How the brain learns to see objects and faces in an impoverished context

R. J. Dolan†, G. R. Fink‡, E. Rolls†, M. Booth, A. Holmes†, R. S. J. Frackowiak & K. J. Priston†

* Wellcome Department of Cognitive Neurology, Institute of Neurology, Queen Square, London WC1N 3BG, UK
† Academic Department of Psychiatry, Royal Free Hospital School of Medicine, London NW3, UK
‡ Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK

A degraded image of an object or face, which appears meaningless when seen for the first time, is easily recognizable after viewing an undegraded version of the same image. The neural mechanisms by which this form of rapid perceptual learning facilitates perception are not well understood. Psychological theories suggest the involvement of systems for processing stimulus attributes, spatial attention and feature binding, as well as those involved in visual imagery. Here we investigate where and how this rapid perceptual learning is expressed in the human brain by using functional neuroimaging to measure brain activity during exposure to degraded images before and after exposure to the corresponding undegraded versions (Fig. 1). Perceptual learning of faces or objects enhanced the activity of inferior temporal regions known to be involved in face and object recognition respectively. In addition, both face and object learning led to increased activity in medial and lateral parietal regions that were implicated in spatial attention, category-independent learning and visual imagery. We observed a strong coupling between the temporal face area and the medial parietal cortex when, and only when, faces were perceived. This suggests that perceptual learning involves direct interactions between areas involved in face recognition and those involved in spatial attention, feature binding and memory recall.

The experimental design involved four conditions: two-tone (black and white) images of objects before (Ob) and after (Oa) exposure to associated grey-scale images of the objects; two-tone images of faces before (Fb) and after (Fa) exposure to associated grey-scale images of the faces. A factorial experimental design, crossing pre- and post-grey-scale exposure with object or face perception, found that the same visual input was maintained across all levels of stimulus presentation, with the condition-specific neural responses being measured with positron emission tomography (PET) indices of local perfusion. Thus, the factors in the experiment are exposure (two levels: before and after learning) and stimulus (two levels: object and face). Each combination of levels was repeated 3 times on 8 subjects (study 1; n = 96) and 6 subjects (study 2; n = 72) as described below.

In study 1, we first examined the main effect of exposure using the contrast (Oa + Fa) − (Ob + Fb) to identify regions activated by the combined detection of objects and faces in study 1. This comparison revealed extensive bilateral activation in the medial parietal region (P < 0.05, corrected), corresponding to the precuneus, and extending laterally to involve the posterior inferior parietal cortex (P < 0.05, corrected; Fig. 2).

We next examined the main effect of stimulus to localize category-specific activations. Comparing object with face conditions, using the contrast (Ob + Oa) − (Fb + Fa), showed activations in the left inferior temporal region (P < 0.05, corrected) (Fig. 3a). The opposite contrast (Fb + Fa) − (Ob + Oa) revealed activation in the right superior temporal region (P < 0.05, corrected) extending into the right inferior temporal region (P < 0.001, uncorrected) in association with the face conditions (Fig. 3b). Finally, to identify regions whose object-specific responses were enhanced by exposure-dependent perception, we tested for the conjunction of object-specific responses (that is, the main effect of objects relative to faces, which equals (Ob + Oa) − (Fb + Fa)) and significantly increased perception-dependent activations (that is, category × exposure interaction) in the left fusiform gyrus (Fig. 3d). The equivalent conjunction analysis for learning-dependent effects due to object perception was seen in the left fusiform region (z = 4.4; P < 0.05 corrected) (Fig. 3c). The equivalent conjunction analysis for learning-dependent effects due to faces demonstrated a similar effect in the right fusiform gyrus (Fig. 3d) (z = 3.22; P < 0.001, uncorrected).

The experimental design interposes an explicit learning phase between first and second stimulus presentation and includes a learning component due to repeated exposure to the same stimuli. To remove the influence of this order-effect learning, we contrasted the patterns of activation associated with explicit learning (study 1) with those elicited in a second experiment (study 2) by interposing non-associated grey-scale images. These contrasts involved a group × effects interaction where the effects comprised those associated with category-dependent learning (combined post-exposure minus pre-exposure) and those specific to learning in the faces and objects conditions (post-exposure minus pre-exposure for faces minus objects and vice versa). The comparisons reach significance at an uncorrected level and are reported descriptively.

The group × exposure interaction (category-independent learning)
identified the medial parietal cortex \( (z = 2.9; P = 0.002, \text{uncorrected}) \). For the group \( \times \) exposure \( \times \) stimulus interaction (category-specific learning) the right fusiform gyrus activation was significant \( (z = 3.6; P < 0.001, \text{uncorrected}) \) for face learning but the left fusiform activation for object learning failed to reach significance even at a threshold \( z = 2.33 \) (corresponding to \( P < 0.01 \)).

We next tested the hypothesis that perceptual learning involved interactions between the medial parietal cortex, an area implicated in higher-order cognition, and the area showing face-learning responses. We tested this hypothesis by creating a statistical parametric map to identify areas where activity can be explained in terms of face-specific interactions with medial parietal cortex activation (Fig. 2). The analysis used a statistical model that includes an effect of faces relative to objects (and vice versa) and a term that represents the interaction between this effect and adjusted activity in the medial parietal cortex. This term is simply the adjusted activity in the parietal region multiplied by 1 when faces were present and \(-1\) when they were not. The significance of this interaction term was assessed using \( t\)-statistics and an uncorrected threshold of \( P < 0.01 \). Because we predicted that this modulatory effect would be expressed in the fusiform gyrus, we selected the regional activation (surviving a threshold of \( P < 0.01, \text{uncorrected} \)) that was closest to the boundary of the fusiform gyrus. This region comprised 340 voxels. The probability of this number of voxels or more being obtained by chance was \( P < 0.039; z = 4.0 \). This can be considered as \( P \) value that has been corrected for the volume of the fusiform gyrus, indicating a significant modulation of face-specific responses in the right fusiform region by medial parietal cortex \( (\text{Fig. 4}) \).

We have made several critical findings. First, learning-facilitated perception of objects or faces in degraded stimuli specifically involves activation of medial parietal cortex. Second, we observed enhanced neural responses in the fusiform gyri to an identical visual input when it elicited a strong visual percept. This effect was most striking for the enhanced percept of faces, for which augmentation of activation in the right fusiform involved a significant interaction with medial parietal cortex. The lack of significance for a similar effect in the object condition, in the between-group analysis, we attribute to greater nonspecific effects of attentional set engendered when subjects attempt a reconstruction of objects following exposure to non-related grey-scale objects in the second study.

Category-dependent lateralized activations, in the direct comparisons of object and face conditions, do not necessarily imply that object and face processing solely involve the left and right fusiform regions. We would expect bilateral activations if we had included a neutral control condition. Selective fusiform activations during face perception have been reported \( ^{10-12} \). Likewise, inferior temporal cortex and nearby occipital regions play a role in pattern and object recognition \( ^{13-17} \). Our findings extend these observations by demonstrating that the right fusiform region alters its response to activation as a result of learning. Moreover, as the same visual inputs were used across all conditions, our findings specifically address what brain areas are activated when faces and objects are perceived.

Another novel finding is the observation that learning-facilitated perception involves not only fusiform regions but also medial (specific to learning) and lateral parietal cortex. The functional role of the latter regions in learning-facilitated perception is open to interpretation. First, perception-related fusiform activations may reflect learning-related tuning of neural activity similar to that seen in monkeys \( ^{18-22} \). The parietal responses could reflect propagated

---

**Figure 1** The experimental design. Binarized images of an object (top row) and face (bottom row) and their associated full grey-scale versions. The process of binarizing images involved transforming grey-scale levels into either black or white (two-tone) with values of either 0 or 1. Exposure to the associated grey-scale version took place 5 min before a second exposure to the two-tone version in study 1. In the pre- and post-learning scans, the subjects were told that they might see faces or objects, respectively, in the stimuli. A significant behavioural learning effect, operationally defined as the facilitation of performance by prior exposure to the grey-scale version, was evident for the object and face conditions. Behavioural data were collected for all conditions at the end of each individual scan. The mean number of resolved percepts pre- and post-exposure to the grey-scale images were 13 and 87% in the object, and 65 and 93% in face conditions, respectively. In study 2, the same sequence and stimuli were used but with pre- and post-exposure being interposed with non-related grey-scale images of objects and faces. Here the mean number of resolved percepts pre- and post-exposure to non-associated grey-scale images were 8 and 10% in the object, and 19 and 25% in the face conditions, respectively.

**Figure 2** Main effect of exposure. Relative activations (for the eight subjects) associated with post-learning perception of objects or faces: that is, main effect of exposure \( (Oa + Fa) - (Ob + Fb) \). Areas of significant activations \( (P < 0.05, \text{corrected}) \) are shown as through-projection onto representations of standard stereotactic space. Coronal, back view; transverse, view from above. Transverse SPM\( z \) maps were superimposed upon the group mean magnetic resonance image (MRI) spatially normalized into the same anatomical space. Coordinates are described with reference to the Talairach system \( ^{30} \). The local maximum within the activated volume is defined by the intersection of red lines. There is bilateral activation centred on the precuneus \( (x = 4 \text{ mm}, y = -74 \text{ mm}, z = 36 \text{ mm}; P < 0.05, \text{corrected}; z\text{-score} = 5.7) \), extending into the lateral aspects of the inferior posterior parietal cortex. Adjusted mean rCBF (to a mean of 50 ml dl\(^{-1}\) min\(^{-1}\)) with s.e.m. per condition displayed for the local maximum at the above coordinates.
outputs from fusiform regions that engage higher-order visual regions to mediate a spatial reconstruction, from fragmentary evidence, of perceived images. The lateral parietal cortex is implicated in spatial attention and in feature binding necessary for spatial representation of objects. The medial parietal cortex (precuneus) on the other hand, is implicated in memory-related imagery. As our experimental tasks involve reconstruction of object (or face) representations from fragmentary evidence we suggest that exposure-related activation of medial parietal cortex reflects this ‘mind’s eye’ reconstruction which may be a necessary process in high-level perception.

An alternative account is that category-specific enhanced fusiform activations reflect neuronal populations modulated by top-down influences, from the parietal cortex, representing an interaction of mnemonic, imagery and attentional processes with category-specific stimuli. This synthesis would necessitate on-going dynamic interactions between category-specific regions (fusiform gyri) and non-category-specific higher-order regions (parietal) necessary for the coherent synthesis of a percept from an insufficiently specified input. Top-down effects are implicated by psychophysical experiments that demonstrate imagery-based facilitation of early sensory representations. Likewise, single-unit recordings demonstrate that memory and attention can tune responses of inferior temporal cortex to repeated encounters with stimuli, and neuroimaging data has shown that directed attention influences early sensory processing. The likely mechanisms can be inferred from experiments that demonstrate behaviourally specific reciprocal interactions between parietal and visual regions, involving synchronization of responses, in the awake cat.

The striking modulatory influences demonstrated here indicate that learning-based facilitation of perception involves response modulation in category-specific extrastriate neuronal populations...
and interactions with higher-order brain regions. Our results support current psychological theories that perception is a conjoint function of current sensory input interacting with memory and possibly attentional processes. The findings extend these psychological observations by providing neurobiological evidence that memory and attention organize the visual input through selective effects on extrastriate visual processing.

Methods

In study 1, eight normal-sighted male volunteers (six right- and two left-handed), and six normal-sighted right-handed male volunteers in study 2 were scanned using a SIEMENS/CPS ECAT EXACT HR+ (model 962) PET scanner (Siemens/CTI) in 3-dimensional mode with a 15-cm axial field of view. Relative cerebral blood flow (rCBF) was measured from the distribution of radioactivity after slow bolus intravenous (i.v.) injection of $\text{H}_2\text{O}$ (9 mCi per scan, each lasting 90 s). Attenuation-corrected data were reconstructed into 63 image planes with a resulting resolution of 6 mm at full-width-half-maximum. For each subject, structural magnetic resonance (MR) images were obtained with a 2 T Magnetom VISION scanner (Siemens).

Twelve regional perfusion measurements were taken during the following four conditions: two-tone images of objects before learning (the learning phase consisted of presentation of the grey-scale images); two-tone images of objects post-learning; two-tone images of faces pre-learning; two-tone images of faces post-learning. The order of presentation of stimuli that represented either objects or faces was counterbalanced but each set of binarized images were grouped back to back. For any set of binarized stimuli, subjects were scanned twice (pre- and post-exposure to the associated grey-scale image). Subjects were aware that binarized images following exposure to faces or objects might relate to the category of stimulus seen in the learning phase. Subjects lay with eyes open in a quiet, darkened room watching a 36-cm video display unit at a viewing distance of 40 cm. Subjects were exposed to individual images for 3 s. Ten images in all were viewed during a single scan window, with a 1-s interstimulus interval. Subjects had 6 scans before and 6 scans after learning (3 scans for objects and 3 for faces, pre- and post-learning). Learning with individual grey-scale images occurred 5 min before re-exposure to the related two-tone images.

Statistical parametric mapping (SPM96, Welcome Department of Cognitive Neurology, London) software was used for image realignment, transformation into standard stereotactic space, smoothing and statistical analysis. All measurements per condition were averaged across subjects. State-dependent differences in global flow were co-varied out using ANCOVA. Main effects and interactions were assessed with contrasts of the adjusted task means using $t$-statistics subsequently transformed into the $z$ statistic. For between-group comparisons, although our block ANCOVA removed random effects due to subjects, this does not constitute a random effects model. Accordingly, we ensured the appropriateness of our fixed effects model by testing explicitly for subject $\times$ contrast interaction effects and demonstrating that they were not significant. Corrected values refer to correction for the whole brain volume based upon magnitude of the response. The resulting set of $z$ values constituted a statistical parametric map (SPM[z]). Localization of maxima are reported within the standard space, as defined by Talairach and Tournoux, and were superimposed on the group mean magnetic resonance image spatially normalized into the same anatomical space.

Received 20 March; accepted 22 July 1997.


Acknowledgements. We thank our volunteers and the radiography staff at the Wellcome Department. R.I.D., K.J.F., R.S.I.F. and G.R.F. are supported by the Wellcome Trust. The bibliographic support of R. Laiz continues to be invaluable. Correspondence and requests for materials should be addressed to R.J.D. (e-mail: r.dolan@pbl.iul.ac.uk).

A hippocampal GluR5 kainate receptor regulating inhibitory synaptic transmission

Vernon R. J. Clarke*, Barbara A. Ballyk†, Ken H. Hoo†, Allan Mandelzys‡, Andrew Pellizzari‡, Catherine P. Bath*, Justyn Thomas†, Erica F. Sharpe†, Ceri H. Davies‡, Paul L. Ornstein‡, Darryle D. Schoepp‡, Rajender K. Kamboj†, Graham L. Collingridge*, David Lodge§ & David Bleakman‡

* Department of Anatomy, University of Bristol, University Walk, Bristol BS8 1TD, UK
† Allelix Biopharmaceuticals, 6850 Grewore Drive, Mississauga, Ontario, LAV 1V7 Canada
‡ Eli Lilly and Company Ltd, Lilly Research Centre, Eel Wood Manor, Windlesham, Surrey GU20 6PH, UK
§ Department of Pharmacology, University of Edinburgh, 1 George Square, Edinburgh EH3 9JZ, UK

Eli Lilly and Company Ltd, Lilly Corporate Centre, Indianapolis, Indiana 46285, USA

The principal excitatory neurotransmitter in the vertebrate central nervous system, L-glutamate, acts on three classes of ionotropic glutamate receptors, named after the agonists AMPA (alpha-amino-3-hydroxy-5-methyl-4-isoxazole-4-propionic acid), NMDA (N-methyl-D-aspartate) and kainate†. The development of selective pharmacological agents has led to a detailed understanding of the physiological and pathological roles of AMPA and NMDA

letters to nature

NATURE | VOL 389 | 9 OCTOBER 1997


599