Detecting faces in pure noise images: a functional MRI study on top-down perception

Hongchuan Zhang[a,c], Jiangang Liu[d], David E. Huber[b], Cory A. Rieth[b], Jie Tian[e,f] and Kang Lee[g]

[a]Center for Human Development, [b]Department of Psychology, University of California, San Diego, La Jolla, California, USA, [c]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, [d]School of Computer and Information Technology, Beijing Jiaotong University, [e]Institute of Automation, Chinese Academy of Science, Beijing, [f]Life Science Research Center, Xidian University, Xi'an, Shaanxi, China and [g] Institute of Child Study, University of Toronto, Toronto, Ontario, Canada

Correspondence to Kang Lee, Institute of Child Study, University of Toronto, 45 Walmer Road, Toronto, Ontario, Canada M5R 2X2
Tel: + 416 934 4597; fax: + 416 934 4565; e-mail: kang.lee@utotonto.ca, or Jie Tian, Institute of Automation, Chinese Academy of Science, Beijing 100080, China; e-mail: tian@ieee.org

Received 16 October 2007; accepted 18 November 2007

To assess the nature of top-down perceptual processes without contamination from bottom-up input, this functional MRI study investigated face detection in pure noise images. Greater activation was revealed for face versus nonface responses in the fusiform face area, but not in the occipital face area. Across participants, positive correlations were found for the degree of greater face-detection activation between the fusiform face area and bilateral inferior frontal gyri, suggesting a top-down pathway generating perceptual expectations. In contrast, the medial frontal, parietal, supplementary motor, parahippocampal, and striatal areas produced negative correlations between degrees of greater face-detection activation and behavioral responses, suggesting a possible role for these areas in selecting and executing appropriate responses that are based on the top-down expectations. *NeuroReport* 19:229–233 © 2008 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Keywords: face processing, functional MRI, (FFA) the fusiform face area, (OFA) the occipital face area, top-down perception

Introduction

One important function of the human visual system is to parse stimuli into coherent percepts that are based on degraded or ambiguous retinal images in everyday life. To do this, rather than passively building information from the bottom up, the brain continuously generates predictive 'codes' corresponding to expected objects [1–3]. According to this view, visual object recognition is accomplished through the combination of top-down expectations and bottom-up input, with a successful match indicating the most likely object. Although the bottom-up stream has been extensively studied, the top-down stream is, however, more theoretical and difficult to quantify.

In most studies, bottom-up input is strong, providing sufficient unambiguous information, thus limiting the role of top-down perceptual processes. This study sought to isolate top-down perceptual processes by asking participants to detect the presence/absence of a face in pure noise images. Similar noise techniques have been employed to study auditory [4] and visual detection [5,6], stereoscopic perception [7] as well as letter and face processing [8]. In most studies, however, a target was presented underneath the noise; whereas, in this one, no systematic bottom-up input was provided. Most of the perceptual work was thus left to top-down processes. Beyond developing 'classification images' to decode the top-down expectations during face detection, which are reported elsewhere, we identified the neural correlates of top-down face perception.

Previous studies found that face processing is enhanced by top-down expectations. For example, a grey oval-shaped stimulus produced greater activation in the fusiform face area (FFA) when placed on human bodies than when presented alone [9]. Presenting faces to one eye and houses to the other, a binocular rivalry study found greater activity in this area when faces dominated [10]. Furthermore, this greater activity occurred if one merely imagined faces versus other visual objects [11,12]. Analogous to binocular rivalry, higher fusiform activity was found when one reported seeing a face rather than a house, in intermixed images of faces and houses [13]. In another study, perceptual set proved to be an important determinant of fusiform and medial frontal cortical activation when comparing detection of faces versus houses, regardless of whether the actual display was of a degraded face, house, or car [14]. This linkage is consistent with a feedback connectivity from midfrontal regions to fusiform area, as revealed in a face-imagery task [15]. In all the above studies, however, either consistent bottom-up information was provided or the top-down processes were studied in a nonperceptual task (i.e. imagery). It is not clear whether bottom-up input was a critical factor for inducing or enabling top-down expectations. Instead, we isolated top-down processes by presenting no consistent bottom-up information in a perceptual-detection task. In the light of earlier results, we focus on face-sensitive areas in the fusiform and occipital regions, the so called 'core system' for face processing [16] and their relationship with other brain areas.

0959-4965 © Wolters Kluwer Health | Lippincott Williams & Wilkins

Methods

Participants

Sixteen right-handed healthy Chinese undergraduate and graduate students (eight women) participated in this study (22–25 years, mean age=22.0, SD=2.12) after giving their informed consent.

Stimuli and procedure

The experiment included an initial training period and a testing period. Participants were scanned only during the testing period. During the training period, participants completed six blocks of a face-detection task. They were told that in each block 50% of the 20 noise images contained a face, and 50% did not. They needed to indicate whether they saw a face or not with their right hand. The first two blocks were composed of faces that are easy to detect, followed by two blocks consisting of faces relatively more difficult to detect, and the last two blocks consisting of pure noise images only. This was to teach participants the nature of the experiment, and to keep them at detecting faces even no face was presented.

After the training period, participants completed four sessions of 120 testing trials with pure noise images only. They were instructed that the task was the same as in the last two training blocks and that 50% of the trials would still include faces. Forty checkerboard images were added in each session as controls. For each trial, a stimulus was presented for 600 ms after a 200-ms fixation cross, followed by a blank screen for 1200 ms.

Noise images were created with the Matlab software (Mathworks Inc., Sherborn, Massachusetts, USA), with a resolution of 480 × 480 pixels. Rather than using simple pixel noises, a more complex forms of noise was used to increase perceptual ambiguity (see Fig. 1). Easy and hard faces were created by mixing 40 front-view Asian face images (half male and half female, approximately same size) with 40 additional noise images. Throughout the entire experiment, every trial involved a different noise image.

After the testing period, a localizer task was administered, to identify the face-sensitive areas. It was separated into eight blocks of four object types (faces, Chinese characters, common objects, and scrambled pictures). In each block, 21 stimuli were presented at a rate of one stimulus per second, of which two repeated stimuli were used as a catch trial. Participants needed to press a button when such trials occurred. Twelve seconds of fixation null trials occurred between each block.

Functional MRI data acquisition and analysis

Structural and functional MRI data were collected using a 3.0 T MR imaging system (EXCITE, General Electric, Milwaukee, USA) at the Huaxi MR Research Center of Western China Hospital affiliated with Sichuan University. The functional fMRI series was collected using a single shot, T2*-weighted gradient-echo echo planar imaging (EPI) sequence (TR/TE=2000/30 ms; 36 slices, 4 mm sickness; matrix=64 × 64) covering the whole brain with a resolution of 3.75 × 3.75 mm. The same sequence was used for the localizer task, except that TR=3000. High-resolution anatomical scans were acquired with a three-dimensional enhanced fast gradient-echo sequence (TR/TE=8.516/3.4 ms, matrix=256 × 256), recording 156 axial images with a thickness of 1 mm and a resolution of 0.94 × 0.94 mm.

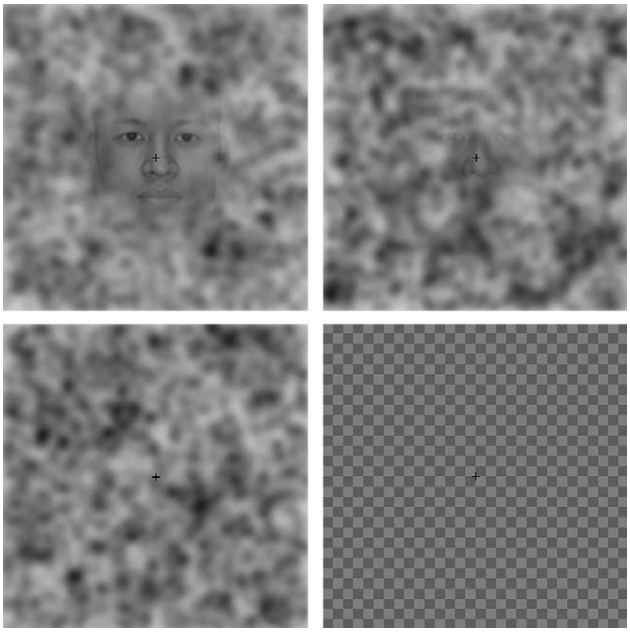

Fig. 1 Depiction of the images used in the experiment. Top left: easy-to-detect face (used in training only); top right: hard-to-detect face (used in training only); bottom left: noise image (used in training and testing); bottom right: checkerboard images. Noise images were created by mixing together multiple dark blobs at three different spatial scales indicated from a bivariate normal distribution. The positions of the blobs were chosen from a uniform distribution across the image. To equate for energy levels between spatial scales, the numbers of blobs were in proportion to the spatial scale of the blobs (i.e. fewer large blobs and more small blobs). All images included a central-fixation cross. To bias attention toward the center of the image, the mixing proportion between face and noise image was done according to a bivariate normal distribution centered on the central cross, such that only the central components of the face were clear.

Spatial preprocessing and statistical mapping were performed with SPM2 software (*www.fil.ion.ucl.ac.uk/spm/spm2.html*). The first three scans of each testing session and the first two scans of the localizer task were excluded from analyses. Functional scans were slice-timing corrected, spatially realigned (six-parameter rigid body), normalized using transformation parameters derived from the high-resolution anatomical scan coregistered to a mean echo planar image, and smoothed using a Gaussian filter of 8 mm.

Data from the localizer task were modeled with four boxcar functions for each object category, convolved with the canonical hemodynamic response function. These regressors were contrasted using a *t*-test in each participant (face > objects), and a *t*-threshold of 2 was chosen to identify face-sensitive areas [17,18]. Regions of interest were created through a 6-mm radius spherical volume centered on the voxel showing peak response in the above comparison.

For the testing period, trials were classified according to whether the response was face or nonface, resulting in two regressors convolved with the hemodynamic response function. Participant-specific parameter estimates were extracted from the first-level (within participant) analysis and compared at the group level, specific to the defined face-sensitive areas using a paired *t*-test. A conventional whole-brain search was also performed at the group level, with a threshold at P=0.001, uncorrected, and k=15.

Results

Behavioral performance

On average, participants responded seeing a face on 35.52% of the 480 pure noise images (minimum=8.33%, maximum=53.54%) with no difference in reaction time for face and nonface responses, t=1.00, $P>0.33$.

Functional MRI results

In the localizer task, all participants showed greater activation in the inferior occipital and fusiform regions for faces, as compared with other objects. Face-sensitive area was located on the right fusiform gyrus for 13 participants and on the left occipital gyrus for eight participants. Within the two areas, paired t-tests between the parameter estimates of face/nonface responses revealed greater activation for face responses in the FFA, $t(15)$=3.586, P=0.003, but no difference in the occipital face area (OFA), t=1.195, $P>0.25$ (see Fig. 2).

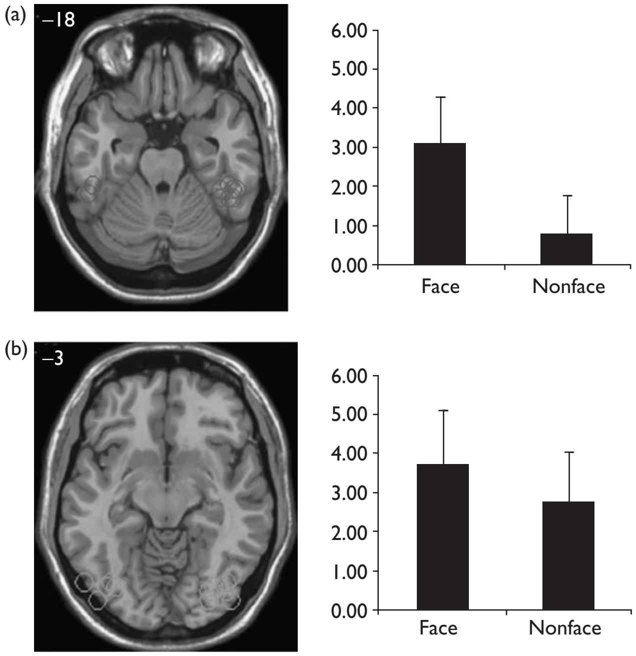

Fig. 2 (a) The fusiform face area defined for all participants (in circles, radius 6 mm) and the corresponding parameter estimates for face and nonface responses. (b) The occipital face area (in circles, radius 6 mm) and the parameter estimates.

A whole-brain search identified a distributed network with higher activation to face responses than to nonface responses, including medial frontal area, bilateral inferior prefrontal gyri, right supplementary motor area and postcentral gyrus, left inferior temporal/fusiform gyrus, precuneus and inferior parietal sulcus, and other areas including bilateral striatal area (caudate nucleus and putamen), right parahippocampal gyrus, and amygdala (see Table 1). The reverse comparison did not yield any significant results.

Correlation analyses examined the relations between the activations of the above-mentioned areas and those of the FFA and the OFA. For each area, parameter estimate of nonface responses were subtracted from that of face responses to generate a difference-activation score for each participant. Significant positive correlations with the FFA, but not the OFA, were found only for bilateral inferior frontal gyri (both Pearson's coefficients=0.684, P=0.003), and the right postcentral gyrus (Pearson's coefficient=0.508, P=0.045; see the bottom row of Fig. 3).

It is possible that higher activation for face responses simply reflects that more face responses were made, thus contributing to the correlating activations between the FFA and the inferior frontal and postcentral gyri. Different analyses thus examined the correlations between the difference scores in the active areas identified by the whole-brain analysis and the percentage of the 'face' responses across participants. Behavioral performance showed no significant correlations with activations in the inferior frontal and postcentral areas as well as in the fusiform and OFA. Instead, it was significantly negatively correlated with activations in the medial frontal, supplementary motor, inferior parietal, parahippocampal, and bilateral striatal areas (Pearson's coefficients ranged from −0.533 to −0.641).

Table 1 Activation peaks from the whole-brain analysis comparing face responses with nonface responses

		Tailarach coordinates				
	Brain regions	*X*	*Y*	*Z*	Cluster size	*Z* score
Frontal	Medial frontal area	3	4	30	221	5.02
		9	16	21		4.57
		6	4	22		4.33
	Left inferior frontal gyrus	−48	13	24	78	4.57
	Right inferior frontal gyrus	53	10	16	36	3.80
	Right supplementary motor area	15	2	50	25	4.68
Temporal	Left inferior temporal/fusiform gyrus	−48	−42	−21	77	4.10
		−48	−56	−7		3.76
		−53	−62	−10		3.61
Parietal	Left precuneus	−30	−62	36	29	3.98
	Left inferior parietal sulcus	−45	−44	44	15	3.34
		−42	−42	33		3.28
	Right postcentral gyrus	62	−13	26	44	3.90
Other	Left caudate nucleus/putamen	−9	11	−11	59	4.01
		3	−9	−5		3.83
	Right caudate nucleus/putamen	12	9	−3	17	3.71
	Right parahippocampal gyrus	15	−21	−19	23	3.67
	Right amygdala	18	−1	−18	15	3.86

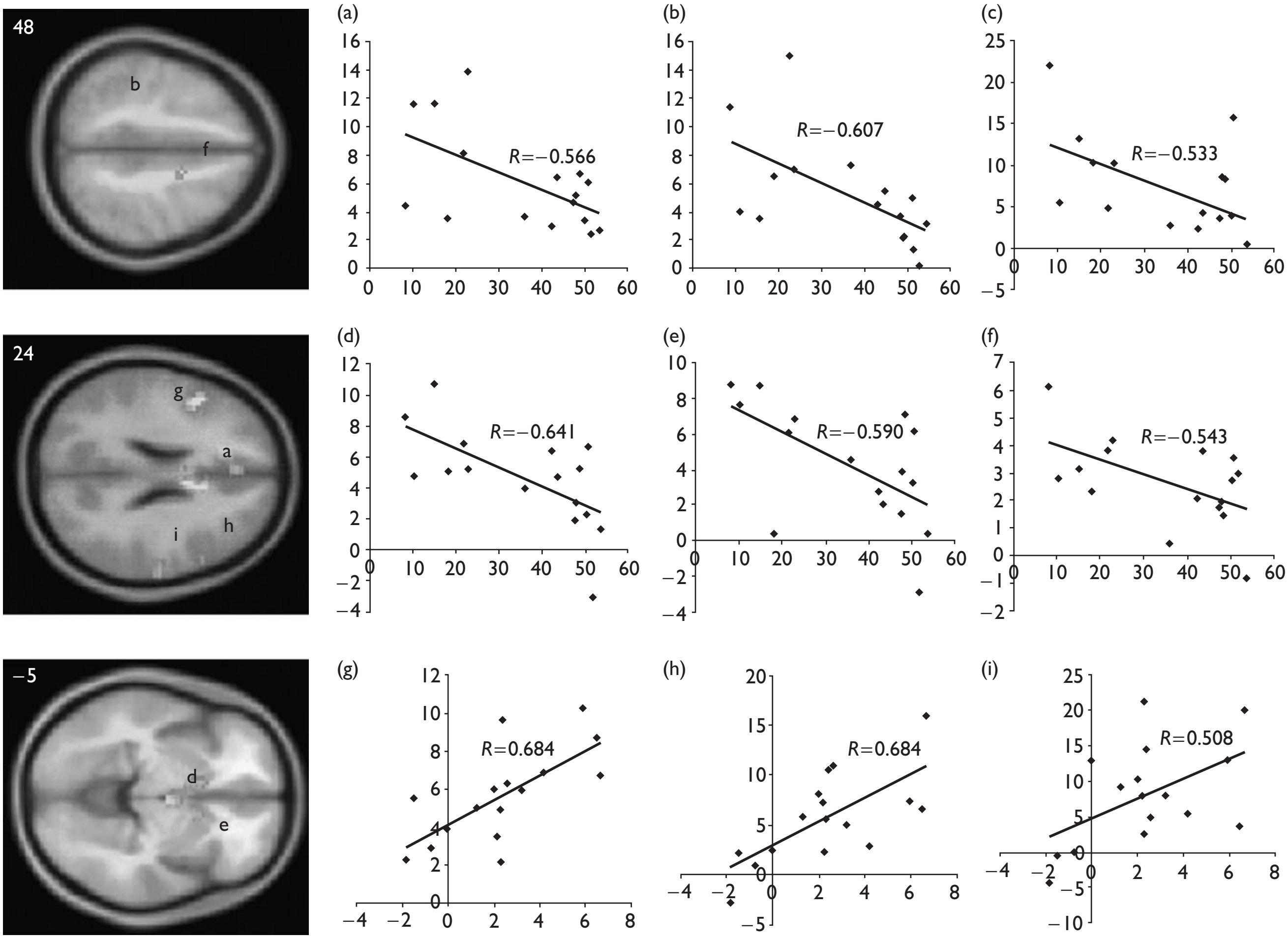

Fig. 3 As identified by whole-brain search, correlations between face/nonface-activation-difference scores are shown in the indicated regions (on the *y* axis), and either percentage of face responses (top and middle rows) or activation-difference scores in the fusiform face area (bottom row). All correlations were across individual participants. The images on the left show brain regions for the corresponding correlations appearing on the right, with *Z* Talairach position being indicated numerically. (a) Medial frontal area; (b) left inferior parietal sulcus; (c) right parahippocampal gyrus; (d) left caudate nucleus/putamen; (e) right caudate nucleus/putamen; (f) right supplementary motor area; (g) left inferior frontal gyrus; (h) right inferior frontal gyrus; and (i) right postcentral gyrus.

Discussion

For the first time, we identified neural correlates of top-down face processing in a perceptual task without systematic bottom-up information: Face detection in pure noise images corresponded to greater activation in the FFA, but not in the OFA, in line with previous findings when participants misperceived a house as a face [13].

It has been suggested that face processing consists of an early structural-encoding stage that identifies physical properties, followed by a later stage that differentiates between individual faces [19]. Previous studies found that the FFA is more sensitive to identity changes, whereas the OFA is more sensitive to retinotropic changes [20–22]. Functional connectivity analyses have identified a feedforward connection from the OFA to the FFA when viewing faces, as well as a feedback connection from frontal regions to the FFA when engaged in face imagery or mental set manipulations [14,15]. These findings suggest that the OFA is more sensitive to bottom-up input whereas, the FFA is the meeting-point between top-down and bottom-up face processing.

Our findings supported this fusiform meeting-point hypothesis. In particular, we identified a widely distributed frontal network involved in top-down face processing. Across participants, bilateral inferior frontal gyri correlated positively with the FFA when activations to face responses were compared with those to nonface responses, suggesting top-down connectivity between these areas. Human inferior frontal gyrus are believed to play an integral part in the semantic analysis of words and pictures [1], analogous to the ventrolateral prefrontal cortex in monkeys, which is known to be involved in visual object analysis [23]. It is believed that there is a rapid magnocellular pathway in monkeys conveying low-spatial-frequency information from the early visual cortex to the inferior temporal cortex and then to the ventrolateral prefrontal cortex, to generate object-related semantic knowledge [24]. It is thus possible that top-down expectations based on object-related semantic information are generated in this area, and are then back-projected to the visual cortices such as the FFA to match with the bottom-up input.

We also found negative correlations between activations in medial frontal, parietal, supplementary motor, parahippocampal, striatal areas, and behavioral performance. Additional work is needed to identify the separate roles of these neural structures within this top-down face-processing network. One possibility is that while the inferior frontal gyri are engaged in the generation of perceptual expectations, some of these negatively correlated areas are involved in resolving decision uncertainty, and in selecting and executing appropriate responses based on expectations [25], consistent with the massive and reciprocal connections between inferior and medial frontal areas.

Conclusion

This study examined the neural correlates of top-down face perception. When participants detected a face in pure noise images, the FFA was more active. A coupling between this area and the bilateral inferior frontal gyri suggests that this pathway perhaps provides top-down expectations that are projected to sensory areas, to assess the match between expectations and observations.

Acknowledgements

This study was supported in part by grants from the Joint Research Fund for Overseas Chinese Young Scholars (30528027-28), the Program for Cheung Kong Scholars and Innovative Research Team in University (PCSIRT, IRT0645), CAS Hundred Talents Program, NBRPC (2006CB705700), NSFC (60532050), and NIH (R01 HD46526). This work is supported in part by grants from the Joint Research Fund for Overseas Chinese Young Scholars (30528027), PCSIRT program (IRT0645), CAS Hundred Talents Program, NBRPC (2006CB705700), NSFC (60532050), and NIH (R01 HD46526).

References

1. Bar M. A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cogn Neurosci* 2003; **15**:600–609.
2. Smith PL, Ratcliff R. Psychology and neurobiology of simple decisions. *Trends Neurosci* 2004; **27**:161–168.
3. Bar M. The proactive brain: using analogies and associations to generate predictions. *Trends Cogn Sci* 2007; **11**:280–289.
4. Ahumada A, Lovell J. Stimulus features in signal detection. *J Acoustic Soc Am* 1971; **49**:1751.
5. Gosselin F, Bacon BA, Mamassian P. Internal surface representations approximated by reverse correlation. *Vis Res* 2004; **44**:2515–2520.
6. Gold JM, Murray RF, Bennett PJ, Sekuler AB. Deriving behavioural receptive fields for visually completed contours. *Curr Biol* 2000; **10**: 663–666.
7. Neri P, Parker AJ, Blakemore C. Probing the human stereoscopic system with reverse correlation. *Nature* 1999; **401**:695–698.
8. Gosselin F, Schyns PG. Superstitious perceptions reveal properties of internal representations. *Psychol Sci* 2003; **14**:505–509.
9. Cox D, Meyers E, Sinha P. Contextually evoked object-specific responses in human visual cortex. *Science* 2004; **304**:115–117.
10. Tong F, Nakayama K, Vaughan JT, Kanwisher N. Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 1998; **21**: 753–759.
11. O'Craven KM, Kanwisher N. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J Cogn Neurosci* 2000; **12**:1013–1023.
12. Ishai A, Ungerleider LG, Haxby JV. Distributed neural systems for the generation of visual images. *Neuron* 2000; **28**:979–990.
13. Summerfield C, Egner T, Mangels J, Hirsch J. Mistaking a house for a face: neural correlates of misperception in healthy humans. *Cereb Cortex* 2006; **16**:500–508.
14. Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J. Predictive codes for forthcoming perception in the frontal cortex. *Science* 2006; **314**:1311–1314.
15. Mechelli A, Price CJ, Friston KJ, Ishai A. Where bottom-up meets top-down: neuronal interactions during perception and imagery. *Cereb Cortex* 2004; **14**:1256–1265.
16. Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. *Trends Cogn Sci* 2000; **4**:223–233.
17. Golby AJ, Gabrieli JDE, Chiao JY, Eberhardt JL. Differential responses in the fusiform region to same-race and other-race faces. *Nat Neurosci* 2001; **4**:845–850.
18. Gauthier I, Skudlarski P, Gore JC, Anderson AW. Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci* 2000; **3**:191–197.
19. Bruce V, Young A. Understanding face recognition. *Br J Psychol* 1986; **77**:305–327.
20. Gilaie-Dotan S, Malach R. Sub-exemplar shape tuning in human face-related areas. *Cereb Cortex* 2007; **17**:325–338.
21. Levy I, Hasson U, Avidan G, Hendler T, Malach R. Center-periphery organization of human object areas. *Nat Neurosci* 2001; **4**:533–539.
22. Andrews TJ, Ewbank MP. Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage* 2004; **23**:905–913.
23. Rainer G, Miller EK. Neural ensemble states in prefrontal cortex identified using a hidden Markov model with a modified EM algorithm. *Neurocomputing Int J* 2000; **32–33**:961–966.
24. Rempel-Clower NL, Barbas H. The laminar pattern of connections between prefrontal and anterior temporal cortices in the rhesus monkey is related to cortical structure and function. *Cereb Cortex* 2000; **10**:851–865.
25. Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S. The role of the medial frontal cortex in cognitive control. *Sci Special Issue Cogn Behav* 2004; **306**:443–447.