed with a 20-ms sliding window) to face (red line) and non-face (blue line) stimuli averaged from all stimulated face-selective ($d' > 1$) sites. Different microstimulation time windows are depicted by vertical lines.

images (for example, human and animal bodies from a back view or with the face cut out) in our non-face set and included artefact faces (for example, masks and sculpture faces) in the face set. The possibility of familiarity being a factor was also controlled by using images of human and monkey bodies (which are presumably as familiar objects as faces) in the non-face set.

Our findings demonstrate a causal relationship between IT neural activity and visual object perception and categorization. Although a general role for IT cortex in object perception has been demonstrated previously in cortical ablation experiments[24,25], our data extend causality to a much finer spatial scale. In addition, our data demonstrate that single neuron response properties provide important clues to the functional role of neurons in perception, even for highly complex stimuli such as faces. The functional role of face-selective neurons in behaviour has been debated, but our data clearly show that this role includes, at the very least, categorization of objects into faces and non-faces.

The 50–100 ms stimulation period overlaps with the earliest time window shown to convey information about stimulus category[22,23,26]. The significant effect of microstimulation in this period suggests that downstream cortical areas can use such early information, in addition to the later 100–150 ms neural activity, for decision making.

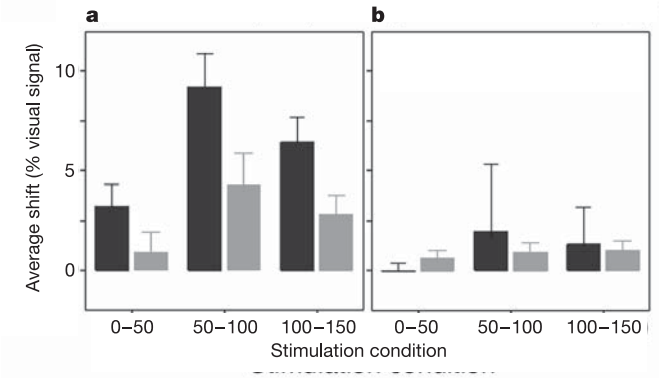

Figure 4 | Effect of stimulus selectivity of neighbouring cortical sites on microstimulation results. Averaged shift in the psychometric function for the three stimulation conditions is shown for all stimulated face-selective (**a**) and non-selective (**b**) sites. Face-selective and non-selective sites were defined by $d' > 1$ and $d' \leq 1$, respectively. Black columns represent sites with face-selective neighbours in their vicinity (±500 μm), and grey columns show sites with non-selective neighbour(s). Positive numbers on the *y*-axis show shifts in favour of face choices. Error bars, s.e.m.

METHODS

Behavioural tasks. Each session started with a passive fixation task in which macaque monkeys (*Macaca mulatta*) were required to maintain fixation in a 4° × 4° window at the centre of the screen. Following 300 ms of fixation, a sequence of visual stimuli (7° × 7° in size) was presented to the monkey. Each image was presented for 200 ms without blank intervals between images[23,27,28], 7–10 times pseudorandomly. The images were greyscale photographs of 30 face objects and 60 non-face objects chosen randomly from a bank of 600 images.

In the second phase of the experiment, monkeys performed a face/non-face categorization task. The monkey started a trial by fixating on the fixation spot for 300 ms. Then a noisy image was presented for 54 ms, followed immediately by two small response targets presented 10° to the left and right of the screen centre. The left and right targets represented face and non-face responses, respectively. The monkey was required to make a saccade to the correct target no later than 660 ms after the onset of targets.

To minimize the monkey's behavioural choice bias, we used a correction scheme[7]. The monkey entered a set of correction trials if it made three consecutive errors within a single category (face or non-face). Upon entering a correction trial, images from the neglected category were presented until the monkey chose that category correctly. All data collected from correction trials were discarded from the analysis. The monkey entered a correction trial in 32 of the 86 sessions, resulting in exclusion of 5.7% of trials in those sessions.

In the categorization task, 11 or 9 signal levels were used in monkey FR and 9 signal levels were used in monkey KH. Each signal level was generated by assigning a uniformly distributed greyscale value to *X*% of image pixels, where *X* is the absolute signal level (positive for faces and negative for non-faces). Noisy face and non-face images create a continuum of task relevant visual signal extending from noiseless faces (100) to full noise images (0) to noiseless non-faces (−100). For each signal level, 16 face and 16 non-face images were randomly selected from the image bank.

Electrophysiology. Recordings were made on an evenly spaced grid, with 1-mm intervals between penetrations over a wide region of the lower bank of STS and TEa cortices[23] (left hemisphere 14 to 21 mm anterior to interauricular line in FR, and right hemisphere 14 to 20 mm anterior to interauricular line in KH). The recording positions were determined stereotaxically by referring to magnetic resonance images acquired before the surgery. Multiunit neural responses were recorded through tungsten microelectrodes (0.4–1.0 MΩ). Neural selectivity of neighbouring sites within ±500 μm from the stimulated site along each recording track was determined as the electrode was advanced. The recorded positions were separated by at least 150 μm (mean, 296 μm). After determining the neighbourhood selectivity, the electrode tip was positioned in the middle of the recorded area and remained there through the rest of the experiment. The neural selectivity in this site was verified before starting the categorization task.

Data analysis. Mean multiunit discharge for each stimulus was measured in a period 70–200 ms after the image onset. The degree of selectivity of each cortical site for face versus non-face images was measured by the d' index[29], based on the following formula:

$$d' = \frac{[M(\mathrm{f}) - M(\mathrm{nf})]}{\sqrt{\frac{\sigma^2(\mathrm{f}) + \sigma^2(\mathrm{nf})}{2}}}$$

where $M(\mathrm{f})$ and $M(\mathrm{nf})$ are the mean multiunit response to face and non-face images, respectively, and $\sigma^2(\mathrm{f})$ and $\sigma^2(\mathrm{nf})$ are the variance of the distributions of neural responses to face and non-face images, respectively.

To calculate the shift in psychometric function, logistic curves were fitted to the monkey's responses in the categorization task, based on the following formula:

$$P(x) = \frac{1}{1 + e^{-(\alpha + \beta x + \lambda_1 I_1 + \lambda_2 I_2 + \lambda_3 I_3)}}$$

where x is the visual signal and $P(x)$ is the probability of face response, I_1, I_2 and I_3 indicate the presence or absence of microstimulation in the three periods (one for stimulation present and zero for stimulation absent condition), and α, β and λ are free parameters (with λ_1 to λ_3 indicating the three stimulation conditions respectively) that were fitted using the maximum likelihood fitting procedure[30]. The fit was performed separately for all of the behavioural data obtained in each experimental session (86 fits). The microstimulation effect in each site was considered significant if λ_i was significantly different from zero ($P < 0.05$). The shift of the psychometric function in each stimulation condition was defined as the change in the visual signal that would have induced a behavioural effect comparable to that of the microstimulation. This is equal to λ_i/β in the logistic fit. Similar methods have been used in other microstimulation studies[7,19]. To reduce the number of the free parameters, our logistic fit assumes a similar slope for the psychometric curves in different stimulation conditions. Allowing different slopes did not improve the fit and was not critical to the results. There was no significant correlation between shift values and slope changes after free fitting of the different conditions (see Supplementary Table 1).

Received 25 January; accepted 16 June 2006.
Published online 26 July 2006.

1. Tanaka, K. Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* **19,** 109–139 (1996).
2. Logothetis, N. K. & Sheinberg, D. L. Visual object recognition. *Annu. Rev. Neurosci.* **19,** 577–621 (1996).
3. Desimone, R., Albright, T. D., Gross, C. G. & Bruce, C. Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* **4,** 2051–2062 (1984).
4. Perrett, D. I., Rolls, E. T. & Caan, W. Visual neurones responsive to faces in the monkey temporal cortex. *Exp. Brain Res.* **47,** 329–342 (1982).
5. Rolls, E. T. Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Hum. Neurobiol.* **3,** 209–222 (1984).
6. Gross, C. G. Processing the facial image: a brief history. *Am. Psychol.* **60,** 755–763 (2005).
7. Salzman, C. D., Murasugi, C. M., Britten, K. H. & Newsome, W. T. Microstimulation in visual area MT: effects on direction discrimination performance. *J. Neurosci.* **12,** 2331–2355 (1992).
8. Romo, R., Hernandez, A., Zainos, A. & Salinas, E. Somatosensory discrimination based on cortical microstimulation. *Nature* **392,** 387–390 (1998).
9. Romo, R. & Salinas, E. Sensing and deciding in the somatosensory system. *Curr. Opin. Neurobiol.* **9,** 487–493 (1999).
10. Cohen, M. R. & Newsome, W. T. What electrical microstimulation has revealed about the neural basis of cognition. *Curr. Opin. Neurobiol.* **14,** 169–177 (2004).
11. Stoney, S. D. Jr, Thompson, W. D. & Asanuma, H. Excitation of pyramidal tract cells by intracortical microstimulation: effective extent of stimulating current. *J. Neurophysiol.* **31,** 659–669 (1968).
12. Asanuma, H., Arnold, A. & Zarzecki, P. Further study on the excitation of pyramidal tract cells by intracortical microstimulation. *Exp. Brain Res.* **26,** 443–461 (1976).
13. Tolias, A. S. *et al.* Mapping cortical activity elicited with electrical microstimulation using FMRI in the macaque. *Neuron* **48,** 901–911 (2005).
14. Nichols, M. J. & Newsome, W. T. Middle temporal visual area microstimulation influences veridical judgments of motion direction. *J. Neurosci.* **22,** 9530–9540 (2002).
15. DeAngelis, G. C. & Newsome, W. T. Perceptual "read-out" of conjoined direction and disparity maps in extrastriate area MT. *PLoS Biol.* **2,** E77 (2004).
16. Perrett, D. I., Hietanen, J. K., Oram, M. W. & Benson, P. J. Organization and functions of cells responsive to faces in the temporal cortex. *Phil. Trans. R. Soc. Lond. B* **335,** 23–30 (1992).
17. Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B. & Tootell, R. B. Faces and objects in macaque cerebral cortex. *Nature Neurosci.* **6,** 989–995 (2003).
18. Zangenehpour, S. & Chaudhuri, A. Patchy organization and asymmetric distribution of the neural correlates of face processing in monkey inferotemporal cortex. *Curr. Biol.* **15,** 993–1005 (2005).
19. Salzman, C. D., Britten, K. H. & Newsome, W. T. Cortical microstimulation influences perceptual judgements of motion direction. *Nature* **346,** 174–177 (1990); Erratum. *Nature* **346**, 589 (1990).
20. Moore, T. & Armstrong, K. M. Selective gating of visual signals by microstimulation of frontal cortex. *Nature* **421,** 370–373 (2003).
21. Bruce, C., Desimone, R. & Gross, C. G. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.* **46,** 369–384 (1981).
22. Sugase, Y., Yamane, S., Ueno, S. & Kawano, K. Global and fine information coded by single neurons in the temporal visual cortex. *Nature* **400,** 869–873 (1999).
23. Kiani, R., Esteky, H. & Tanaka, K. Differences in onset latency of macaque inferotemporal neural responses to primate and non-primate faces. *J. Neurophysiol.* **94,** 1587–1596 (2005).
24. Horel, J. A., Pytko-Joiner, D. E., Voytko, M. L. & Salsbury, K. The performance of visual tasks while segments of the inferotemporal cortex are suppressed by cold. *Behav. Brain Res.* **23,** 29–42 (1987).
25. Buffalo, E. A. *et al.* Dissociation between the effects of damage to perirhinal cortex and area TE. *Learn. Mem.* **6,** 572–599 (1999).
26. Tovee, M. J., Rolls, E. T., Treves, A. & Bellis, R. P. Information encoding and the responses of single neurons in the primate temporal visual cortex. *J. Neurophysiol.* **70,** 640–654 (1993).
27. Keysers, C., Xiao, D. K., Foldiak, P. & Perrett, D. I. The speed of sight. *J. Cogn. Neurosci.* **13,** 90–101 (2001).
28. Foldiak, P., Xiao, D., Keysers, C., Edwards, R. & Perrett, D. I. Rapid serial visual presentation for the determination of neural selectivity in area STSa. *Prog. Brain Res.* **144,** 107–116 (2004).
29. Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics* (Peninsula Publishing, Los Altos, California, 1988).
30. Meeker, W. Q. & Escobar, L. A. Teaching about approximate confidence regions based on maximum likelihood estimation. *Am. Stat.* **49,** 48–53 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S.-M. Noorbakhsh for technical assistance, and Y. Najian for training one of the monkeys. We also thank W. T. Newsome, M. N. Shadlen and N. Kanwisher for comments on the manuscript.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to H.E. (esteky@ipm.ir).