fMRI–Adaptation Studies of Viewpoint Tuning in the Extrastriate and Fusiform Body Areas

John C. Taylor, Alison J. Wiggett, and Paul E. Downing
Wales Institute of Cognitive Neuroscience, School of Psychology, Bangor University, Bangor, United Kingdom

Submitted 21 July 2009; accepted in final form 19 December 2009

Taylor JC, Wiggett AJ, Downing PE. fMRI–adaptation studies of viewpoint tuning in the extrastriate and fusiform body areas. J Neurophysiol 103: 1467–1477, 2010. First published December 23, 2009; doi:10.1152/jn.00637.2009. People are easily able to perceive the human body across different viewpoints, but the neural mechanisms underpinning this ability are currently unclear. In three experiments, we used functional MRI (fMRI) adaptation to study the view-invariance of representations in two cortical regions that have previously been shown to be sensitive to visual depictions of the human body—the extrastriate and fusiform body areas (EBA and FBA). The BOLD response to sequentially presented pairs of bodies was treated as an index of view invariance. Specifically, we compared trials in which the bodies in each image held identical poses (seen from different views) to trials containing different poses. EBA and FBA adapted to identical views of the same pose, and both showed a progressive rebound from adaptation as a function of the angular difference between views, up to ~30°. However, these adaptation effects were eliminated when the body stimuli were followed by a pattern mask. Delaying the mask onset increased the response (but not the adaptation effect) in EBA, leaving FBA unaffected. We interpret these masking effects as evidence that view-dependent fMRI adaptation is driven by later waves of neuronal responses in the regions of interest. Finally, in a whole brain analysis, we identified an anterior region of the left inferior temporal sulcus (l-aITS) that responded linearly to stimulus rotation, but showed no selectivity for bodies. Our results show that body-selective cortical areas exhibit a similar degree of view-invariance as other object selective areas—such as the lateral occipitotemporal area (LO) and posterior fusiform gyrus (pFs).

INTRODUCTION

Human bodies are ubiquitous, salient, and easily recognized from many viewpoints, but the structure of neural body representations is not well understood. The extrastriate and fusiform body areas—EBA (Downing et al. 2001) and FBA (Peelen and Downing 2005; Schwarzlose et al. 2005), respectively—are two regions of the visual system that are highly sensitive to visual depictions of the human body (for a review, see Peelen and Downing 2007). EBA is located in the posterior inferior temporal sulcus, and FBA is found in the fusiform gyrus, where it overlaps with the fusiform face area (Peelen and Downing 2005; Schwarzlose et al. 2005). Recent studies have provided some clues regarding the respective contributions of these regions to representing body form and action (Moro et al. 2008; Schwarzlose et al. 2005, 2008; Taylor et al. 2007; Urgesi et al. 2007b). Furthermore, two studies have investigated the viewpoint dependency of EBA (Chan et al. 2004; Saxe et al. 2006), showing a small right hemisphere bias in this region for allocentric over egocentric views of bodies and body parts. However, to date, no comprehensive study of the view invariance of whole-body representations in EBA and FBA has been reported.

One potential source of clues about the representations in EBA and FBA is the evidence on nearby object representations in the lateral occipital complex (LOC). The LOC is segregated into discrete subregions—a dorsal-caudal region (LO) and a ventral region (pFs), extending into the fusiform gyrus (Grill-Spector et al. 1999). Anatomically, LO falls close to EBA (Downing et al. 2007), and pFs is near FBA. There may be some functional similarities between areas generally involved in object perception and neighboring areas that are selective for the human form. LO (relative to pFs) is sensitive to object size (Grill-Spector et al. 1999) and shows greater sensitivity to image scrambling (Lerner et al. 2001, 2002). Area pFs, but not LO, responds to stimulus changes that alter 3D form but leave the 2D outline intact (Kourtzi et al. 2003). Furthermore, LO carries more information about local object parts than pFs (Drucker and Aguirre 2009) and favors physical aspects of the stimulus rather than the resulting subjective percept (Hauk and Seltzer 2008).

Taken together, the above data imply a greater sensitivity to local features in lateral object areas, compared with “higher level,” more holistic representations in their ventral counterparts, a distinction that may extend to body-selective regions. Indeed, EBA responds selectively to individual body parts, in contrast to FBA, which shows selectivity for larger assemblies of parts and whole bodies (Taylor et al. 2007). The claim for a stronger representation of parts in EBA is supported by transcranial magnetic stimulation (TMS) evidence (Urgesi et al. 2007a). In this study, match to target performance for inverted (but not upright bodies) was impaired by TMS over EBA. TMS over ventral premotor cortex (vPMC) gave the opposite pattern. The authors concluded that performance in the upright condition was driven by configural (template based) representations, whereas inversion disrupts template-based perception and is reliant on local feature analysis. Disruption of EBA therefore leads to a deficit in local feature processing.

Behavioral studies of the perception of human-like mannequins across changing viewpoints suggest that body representations are somewhat view-dependent. Kourtzi and Shiffrar (1998) showed priming of unseen views of stimuli for rotations of ±30° from a previously seen view, but not for rotations of 60°, provided the views were not linked by motion. In subsequent functional MRI (fMRI) studies using synthetic objects, Kourtzi et al. (2003) showed response invariance to rotations of 30° in object-selective regions of LOC. Here we used a similar approach to determine experimentally the degree of

Address for reprint requests and other correspondence: J. Taylor, School of Psychology, Bangor Univ., Adeliaid Brigantia, Penrallt Rd., Bangor LL57 2AS, UK (E-mail: psp223@bangor.ac.uk).
view invariance of human bodies in lateral and ventral body-selective regions.

To do so, we used an fMRI adaptation (fMRI-A) design. fMRI-A is a robust feature of cortical dynamics that manifests as an attenuation of BOLD signal response to repeated presentations of the same stimulus. Although the precise mechanisms behind fMRI-A are still debated (Epstein et al. 2008; for reviews, see Grill-Spector et al. 2006; Henson 2003), there is a consensus that the smaller BOLD signal response obtained for repeated (compared with different) stimuli can be treated as an index of stimulus similarity along a given dimension (Grill-Spector and Malach 2001). The effect of varying any stimulus parameter can be monitored in this way, providing researchers with a powerful tool for characterizing the functional selectivity of brain areas.

Adaptation paradigms have been widely applied in neuroimaging studies. Considering face perception alone, this approach has been used to study: the role of the occipital and fusiform face areas in holistic and part-wise analysis (Betts and Wilson 2007; Harris and Aguirre 2008; Schlitz and Rossion 2006); dissociations between identity and expression (Andrews and Ewbank 2004; Winston et al. 2004); the modulatory effects of emotion (Rotshtein et al. 2001); contributions of nonface cortical areas to face discrimination in prosopagnosic patients (Dricot et al. 2008); scale invariance (Eger et al. 2004); and viewpoint invariance (Fang et al. 2007). fMRI-A has proved to be of particular use in the study of representational invariance in the visual system, across a wide range of stimulus categories: e.g., sinusoidal gratings (Boynton and Finney 2003), geometric structures (Weigelt et al. 2007), abstract shapes (Kourtzi et al. 2003), object contours and perceived shape (Kourtzi and Kanwisher 2001), and natural and man-made objects (James et al. 2002). However, although fMRI-A has previously been applied to study some body properties in EBA—e.g., human action/identity pairings (Kable and Chatterjee 2006), self/other discrimination of body parts (Myers and Sowden 2008)—to date, view invariance has not been examined, and the adaptation characteristics of FBA remain unknown.

We used an fMRI-A paradigm to examine BOLD response changes to sequentially presented pairs (Buckner et al. 1998; Kourtzi et al. 2003) of human bodies. These differed either in the pose held by the figure or the angle at which they were viewed. Identity was held constant within each trial and varied across trials. For both EBA and FBA, we characterized the modulation of adaptation by changes in viewpoint. The extent of viewpoint change at which the response to a repeated pose reaches that produced by two different poses is taken as an index of the limit of view invariance for a given area. Our main aim was to assess whether view invariance in these regions was comparable to that seen in behavioral studies and in fMRI studies of object representation in the lateral occipital complex. Furthermore, on the basis of the evidence reviewed above, we hypothesized that the body representation in FBA may be more invariant to changes in view than EBA.

METHODS

Experiment 1

PARTICIPANTS. Thirteen right-handed participants (6 female) were recruited via the Bangor University Community Panel. All participants were briefed, and informed consent was obtained in compliance with procedures set by the School of Psychology Ethics Panel, which approved these studies. Participants were prescreened by questionnaire to identify any factors preventing inclusion, such as preexisting medical or neurological conditions or claustrophobia. All participants were right handed with normal (or corrected to normal) vision.

Data were acquired using a 3T Philips MRI scanner, fitted with aSENSE parallel head coil (Philips, Best, The Netherlands). Stimulus images were presented using a Sanyo LCD projector (Sanyo, Osaka, Japan) focused onto a translucent, rear projection screen. Presentation was controlled using MatLab Psychophysics Toolbox (Mathworks, Natick, MA), running on an Apple iBook G3 laptop computer with Macintosh OS9 operating system (Apple, Cupertino, CA). 1-back and target stimulus responses were recorded using a nonferrous, fiber optic response pad (Current Designs, Philadelphia, PA). Preprocessing, region of interest (ROI) definition and ROI/General Linear Model (GLM) calculations were performed using BrainVoyager QX 1.9 software (Brain Innovation, Maastricht, The Netherlands).

MAIN EXPERIMENT STIMULI. Stimuli were images of headless male and female human bodies (2 exemplars of each). Each body was arranged into eight different poses and positioned on a minimally defined floor plane (Poser, Curious Labs, Santa Cruz, CA). Each pose was rotated through 24 successive 15° increments—giving a total of 768 images (24 views per pose, 8 poses per figure, 4 figures). All images were scaled to 400 × 400 pixels, 72 dpi, and default Poser image colors were retained. Seven poses were selected as experimental stimuli, whereas the eighth was a “target” pose for the detection task. Masks were prepared from each of the four figures, to allow the masking of each figure to be matched for low-level image properties. One exemplar of each figure was selected and divided into an 11 × 11 matrix. The spatial arrangement of the resulting 121 tiles was then scrambled. Examples of the four stimulus figures and mask images are given in Fig. 1.

LOCALIZER STIMULI. EBA and FBA localizer stimuli comprised 20, 400 × 400 pixel, 72 dpi grayscale images each of headless human bodies (in a variety of postures) and chairs. Image positions were jittered slightly on alternate presentations to prevent reliance on low-level transient detection.

SCANNING PARAMETERS. Functional data were obtained using T2*-weighted scans. A single shot echo planar (EPI) sequence was used because this approach minimizes potential motion artifacts resulting from participant head movement such as image ghosting and blurring. To capture all cortical ROIs, 15 oblique-axial slices were scanned to include, bilaterally, the inferior temporal and occipital lobes including the fusiform gyrus. Acquisition parameters for all participants were 112 × 112 matrix, voxel dimensions = 2 mm isotropic; echo time (TE) = 35 ms; repetition time (TR) = 1,500 ms; flip angle = 65°. Participants viewed rear-projected stimuli via a mirror positioned above the SENSE head coil. Functional scans were recorded for the duration of image presentation. Parameters for T1-weighted anatomical scans were: 256 × 256 matrix; voxel dimensions = 1 × 1 × 1 mm; TR = 12 ms, TE = 3 ms; flip angle = 8°.

MAIN EXPERIMENT. For the main experiment, participants were scanned on 3 of 32 possible sequences of a rapid event-related adaptation design. Each image sequence comprised 128, 3-s trials, including 16 fixation baseline trials. Nonbaseline trials comprised a serial presentation of two frames, each depicting the same human figure, for a duration of 300 ms each (Fig. 1). The first frame could be any of the eight poses in any of the 24 possible views. The second frame was either the same pose (5/6 trials) or a different pose (1/6 trials). Stimuli in frame 2 were either viewed from the same angle as frame 1 (0° condition) or rotated by ≥60° clockwise or counterclockwise, in 15° increments (this applied to both same and different poses). In this way, the absolute viewpoint of the frames was rendered
orthogonal to the extent of subsequent rotation. To disrupt an apparent motion percept, successive frames were jittered by 100 pixels on the x-axis—equivalent to 3.6° of visual angle, randomly from either right to left or left to right (cf. Kourtzi and Shiffrar 1999). Each frame was immediately followed by a stimulus specific mask (400 ms). Trials concluded with a question mark task response cue (1,600 ms). Before the start of the scan, participants were shown all 24 views of the target pose and instructed that they should press a response button (when the question mark appeared) if they had seen the target pose held by any of the actors, in either trial frame, viewed from any angle. Each run was bracketed with an additional five fixation-only epochs, giving a total run time of 276 s (138 TR). Target RTs were recorded, allowing independent modeling of both target trials and RTs. Trial orders were counterbalanced such that every stimulus condition was preceded equally often by every other condition, allowing an unbiased estimation of stimulus dependent HRFs (per subject, per run).

LOCALIZERS. EBA and FBA were localized individually in two separate runs, each comparing bodies to chairs. Each localizer run consisted of 21 15-s blocks. Blocks 1, 6, 11, 16, and 21 were fixation-only baseline epochs. In each of the remaining blocks, 20 images of either headless human bodies and images of chairs were presented. Stimuli were each displayed for 300 ms, followed by a blank screen for 450 ms.

DATA PREPROCESSING. To reduce the potential for T1 saturation, seven dummy volumes (not processed) were acquired before each experimental and localizer scan. Preprocessing entailed motion correction, with a maximum adjustment of ±2 mm (runs containing corrections that exceeded these limits were rejected) and high-pass filtering (0.006 Hz) to remove low-frequency and linear trends. The analysis was performed without spatial smoothing. T1-weighted anatomical scans were transformed into Talairach space (Talairach and Tournoux 1988). T2*-weighted, functional volumes were internally aligned to the first volume acquired for each participant manually co-registered, within subjects, to T1-weighted anatomical scans. In this way regional activation identified in functional scans could be accurately attributed via spatial comparison with higher resolution anatomical scans for individual participants.

ROI ANALYSIS. For each participant, GLMs were created for each localizer and main experiment using a boxcar predictor, convolved to a hemodynamic response function (HRF) for each stimulus condition. EBA and FBA were localized in each participant by contrasting the BOLD signal response obtained for bodies with that obtained for chairs [bodies – chairs]. ROIs were localized in both hemispheres where present.

Objective definition of ROIs was ensured by a statistical attribution method. The voxel exhibiting the strongest BOLD signal was identified within a relevant target region of the cortex suggested by current literature (EBA: Downing et al. 2007; FBA: Peelen and Downing 2005). Each ROI was defined as the voxels exhibiting significant activation (P < 0.0001, uncorrected), within a 9-mm cube surrounding the most strongly activated voxel. This approach further served to demarcate ROIs from adjacent regions of selective activation and limit them to only the most strongly activated voxels.

MAIN EXPERIMENT: GLM ANALYSIS. In the main experiment, predictors were defined as 300-ms events corresponding to the second frame in each trial. Predictors were classified as follows: Different Pose; Same/0°; Same/15°; Same/30°; Same/45°; Same/60°; Target trial; and Button-press. The button-press predictor was modeled individually, by trial, from the button-press data and was defined as the period spanning the offset of the second stimulus image in every target image pair and the response time recorded for that target (i.e., the duration of this predictor varied with task performance). Within each ROI, GLMs were calculated for all eight predictors, for each participant.
The time window used to define the button press predictor was calculated on a per trial basis, varying with response time to the target. Stimulus-dependent BOLD signal changes were expressed as standardized regression coefficients ($\beta$) and subjected to repeated-measures ANOVA with planned contrasts.

MAIN EXPERIMENT: TIME-COURSE MODELING. Raw MR signal was extracted from each ROI (i.e., 1 value per TR) and assigned to individual stimulus conditions. Alternating TRs—approximating to the task response period in each trial—were coded separately. Mean percent signal change (PSC) time-courses were calculated for every stimulus condition in every trial: PSC (%) = \left(\frac{MR_c - MR_f}{MR_f}\right) \times 100

Where $MR_c$ is the raw MR signal for an experimental condition and $MR_f$ is the raw MR signal for the fixation condition. Average time course data were used to confirm the presence of a stimulus dependent HRF across a complete run for each subject. This was considered necessary because of the possibility of both long-term adaptation and participant fatigue in the final experimental run. Average time courses for each run were therefore collapsed across trial conditions, and the run was rejected if the time course showed no discernable stimulus dependent profile.

Experiment 2

PARTICIPANTS. Ten right-handed participants (6 female) were recruited via the Bangor University Community Panel. Ethics procedures were identical to experiment 1.

STIMULI AND SCANNING PARAMETERS. The stimulus set was identical to that used in experiment 1, with the exception that one of the stimulus poses was replaced, as it was reported by participants to be very similar to another pose in the set. All scanning and data preprocessing procedures were the same as experiment 1. All instrument and stimulus presentation protocols, main experiment, and EBA/FBA localizer stimuli were identical to experiment 1.

MAIN EXPERIMENT. For the main experiment, participants were scanned on 6 of 12 possible orders of a rapid event-related adaptation design. Each image sequence comprised 91, 3-s trials, including 16 fixation-only, baseline trials. Nonbaseline trials comprised a serial presentation of two frames, each depicting the same human figure (300 ms each). The first frame in each trial comprised any one of the eight poses in any of the 24 possible views (randomized). The second frame was either the same pose (2/3 trials) or a different pose (1/3 trials). The viewpoint of the first frame was randomized and the orientation of the second frame was defined relative to the absolute viewpoint of the first frame—rendering the extent of rotation orthogonal to the absolute viewpoints tested (as per experiment 1). Run orders were ($n - 1$) counterbalanced such that all conditions were preceded equally often by all other conditions.

To attain statistical power comparable to experiment 1, the number of pose type conditions was reduced to three: different poses; same pose (same angle: $0^\circ$ condition); and same pose (different angle: $30^\circ$ condition). Each frame was followed by a stimulus specific mask (400 ms), either 0 (immediate condition) or 300 ms (delayed condition) after the offset of each body image (Fig. 2). In the delayed condition, the 300-ms poststimulus interval was filled with a blank frame in the background color. Runs were bracketed with an additional five fixation-only epochs, giving a total run of 273 s (182 TR). In other respects the procedure matched that of experiment 1.

Experiment 3

PARTICIPANTS. Twenty-two right-handed participants (12 female) were recruited via the Bangor, School of Psychology Community Panel. Ethics procedures were identical to experiments 1 and 2.

STIMULI, SCANNING, AND ANALYSIS. All apparatus and stimuli were the same as experiment 1. The trial structure was identical to experiment 1, with the exception that the masks were replaced with blank frames in the same color as the background. All scanning and data preprocessing procedures were the same as experiments 1 and 2. The trial structure is as shown in Fig. 2. Localization of EBA and FBA used identical procedures to experiment 1. Data analysis procedures were identical to experiment 1.

RESULTS

Experiment 1

Because of a scanner malfunction, incomplete data sets were obtained for two participants. Two further data sets were

FIG. 2. Example of functional regions of interest (ROIs: left to right, sagittal, transverse, and coronal views in 1 subject, $P < 0.0001$, uncorrected). The crosshairs indicate extrastriate body areas (EBAs) bilaterally (top) and extrastriate and fusiform body areas (FBAs) in the right hemisphere (bottom). Talairach coordinates are given (top left) for each slice.
rejected on the basis of participant drowsiness. Only data from the remaining nine participants was analyzed further.

ROI LOCALIZATION. The left and right EBA and right FBA were successfully identified in all the remaining participants. The mean spatial coordinates (with SE) of these regions, in Talairach space, were as follows: left EBA: X: −45(1.2), Y: −73(0.6), Z: 1.5(1.9); right EBA: X: 47(0.6), Y: −68(1.2), Z: 1.5(1.4); right FBA: X: 39(1.8), Y: −46(1.7), Z: −16(1.0). Example ROI data are shown in Fig. 2.

BEHAVIORAL DATA. Target detection was 84%, with all participants detecting a minimum of 40/48 targets. False alarm rate was 2.5%, averaged across all trials.

QUALITATIVE ASSESSMENT OF FMRI MAIN EXPERIMENT DATA. Stimulus-dependent HRF curves were obtained in all ROIs for each of the nine participants. On this basis, all data were included in the subsequent GLM analysis.

ADAPTATION TO CHANGES IN BODY-VIEW: COMPARING DIFFERENT AND PROGRESSIVELY ROTATED SAME POSES IN BILATERAL EBA AND RIGHT-FBA. Comparing right versus left-EBA, a [2 (hemisphere: L, R) × 6 (trial type: different, 0, 15, 30, 45, 60°)] repeated-measures ANOVA showed no significant main effects or interactions: [hemisphere, F(1,8) = 0.04, P = 0.85; trial type, F(5,40) = 5.45, P = 0.74; ROI × trial type, F(5,40) = 1.12, P = 0.37]. Because no hemisphere-dependent interaction was found, the EBA data were collapsed across hemispheres for comparison with right-FBA (Fig. 3). Again, no significant main effects [ROI, F(1,8) = 4.50, P = 0.067; trial type, F(5,40) = 1.16, P = 0.35] or interactions [ROI × trial type, F(5,40) = 1.04, P = 0.41] were found. Within each individual ROI, the effect of trial type was not significant and all simple contrasts (0, 15, 30, 45, 60°), each vs. a reference category of different pose) were not significant: F(1,8) = 0.22, P = 0.65; F(1,8) = 2.62, P = 0.15; F(1,8) = 0.00, P = 1; F(1,8) = 0.28, P = 0.61 and F(1,8) = 0.05, P = 0.84, respectively.

Finally, to establish whether the raw effect of adaptation was present, we used a paired samples t-test to compare the two most extreme conditions (Same, 0° view change vs. Different pose). All comparisons were not significant: in right-EBA: t(8) = 0.996, P = 0.35; in left-EBA t(8) = −0.027, P = 0.98; and in FBA t(8) = 1.39, P = 0.20.

Experiment 2

BEHAVIORAL DATA. Because of an instrument fault, no behavioral data were recorded for one participant. Target detection was at ceiling for all remaining participants in both masking conditions, with a total of five false alarms (<1%) across all trials and subjects. However, a comparison of RTs showed a 75-ms performance advantage in the delayed masking condition, t(7) = 3.51, P < 0.01: delayed, 507 ms; immediate, 581 ms.

MRI DATA. In the case of one participant, inconsistent shimming gave rise to unacceptable distortion in the L>R plane of the functional images in the first localizer run. This participant was excluded from further analyses. All subsequent analysis was restricted to the eight participants satisfying motion correction and task performance criteria and showing right-FBA and bilateral EBA.

ROI LOCALIZATION. The left and right EBA and right FBA were successfully identified in all of the remaining eight participants at a threshold of P < 0.0001 (uncorrected). The mean spatial coordinates (with SE) of these regions, in Talairach space, were as follows: left EBA: X: −45(1.5), Y: −70(2.5), Z: 2(1.9); right EBA: X: 45(1.2), Y: −70(2.0), Z: −2(2.3); right FBA: X: 41(1.3), Y: −47(2.7), Z: −18(1.2).

MAIN EXPERIMENT. Mean betas (Fig. 4) for the six predictors of interest were entered into a [2 (hemisphere, left-EBA, right-EBA) × 2 (trial type: delayed, immediate) × 3 (pose type: same 0°; same 30°; different)] ANOVA. In line with experiment 1, no significant interaction of hemisphere was detected in EBA, and therefore the EBA data were collapsed across hemispheres. Comparisons between collapsed EBA and right-FBA showed significant main effects of trial type (delayed > immediate), F(1,7) = 8.16, P = 0.024 and ROI (EBA > FBA) F(1,7) = 6.06, P = 0.043 and a significant interaction of ROI × trial type: F(1,7) = 23.30, P < 0.001. The main effect of trial type was marginally significant in right-FBA: F(1,7) = 5.15, P = 0.058. The main effect of pose type and the interaction of pose type × trial type were not significant (P > 0.05) in all cases.

Experiment 3

BEHAVIORAL MEASURES. The data for one participant was rejected on the basis of abnormal false alarm scores (16/42 targets identified vs. 173 false alarms). Data from one of three scans by a further participant were also excluded because of incorrect task switching (confirmed during debriefing). All remaining subjects performed close to ceiling, correctly identifying ≥43/48 targets (<1% false alarms).

FIG. 3. Results of experiment 1. Graphs show response magnitude (beta) vs. trial condition for 6 trial categories in EBA (collapsed across hemispheres) and right-FBA (n = 9). Error bars denote SE. Categories denote pairs of different pose pairs (Diff) and 5 magnitudes of rotation applied to same pose pairs (0–60°).
Qualitative Assessment of fMRI Data. Excessive head motion (4–9 mm) was found in four participant data sets, which were dropped from further analyses. One further participant’s data set was rejected on the basis that there was no stimulus dependent HRF in right-FBA for two of the three experimental runs.

ROI Localization. All analysis was restricted to the 15 participants satisfying motion correction and task performance criteria and demonstrating right-FBA and bilateral EBA (a threshold of $P < 0.005$ was required to obtain right FBA in 1 participant). The mean spatial coordinates (with SE) of these regions, in Talairach space, were as follows: left EBA: $X$: $-46(1.1), Y$: $-71(1.2), Z$: $3.4(1.1)$; right EBA: $X$: $43(4.5), Y$: $-71(0.8), Z$: $1.7(1.2)$; right FBA: $X$: $33(5.5), Y$: $-44(1.2), Z$: $-17(0.8)$.

ROI Analyses. The response pattern obtained for each ROI is shown in Fig. 5. Comparing right versus left EBA, a 2 (hemisphere: L, R) × 6 (pose type: different, 0, 15, 30, 45, 60°) repeated-measures ANOVA showed a significant main effect of pose type only, $F(5,70) = 4.93, P = 0.001$. The interaction of pose type × hemisphere was not significant, so the EBA response was collapsed across hemispheres for comparison with right-FBA. In this analysis, the main effect of pose type only was significant, $F(5,70) = 5.19, P < 0.001$. The interaction of pose type × ROI was not significant $F(5,70) = 0.44, P = 0.82$. Separate analyses of pose type in each ROI showed significant effects in both: EBA, $F(5,70) = 4.93, P = 0.001$; FBA, $F(5,70) = 3.71, P = 0.005$.

Planned (simple) contrasts were undertaken in each ROI to quantitatively assess the effect of viewpoint change on adaptation. In EBA, simple contrasts (0, 15, 30, 45, 60°, each vs. the reference category of different pose) were significant over the range 0–30°: $F(1,14) = 12.77, P = 0.003$; $F(1,14) = 10.29, P = 0.006$; $F(1,14) = 6.46, P = 0.024$, marginally significant at 45° $F(1,14) = 4.60, P = 0.05$, and not significant at 60°: $F(1,14) = 1.03, P = 0.33$. In FBA, the same contrasts were significant over the range 0–45°: $F(1,14) = 14.45, P = 0.002$; $F(1,14) = 12.70, P = 0.003$; $F(1,14) = 9.47, P = 0.008$; $F(1,14) = 5.42, P = 0.035$, respectively, and not significant at 60°: $F(1,14) = 1.27, P = 0.28$.

To further characterize the profile of viewpoint adaptation, the same body views were considered apart from the (unadapted) different views. EBA response was collapsed across hemispheres for comparison with right-FBA. Only the main effect of angle was significant, $F(4,56) = 2.75, P = 0.04$. The interaction of angle × ROI was not significant, $F(4,56) = 0.55, P = 0.70$.

The demonstration of a systematic, view-dependent adaptation profile in the absence, but not in the presence, of pattern masking (see experiment 1) suggests that pattern masking significantly disrupts viewpoint dependent adaptation. However, a comparison between the findings of the two experiments is complicated by the differing sample sizes (experiment 1: $n = 9$; experiment 3, $n = 15$). It is therefore possible that the significance of the results obtained in experiment 3 may be partially attributable to an increase in statistical power rather than solely to the removal of the mask. To address this possibility, we undertook a simple Monte Carlo analysis of the experiment 3 data set. We repeatedly ($n = 100$) randomly sampled subsets of 9 data sets from the 15 available in experiment 3 and compared the effect of raw adaptation (Same, 0° view change vs. Different pose) for each subsample. We
found that 96% of EBA and 93% of FBA subsamples showed a significant effect of adaptation.

**WHOLE BRAIN ANALYSIS.** If the pattern of response obtained in the ROI analysis was caused by attentional artifacts or to low-level retinotopic image changes, we would expect to see effects of angle outside of the body-specific ROIs we examined. To test this, a random-effects whole brain group analysis was performed, testing the linear effect of rotational angle in the “same” conditions: 0, 15, 30, 45, 60°,\[H_11002\]/H_11021,\[H_11002\]/H_11021,\[H_11002\]/H_11021,\[H_11002\]/H_11021,\[H_11002\]/H_11021. A cluster threshold of four contiguous voxels and a statistical threshold of \(P < 0.001\) (uncorrected) were applied. For comparison, a second whole brain random effects analysis was performed using localizer runs. Clusters were defined by the contrast [bodies – objects] at a threshold of \(P = 0.05\) (Bonferroni corrected). The more lenient threshold used for the linear contrast reflected the relatively low power per-condition of a five-level contrast derived from an event related design, relative to the two condition blocked localizer scans.

Three clusters responded significantly to the linear effect of viewpoint: Bilaterally, two areas of the lateral occipitotemporal cortex, close to EBA, and an anterior portion of the left anterior inferior temporal sulcus (aITS), at the boundary of BA20/21. Peak Talairach coordinates and cluster sizes are summarized in Table 1. To confirm that the lateral occipital activation was similar to EBA, the linear contrast was superimposed on a map of the group-average localizer contrast (Fig. 6). Extremely good agreement was found between the two areas, suggesting that the viewpoint tuning effect is restricted to body-selective regions. No significant clusters for the linear effect of viewpoint were observed in the fusiform gyrus and no body-selective clusters were found in the vicinity of l-aITS. Furthermore, l-aITS was not body-selective: Mean beta values \((n = 15)\) for bodies and chairs: \(-0.16 (SE = 0.03)\) and \(-0.15 (SE = 0.03)\), respectively. The contrast [bodies – chairs] was not significant, \(t(14) = -0.46 (P = 0.65)\).

**DISCUSSION**

**Experiment 1**

No view-dependent adaptation was obtained in any of the body-selective ROIs that were examined. Indeed, no significant response differences were seen even for the comparison of an identical repetition of a pose to two different poses. Task performance was strong, suggesting that participants were attending to the relevant stimuli features. A qualitative assessment of stimulus-dependent time-courses confirmed that, in all participants, in all runs, the stimuli were effectively engaging the ROIs. Furthermore, the magnitude of the \(t\)-values obtained for the comparison of (Same, 0° view change vs. Different pose) indicate that effect sizes are so small that no significant

<table>
<thead>
<tr>
<th>Region</th>
<th>(x)</th>
<th>(y)</th>
<th>(z)</th>
<th>Number of Voxels</th>
<th>(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left occipitotemporal cortex</td>
<td>-42.00</td>
<td>-67.00</td>
<td>7.00</td>
<td>144</td>
<td>5.10</td>
</tr>
<tr>
<td>Left inferior temporal sulcus</td>
<td>-60.00</td>
<td>4.00</td>
<td>-17.00</td>
<td>244</td>
<td>4.60</td>
</tr>
<tr>
<td>Right occipitotemporal cortex</td>
<td>51.00</td>
<td>-64.00</td>
<td>- 2.00</td>
<td>377</td>
<td>5.10</td>
</tr>
</tbody>
</table>
effect of raw adaptation would be detectable at any practically attainable sample size. Thus the absence of the predicted stimulus adaptation seems unlikely to be attributable either to statistical power, attentional factors (Murray and Wojciulik 2004) or to an inadequate response from the regions of interest.

Although the rapid fMRI-A paradigm has been used in many previous studies, the stimuli are not routinely masked in designs of this type (Fang et al. 2005; Kourtzi et al. 2003; Schiltz and Rossion 2006; Weigelt et al. 2007). The purpose of the mask in the current study was to assist in fully disrupting any percept of apparent motion between the two frames of the display, to prevent interpolation between successive body image frames in higher visual areas (Liu et al. 2004). This was because such interpolation would interfere with the aim of the study to examine the view specificity of body-selective neural populations (Kourtzi and Shiffrar 1997, 1999).

However, in this study, it is likely that the mask also interrupted ongoing visual analysis of the body images in EBA and FBA. When a visual image containing a human body is attended, the neural activity in body-selective areas presumably develops systematically over a brief time window, normally culminating in a full representation of the stimulus. In experiment 1, the mask appeared 300 ms after the offset of each body stimulus. Intracranial (McCarthy et al. 1999; Pourtois et al. 2007) and scalp (Gliga and Deheane-Lambertz 2005; Stekelenburg and de Gelder 2004; Thierry et al. 2006) ERP studies show a body-selective negativity peaking at around 200 ms poststimulus. While little further evidence is available for body stimuli, a recent study of the MEG analog of the face-selective N170 (M170: Liu et al. 2002), showed this component to be significantly viewpoint dependent over small rotational angles (Ewbank et al. 2007). These findings raise the possibility that the emergence of a complete representation in EBA and FBA extends beyond 200 ms and may have been disrupted by masking.

Experiment 2 further examined the role of the mask in eliminating adaptation effects in experiment 1 by manipulating the timing of the mask onset. Masks were either presented at 0 ms relative to stimulus offset (immediate mask trials, comparable to experiment 1) or delayed by 300 ms (delayed mask trials). Two specific predictions follow from this manipulation. First, RTs should be faster in the delayed mask condition, because more time is available after the second body stimulus in which to select and prepare the motor response. Second, and more important, if delaying the mask weakens or abolishes its interfering effects on the creation of a body representation in the regions of interest, we would expect a significant modulation of the BOLD responses in EBA/FBA as a function of the change in viewpoint from the first to the second body image.

Experiment 2

The data from experiment 2 showed that immediate pattern masking, relative to delayed pattern masking, produced a decrement in task performance as indexed by RT. In EBA bilaterally, but only marginally in FBA, delaying the mask onset significantly increased the BOLD response to bodies. However, in both ROIs, the absence of any view-dependent adaptation, even in the delayed mask condition, indicates that adaptation effects on body representations in these areas can still be substantially disrupted as late as 600 ms after stimulus onset. Notably the effect of masking differed between the two ROIs, suggesting a different time course of developing representations in these areas.

In experiment 3, we tested whether removing the mask altogether would show the sensitivity (or lack thereof) to changes in body viewpoint in the EBA and FBA. We repeated the design used in experiment 1, replacing the pattern mask with a blank, grayscale frame in the same color as the background.
Experiment 3

In experiment 3, we showed that bilateral EBA and right FBA undergo view-dependent adaptation to bodies and that the pattern of view invariance is significantly linear in EBA but not in FBA (although the patterns in these regions are highly similar in both shape and extent). Comparison of these results to the first two experiments shows that masking the stimuli substantially disrupts the analysis of bodies in EBA and FBA and specifically prevents these areas from establishing a representation that is partially invariant to viewing angle. Furthermore, both the Monte-Carlo analysis of experiment 3 data and the extremely small effect sizes indicated by the  $t$-test of raw adaptation in experiment 1 indicate that the different patterns of results from the two experiments cannot be attributed simply to differences in sample size.

The results of the functional ROI analyses were confirmed in a whole brain analysis, which showed that the linear response to increasingly large changes in viewpoint did not extend widely throughout visual cortex but was restricted to lateral occipitotemporal cortex. The absence of significant linear effects in the early visual cortex indicates that the origin of the effect is not retinotopic. Furthermore, the highly focal responses in lateral OTC, consistent with the location of EBA, suggest that the monotonic pattern is not a global attentional artifact, which would be expected to have widespread effects.

We also considered the close proximity of EBA to the motion sensitive hMT$. Because there is a potential confound between the angular difference between two images and the magnitude of any perceived motion attributable to that difference, could a motion percept account for the pattern of results shown in experiment 3? This possibility can be rejected on two counts: first previous studies (Kourtzi and Shiffrar 1997, 1999) have shown that apparent motion is reliably disrupted under similar interstimulus interval (ISI) and jitter conditions to those used here—indeed we used a larger spatial jitter than has been shown to be necessary. Furthermore, in this study (and during piloting), no motion percept was visible and none was reported. Second, we also obtained strong view-dependent adaptation in FBA, which is not spatially adjacent to hMT$ and is distant from other classical motion-selective areas, and hence the effect in this area is unlikely to be an artifact of apparent motion.

In this study, as in the previous two, behavioral performance was near ceiling on the target-detection task. This may seem at first to be paradoxical—performance might be expected to drop when masking impairs the formation of neural body representations. However, given the extensive practice participants had with the target images, it is likely that they performed the task on the basis of detecting simple local features that co-vari with the particular target pose used. Also, note that the target trials were not included in the fMRI data analyses so performance on these trials does not contribute to the patterns of adaptation observed.

General Discussion

The extent of viewpoint sensitivity in EBA and FBA is in close agreement with behavioral data examining the perception of synthetic objects (Kourtzi and Shiffrar 1999) and mannequins (Kourtzi and Shiffrar 1998) across changing viewpoints and with fMRI data examining object representations in the LO (Kourtzi et al. 2003). The studies present new evidence that the viewpoint invariance of cortical body representations extends to $\pm 30^\circ$ but to $<60^\circ$. Similarly, a recent imaging study examining sensitivity to changes in viewpoint for vehicles and animals (Andresen et al. 2009) found that the rotational limit for adaptation effects in these categories was also $<60^\circ$ in both LO and pFs. Thus our data on body representations in posterior and anterior body-selective regions appear broadly consistent with current data on general object-viewpoint selectivity.

One aim of the study was to compare the properties of the two body-selective regions of extrastriate cortex. The contrast analysis of the linear effect of angle was significant in EBA but not FBA, and this observation was supported by whole brain analysis. However, no interaction was found between EBA and FBA—indeed, the overall patterns of response were visually similar—and the patterns of adaptation suggest broadly similar tuning properties in these areas. Furthermore, in experiment 2, delaying the mask produced an increase in the response to the body stimuli in bilateral EBA but not FBA. Thus the additional 300 ms facilitated a measurable strengthening of the representation in EBA but not FBA. Taken together, these two findings hint at distinctions between EBA and FBA but do not provide strong evidence for a functional dissociation (cf. Hodzic et al. 2009; Taylor et al. 2007).

In addition to the effects seen in our a priori regions of interest, we also found a significant linear effect of viewpoint in a left-anterior portion of ITS (corresponding to Brodmann’s area 21). This region did not coincide with body-selective areas. One previous study (Vanrie et al. 2001) implicated a similar region in a subset of participants, in a task requiring the detection of local feature differences when making decisions about the similarity of objects viewed from different angles (our peak coordinates: $x$, $y$, $z$: $-60$, $-4$, $-17$; Vanrie et al. $x$, $y$, $z$: $-67$, $-6$, $-16$). Although a significant body of literature posits a primary role for nearby areas in the functional organization of semantic information about person identity and naming (Damasio et al. 1996; Tsuchiya et al. 2002), the Talairach coordinates reported are typically more lateral (Tsukiura et al. 2006, $x$, $y$, $z$: $-42$, $-11$, $22$; Grabowski et al. 2001: $-37$, $-14$, $-20$) compared with those identified here. Although the evidence remains tentative, our findings and those of Vanrie et al. (2001) are suggestive of a possible category-general role in the decoding of spatial rotation.

The unexpected effects of pattern masking in experiments 1 and 2 allow us to make some speculative inferences about the time course of neural processes of the body-selective regions we tested here. The pair of body images jointly produced a strong BOLD response in the ROIs, and the ability to detect a prespecified target in a single frame was intact. Therefore—to the extent that neural activity in these regions supports body perception (Moro et al. 2008; Peelen and Downing 2007; Pitcher et al. 2009; Urgesi et al. 2007a)—masking did not seem to abolish that activity. However, there was no fMRI adaptation effect in either region, even for an exact repetition (in different spatial locations) of a body image. It is likely, however, that the masks interrupted the later stages of neural activity in these regions, which perhaps are influenced by reciprocal interactions with higher level cortical areas (Fahrentz et al. 2007). Therefore a working hypothesis is that the BOLD adaptation effect observed in experiment 3 was driven...
primarily by these later aspects of the neural response, which were interrupted in experiments 1 and 2. One (indirect) line of supporting evidence is that event-related potential responses to faces show repetition effects on later but not earlier face-selective components (Boehm and Paller 2006). (Of course a further possibility is that neuronal and BOLD dynamics are coupled in a complex way such that a preserved neuronal adaptation effect is not transmitted via concomitant changes in the hemodynamic response; cf. Sawamura et al. 2006). An additional useful test would be to present masks at still further poststimulus delays, with the prediction that fMRI adaptation effects would return when later neural responses were no longer influenced by masking. However, this would require significant changes to the trial design used here, confounding comparisons with the present experiments.

One additional consideration arises from the trial structure in experiments 1 and 2: Although both frames were backward masked, frame 2 was also effectively forward masked with a −400-ms SOA (i.e., by the mask following frame 1). Because the design of experiment 3 did not directly address this issue, it is necessary to consider the possible implications of forward masking on the abolition of adaptation. Perrett et al. (2009) showed the temporally extended effects of forward masking on the firing rate of category selective neurons in macaque STS. Masking was shown to attenuate firing rates evoked by target stimuli for as long as 400 ms following 125-ms mask presentations. However, the effectiveness of forward masking of this type was directly related to the similarity (determined independently by stand-alone firing rate) between the mask and the target. In this case, the finely scrambled patterns used in the masks carry little object structure and so probably do not effectively drive activity in the body-selective regions tested here (see also Downing et al. 2001). This suggests that scrambled images may not be very effective forward masks on neural activity in EBA and FBA; however, more direct manipulations would be required to separately assess the contributions of forward and backward masking on adaptation in these areas. Finally, task performance (as indexed by hit rate) was similarly high in experiments 1 and 2, suggesting that participants were easily able to identify the pose (structure) of the second frame in the majority of target trials. This also suggests that the effect of forward masking on explicit perception was very weak.

Our understanding of high-level visual areas benefits from considering category-selective response properties not as fixed, but as evolving over time. Single-unit recording studies in the macaque IT cortex suggest a progressive emergence of category selectivity, with greater selectivity during later spiking (~240 ms following stimulus onset) compared with early transient responses at ~130 ms (Tamura and Tanaka 2001). Convergent findings are reported by Sugase et al. (1999). These authors reported a faster emergence of global compared with fine category information, with former being maximal during the spiking phase and the latter maximal during the decay phase. Significant information transmission was sustained to ~500 ms following onset. In human neuroimaging studies, comparisons of short-term and long-term adaptation indicate that the adaptation properties of face selective regions of visual cortex (i.e., OFA/FFA) may differ depending on the duration of the adaptation period (Fang et al. 2007; Kovács et al. 2008). Furthermore, the magnitude of both repetition suppression (in LO) and suppression and enhancement (in pFs/FFA) have been shown to decrease as lag-time increases. This effect is stable across multiple stimulus categories, suggesting that such temporal contingencies may be a general property of object sensitive regions (Henson 2003; Henson et al. 2000, 2002). This study adds to this picture, indicating that examining adaptation effects, coupled with systematic manipulation of stimulus masking, may prove a useful tool to further investigate category-selective regions of the extrastriate cortex. This combination has the potential to add information about how representations evolve over time and how this relates to neurophysiological evidence and to the temporal demands of real world viewing tasks.

In summary, we showed viewpoint selective tuning to bodies in EBA and FBA and found some provisional evidence for a category-independent effect in the left aITS. The close agreement between the viewpoint-dependent tuning in EBA and FBA reported here and that seen for other object categories in adjacent regions of the LOC suggests analogous representational processes, operating across stimulus classes. We interpret the abolition of view dependent fMRI adaptation by pattern masking as evidence that the viewpoint-dependent fMRI adaptation effect seen in experiment 3 is driven by later waves of neuronal responses in the regions of interest.

GRANTS

This research was supported by a Biotechnology and Biological Sciences Research Council grant to P. E. Downing, an Economic and Social Research Council (ESRC) studentship to J. C. Taylor, and a Wales Institute of Cognitive Neuroscience grant to P. E. Downing and J. C. Taylor.

REFERENCES


