

COGNITIVE NEUROSCIENCE

Activity patterns in the category-selective occipitotemporal cortex predict upcoming motor actions

Jason P. Gallivan,¹ Craig S. Chapman,² D. Adam McLean,³ J. Randall Flanagan¹ and Jody C. Culham^{3,4}

¹Centre for Neuroscience Studies, Department of Psychology, Queen's University, Kingston, ON, Canada

²Faculty of Physical Education and Recreation, W1-34 Van Vliet Centre, University of Alberta, Edmonton, AB, Canada

³Brain and Mind Institute, Natural Sciences Centre, University of Western Ontario, London, ON, Canada

⁴Department of Psychology, University of Western Ontario, London, ON, Canada

Keywords: action, decoding, fMRI, occipitotemporal cortex, planning, ventral visual stream

Abstract

Converging lines of evidence point to the occipitotemporal cortex (OTC) as a critical structure in visual perception. For instance, human functional magnetic resonance imaging (fMRI) has revealed a modular organisation of object-selective, face-selective, body-selective and scene-selective visual areas in the OTC, and disruptions to the processing within these regions, either in neuropsychological patients or through transcranial magnetic stimulation, can produce category-specific deficits in visual recognition. Here we show, using fMRI and pattern classification methods, that the activity in the OTC also represents how stimuli will be interacted with by the body – a level of processing more traditionally associated with the preparatory activity in sensorimotor circuits of the brain. Combining functional mapping of different OTC areas with a real object-directed delayed movement task, we found that the pre-movement spatial activity patterns across the OTC could be used to predict both the action of an upcoming hand movement (grasping vs. reaching) and the effector (left hand vs. right hand) to be used. Interestingly, we were able to extract this wide range of predictive movement information even though nearly all OTC areas showed either baseline-level or below baseline-level activity prior to action onset. Our characterisation of different OTC areas according to the features of upcoming movements that they could predict also revealed a general gradient of effector-to-action-dependent movement representations along the posterior–anterior OTC axis. These findings suggest that the ventral visual pathway, which is well known to be involved in object recognition and perceptual processing, plays a larger than previously expected role in preparing object-directed hand actions.

Introduction

An influential view of the primate visual system argues for a neuro-anatomical division between the brain areas that support visual object recognition, located ventrally in the occipitotemporal cortex (OTC), and the brain areas that support action planning and control, located dorsally in the frontoparietal cortex (Goodale & Milner, 1992). Under this view, the ventral pathway is specialised for extracting visual details of stimuli in the environment (e.g. size, shape, and color) and, consistent with the selectivity in the OTC for different categories of stimuli (e.g. objects, faces, scenes, and bodies) (see Grill-Spector & Malach, 2004 for a review), providing a visual analysis of ‘what’ is being viewed. In contrast, the dorsal pathway is specialised for transforming visual details related to real-world metrics (e.g. size and orientation) into motor commands to move the body’s effectors (e.g. eyes and limbs), and, in this sense, is thought to provide a sensorimotor analysis of ‘how’ to interact with objects.

Despite the utility of this theoretical distinction and its use as a guiding neural framework for studies involving human functional magnetic resonance imaging (fMRI) and monkey neurophysiology (Kravitz *et al.*, 2011), many everyday behaviors require direct communication between ventral-stream and dorsal-stream processes (e.g. recognising an object and generating the appropriate motor commands). At the neuroanatomical level, this cross-talk is facilitated by direct interconnections between the anterior intraparietal (AIP) area involved in grasping and the OTC structures (namely, the inferotemporal cortex) involved in object processing (Borra *et al.*, 2008, 2010). Whereas, intuitively, this connectivity allows parietal areas to gain access to object-related information processed by the OTC, the opposite may also be true; OTC structures may have access to ongoing movement-related computations occurring in the parietal cortex.

One intriguing possibility, suggested by fMRI activation in the OTC when individuals perform hand and arm movements (e.g. Astafiev *et al.*, 2004; Dinstein *et al.*, 2007; Filimon *et al.*, 2009; Orlov *et al.*, 2010), is that OTC areas may receive efference copies of motor commands prior to movement. One purpose of these efference copy signals could be to anticipate the proprioceptive, tactile and/or visual consequences of movement (Astafiev *et al.*, 2004; David *et al.*, 2007; Orlov *et al.*, 2010), a notion closely related to the types

Correspondence: Jason P. Gallivan, as above.

E-mail: jasongallivan@gmail.com

Received 8 March 2013, accepted 10 March 2013

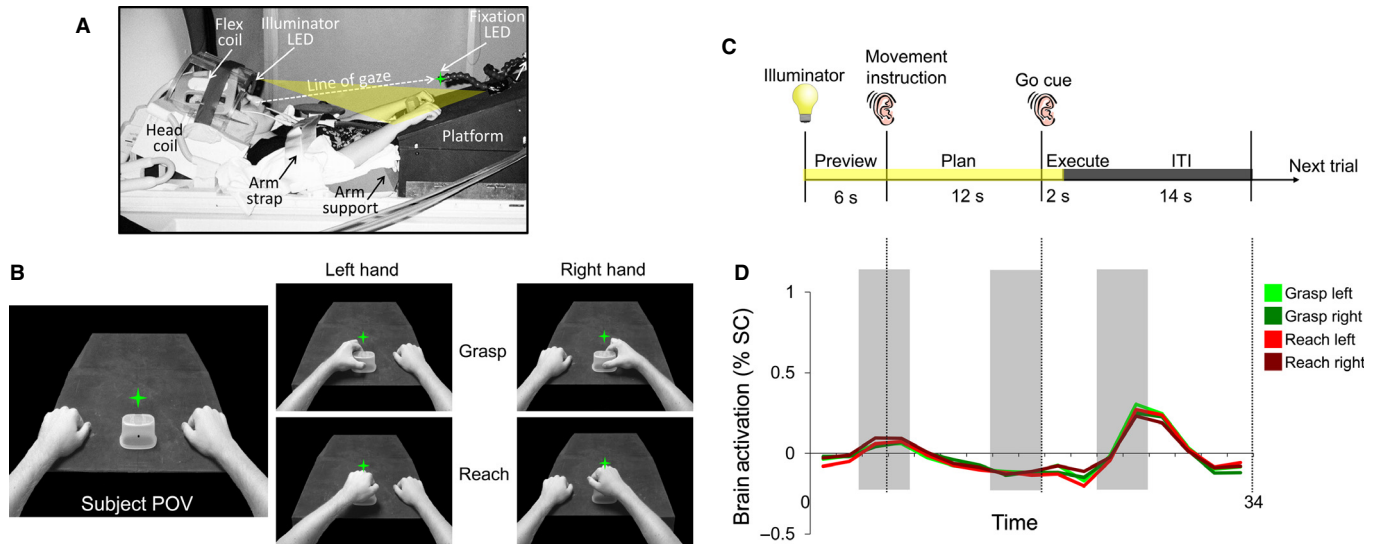


FIG. 1. Methods. (A) Subject setup shown from side view. (B) Experimental setup (left) and hand movements (right) for the motor experiment, shown from the subject's point of view (POV). Note that the centrally located target object never changed position from trial to trial. The green star represents the fixation LED. Left: the hand(s) are positioned at their starting locations. (C) Timing of event-related trials. Trials began with the 3D object being illuminated while the subject maintained fixation (preview phase; 6 s). Subjects were then instructed via headphones to perform one of the four movements in B. This auditory cue initiated the plan phase of the trial. Following a fixed delay interval (12 s), subjects were then cued to perform the instructed hand movement (initiating the execute phase). Two seconds after the go cue, vision of the workspace was extinguished, and participants waited for the following trial to begin (14 s; ITI). (D) Averaged neural activity, time-locked to events in C, from the left pFs functionally localised by the contrast Objects > Scrambled in the separate testing session. MVPA was performed on single trials, based on the windowed average (4 s) of the percentage signal change (% SC) activity corresponding to the three different time periods denoted by each of the gray shaded bars (each corresponding to activity elicited from the three distinct trial phases: preview, plan, and execute).

of forward-state estimations considered to be essential in motor planning and control (Wolpert & Flanagan, 2001; Wolpert *et al.*, 2011), and, to date, largely thought to lie within the purview of sensorimotor-related brain structures (Andersen & Cui, 2009). A more parsimonious explanation of the fMRI findings, however, is that any 'motor'-related activity in OTC merely reflects the visual and/or proprioceptive sensory feedback responses that accompany movement. That is, in accordance with the current understanding of the processing in the ventral stream being primarily perceptual in nature, the OTC might simply proprioceptively and/or visually track the position of the limbs in space (Downing *et al.*, 2001; Orlov *et al.*, 2010; Downing & Peelen, 2011). One way to disentangle these possibilities is to determine whether, prior to movement onset and the accompanying proprioceptive, tactile and visual feedback responses, OTC activity, like that in the frontoparietal cortex (e.g. Gallivan *et al.*, 2011b), can be used to predict upcoming motor actions. Such a result would suggest a much larger than previously expected role for the OTC in planning hand-related behaviors.

Materials and methods

Although the current findings have not been presented before, here we performed completely new analyses upon a dataset that was previously used to examine sensorimotor representations contained only in the frontoparietal cortex (Gallivan *et al.*, 2013). Here, with the addition of object-selective, face-selective, body-selective and scene-selective localisers along with a new set of analyses, our aim was much different: to investigate whether action intentions (the hand action to be performed and/or the limb to be used) can be decoded from various localiser-defined regions of interest (ROIs) in the OTC.

In brief, our study required participants to perform grasp or reach movements (i.e. two different types of hand action) with the right or left hand (i.e. using two different effectors) towards a centrally

located target object in each trial (Fig. 1B). The sequence of events for each trial was as follows. Participants were first cued to the specific action to be performed, and then, following a delay period, they executed the action (Fig. 1C). Importantly, the target object was always presented throughout the full trial sequence and positioned in the same central location throughout the length of the experiment, such that visual presentation of the object remained constant both within and between trials. The delayed timing of this experiment allowed us to isolate the pre-movement activity patterns (plan phase) in each OTC ROI from the transient visual (preview phase) and movement execution responses (execute phase) that directly accompany presentation of the target object and action initiation, respectively (Fig. 1D).

We used multi-voxel pattern analysis (MVPA) to examine whether we could predict upcoming motor actions from a range of well-established areas in the OTC. In particular, we examined the pre-movement activity from each of the following category-selective areas: (i) object-selective areas – the lateral occipital (LO) area and the posterior fusiform sulcus (pFs); (ii) face-selective areas – the occipital face area (OFA) and the fusiform face area (FFA); (iii) body-selective areas – the extrastriate body area (EBA) and the fusiform body area (FBA); and (iv) a scene-selective area – the parahippocampal place area (PPA). We reasoned that, if we could decode, prior to movement, motor-specific aspects of the upcoming object-directed task from some of these areas (i.e. the intention to perform a grasping vs. reaching hand action and/or the actual limb to be used in the movement), then we would be able to rule out previous interpretations of the action-related responses in the OTC as being purely related to sensory feedback responses (e.g. proprioceptive, tactile, and visual) (e.g. Astafiev *et al.*, 2004; Filimon *et al.*, 2009; Cavina-Pratesi *et al.*, 2010; Orlov *et al.*, 2010). Above-chance decoding prior to movement onset would provide an unequivocal demonstration of a predictive component to the action-related

responses in the OTC. At a more general level, a central goal of this investigation was to update the currently limited understanding of the role that the OTC structures serve in sensorimotor processing.

Subjects

Eleven right-handed volunteers (five females; mean age, 25.7 years), recruited from the University of Western Ontario (London, ON, Canada), participated in a motor-related experiment and additional localiser experiments, conducted on separate testing days. The experiments were undertaken with the understanding and written consent of each subject, obtained in accordance with ethical standards set out by the Declaration of Helsinki (1964) and with procedures approved by the University of Western Ontario's Health Sciences Research Ethics Board (ethics review number: 13507).

Motor experiment

The complete methods for the motor experiment have been previously described in detail (Gallivan *et al.*, 2013). Here, we provide a more concise description of the methods relevant for the new analyses.

Setup and apparatus

Subjects were scanned in a head-tilted configuration (allowing direct viewing of the hand workspace) while they performed the object-directed delayed movement task (see Fig. 1 for setup and timing). During the experiment, the target object was illuminated from the front with a bright white light-emitting diode (LED) attached to a flexible plastic stalk (Loc-Line; Lockwood Products, Lake Oswego, ON, Canada), located over the participant's left shoulder. Each trial was preceded by a period when participants were in complete darkness. Experimental timing and lighting were controlled with in-house software created with MATLAB (The Mathworks). To control for eye movements during MRI scanning, a small green fixation LED attached to a flexible plastic stalk was placed above and behind the target object, and participants were required to always foveate the fixation LED during experimental testing. Considerable care was taken to ensure that the green fixation LED was as close to the object's location as possible (in the dimensions of both depth and height) without actually obstructing any of the hand movements performed. Throughout the experiment, the subject's eye and arm movements were recorded with two magnetic resonance-compatible infrared-sensitive cameras (borecameras; MRC Systems), each attached to a flexible plastic stalk (note that both of these borecameras are not shown in Fig. 1A). The videos captured during the experiment were then analysed off-line in order to exclude error trials from fMRI analysis.

For each trial, subjects were required to perform one of four actions upon the target object: (i) reach towards and precision grasp the object (between their thumb and index finger) with the left hand, without lifting ['grasp left' auditory command (GraspL)]; (ii) reach towards and precision grasp the object with the right hand, without lifting ['grasp right' auditory command (GraspR)]; (iii) reach towards and manually touch the top of the object (without hand pre-shaping) with the left hand ('reach left' auditory command (ReachL)); and (iv) reach towards and manually touch the top of the object with the right hand ('reach right' auditory command (ReachR)). Other than the execution of these hand actions, participants were instructed to keep their hands still and in pre-specified 'home' positions throughout all other phases of the

trial [i.e. preview phase, plan phase, and intertrial interval (ITI); Fig. 1C].

Experiment design and timing

To extract the sustained planning response from the transient visual and motor execution responses, we used a slow event-related delay paradigm with 34-s trials, each consisting of three distinct phases: 'preview', 'plan' and 'execute' (Fig. 1C). Each trial began with the preview phase, in which the subject's workspace was illuminated, revealing the centrally located target object. After 6 s of the preview phase, subjects were given an auditory cue (duration, 0.5 s; the specific auditory instructions were mentioned above), informing them of the upcoming movement required; this cue marked the onset of the plan phase. Although there were no visual differences between the preview and plan phase portions of the trial (that is, the target object was always visually present), only in the plan phase did participants have the information necessary (i.e. type of movement to be performed) to prepare the upcoming action. After 12 s of the plan phase, a 0.5-s auditory beep cued participants to immediately execute the planned action (for a duration of ~2 s), initiating the execute phase of the trial. Two seconds following the beginning of this go cue, the illuminator was turned off, providing the cue for subjects to return the hand to its respective starting position. After the illuminator was turned off, subjects waited in the dark while maintaining fixation for 14 s, allowing the blood oxygen level-dependent response to return to baseline prior to the next trial (ITI phase). The four trial types, with five repetitions per condition (20 trials in total), were randomised within a run and balanced across all runs, so that each trial type was preceded and followed equally often by every other trial type across the entire experiment.

Localiser experiments

In a separate testing session, each subject from the motor experiment participated in two types of functional localiser runs. The first included stimulus blocks of black-and-white photographs consisting of faces, scenes, objects, and scrambled versions of these stimuli (modified from Cant & Goodale, 2007, 2011) [see faces, places and objects (FPO) localiser details below]. The second included stimulus blocks of color photographs consisting of headless bodies, tools, non-tool objects, and scrambled versions of these stimuli (Valyear & Culham, 2010) [see bodies, objects and tools (BOT) localiser details below]. The purpose of these localiser scan sessions was to independently identify well-documented visual-perceptual ROIs involved in object-selective, face-selective, scene-selective and body-selective processing, and then examine, using the motor-related data from these same subjects, whether upcoming hand movements could be decoded from the pre-movement activity patterns in each of these category-specific areas (see Fig. 2A for the category-selective ROIs localised in one representative subject's brain).

In both localisers, participants were required to maintain fixation on a dot (a small black circle) superimposed on the center of each image. Each image subtended ~15° of visual angle. Photographs were repeated across runs, and the stimulus and epoch orders were pseudo-randomised and balanced across runs. To encourage participants to maintain attention throughout the localiser scans, subjects performed a one-back task throughout, whereby responses were made, via a right-handed button press, whenever two successive photographs were identical. In addition to these localisers, we also

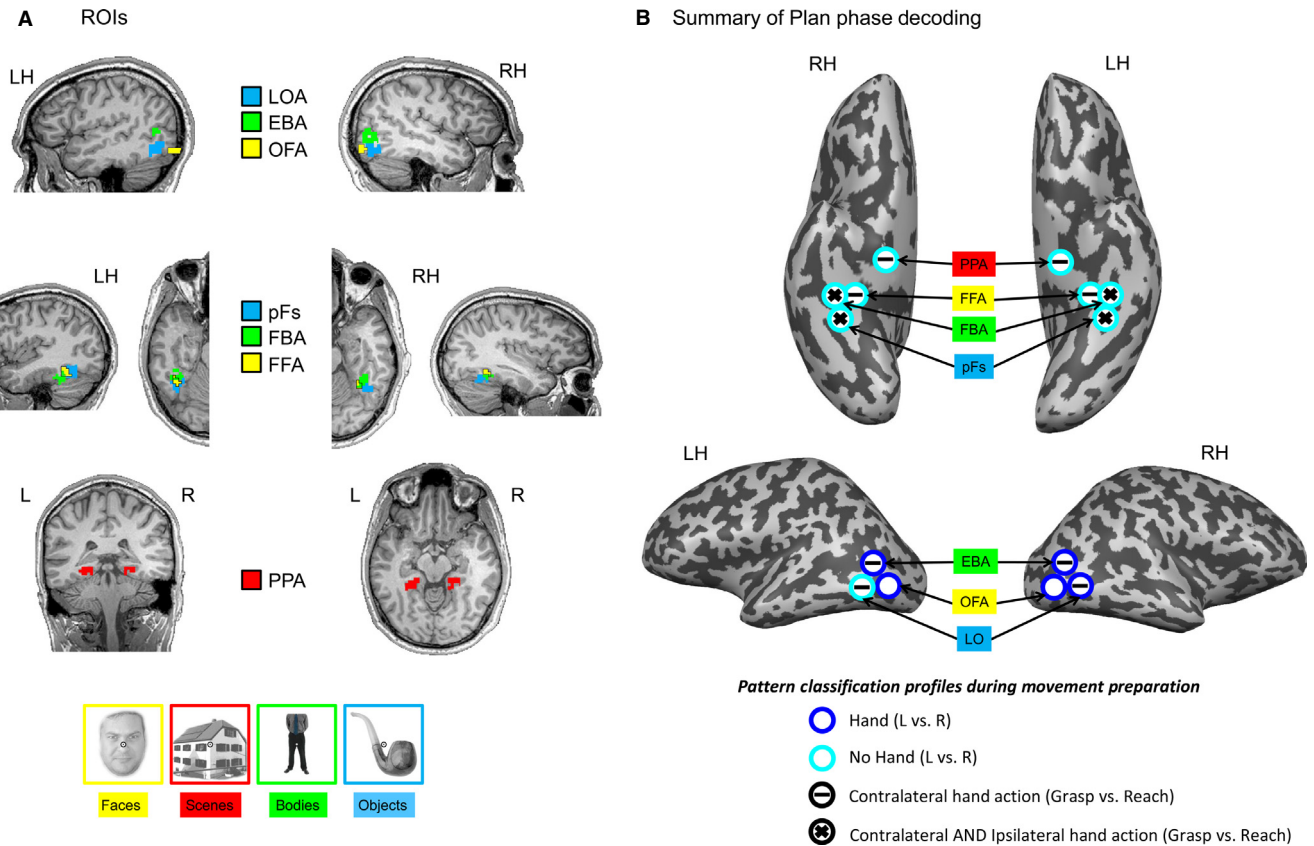


FIG. 2. Cortical organisation of category-selective ROIs and summary of decoding results. (A) Category-selective ROIs (for objects, faces, bodies, and scenes) overlaid on the transverse and coronal cortical slices of a representative subject ($t = 3$, $P < 0.005$, corrected). Blue (object-selective) ROIs were defined by the contrast Objects > Scrambled. Yellow (face-selective) ROIs were defined by the conjunction contrast [(Faces > Objects) AND (Faces > Scenes) AND (Faces > Scrambled)]. Green (body-selective) ROIs were defined by the conjunction contrast [(Bodies > Objects) AND (Bodies > Tools) AND (Bodies > Scrambled)]. Red (scene-selective) ROIs were defined by the conjunction contrast [(Scenes > Objects) AND (Scenes > Faces) AND (Scenes > Scrambled)]. Owing to the partial overlap of both face-selective and body-selective activity and object-selective and body-selective activity in the ventral-temporal cortex, face-selective activity (in yellow) is outlined in black to delineate the ROIs (for further details, see Materials and methods). (B) Summary of plan phase decoding findings. General anatomical locations of the category-selective ROIs are displayed on one representative subject's inflated hemispheres (top, ventral views of the hemispheres; bottom, side views of the hemispheres). The outline and center symbols of each ROI are coded according to the general pattern of pair-wise discriminations made during movement planning (with respect to the black asterisks in Figs 3–6); see key at the bottom for classification profiles. Note that, in order for the activity in an ROI to be considered hand-selective (dark blue-bordered ROIs), it must significantly decode the left limb vs. right limb for grasp and/or reach hand actions. L, left; R, right; LH, left hemisphere; RH, right hemisphere.

collected a high-resolution anatomical image from each of the participating subjects.

FPO localiser

This localiser was performed in all 11 subjects from the motor experiment. Stimuli were organised into separate 16-s blocks, with 16 photographs per block, presented at a rate of 400 ms per photograph with a 600-ms inter-stimulus interval. Each run lasted for 7 min 30 s, and was composed of four stimulus blocks per condition, with each stimulus block separated by a scrambled block. Two fixation/baseline blocks (20 s) were placed at the beginning and end of each run. All subjects participated in at least three of these localiser scans. Each stimulus block included two repeated photographs.

BOT localiser

This localiser was performed in nine of 11 subjects from the motor experiment (two of the 11 subjects were unavailable to participate in a separate testing session for the BOT localiser). Stimuli were

organised into separate 16-s blocks, with 18 photographs per block, presented at a rate of 400 ms per photograph with a 490-ms inter-stimulus interval. Each run lasted 7 min 30 s, and was composed of six stimulus blocks per condition, seven scrambled blocks, and two fixation/baseline blocks (20 s) placed at the beginning and end of each run. Stimulus blocks were organised into sets of three, separated by scrambled blocks, and balanced for prior-block history within a single run. All subjects participated in at least three of these localiser scans. Each stimulus block included either three or four repeated photographs, balanced across conditions.

Eye-tracking control experiment

Four of the participants completed an additional testing session outside of the magnetic resonance imaging (MRI) scanner, in which their eye fixations were monocularly monitored with an Eye-Link II eye-tracker (SR Research). The methods, analyses and results of this eye-tracking control experiment have been previously described in detail (Gallivan *et al.*, 2013). Here, we only note that this control study revealed negligible evidence of eye movements in our participants. Therefore, subtle differences in eye stability between the trial types is

unlikely to account for any accurate decoding performance found throughout the OTC.

Magnetic resonance image acquisition and preprocessing

Subjects were scanned with a 3-T Siemens TIM MAGNETOM Trio MRI scanner. The T1-weighted anatomical image was collected with an ADNI MPRAGE sequence (time to repetition, 2300 ms; time to echo, 2.98 ms; field of view, $192 \times 240 \times 256$ mm; matrix size, $192 \times 240 \times 256$; flip angle, 9° ; 1-mm isotropic voxels). Functional MRI volumes were collected with a T2*-weighted single-shot gradient-echo echo-planar imaging acquisition sequence [time to repetition, 2000 ms; time to echo, 30 ms; field of view, 240×240 mm; matrix size, 80×80 ; flip angle, 90° ; 3-mm isotropic voxels; and acceleration factor (integrated parallel acquisition technologies) of 2, with generalised auto-calibrating partially parallel acquisitions reconstruction]. Each volume comprised 34 contiguous (no gap) oblique slices acquired at an approximately 30° caudal tilt with respect to the plane of the anterior and posterior commissure (ACPC), providing nearly whole brain coverage. For the motor experiment we used a combination of imaging coils to achieve a good signal/noise ratio and to enable direct viewing of the workspace without mirrors or occlusion. Specifically, we tilted ($\sim 20^\circ$) the posterior half of the 12-channel receive-only head coil (six channels), and suspended a four-channel receive-only flex coil over the anterior–superior part of the head (Fig. 1A). In the functional localiser experiments, subjects were scanned with a conventional setup (i.e. stimuli projected onto a two-dimensional screen and viewed through a mirror), with a 32-channel receive-only head coil. The cortical surface from one subject was reconstructed from a high-resolution anatomical image, a procedure that included segmenting the gray and white matter and inflating the boundary surface between them. This inflated cortical surface was only used to summarise the experimental findings (Fig. 2B). For analyses, all data (from the motor and localiser experiments) were spatially aligned to the corresponding individual's high-resolution anatomical image collected during the localiser experiments. All preprocessing and univariate analyses were performed with BRAIN VOYAGER QX version 2.12 (Brain Innovation, Maastricht, The Netherlands).

Following slice scan–time correction, three-dimensional (3D) motion correction (such that each volume was aligned to the volume of the functional scan closest in time to the anatomical scan), high-pass temporal filtering (four cycles per run for the motor experiment, and three cycles per run for the localiser experiments), and functional-to-anatomical co-registration, functional and anatomical images were rotated such that the axial plane passed through the ACPC space. Other than the sinc interpolation inherent in all transformations, no additional spatial smoothing was performed on the data.

For each participant, functional data from each session were screened for motion and/or magnet artefacts by examining the time-course movies and the motion plots created with the motion correction algorithms. None of the runs revealed head motion that exceeded 1 mm of translation or 1° of rotation. Error trials – trials where the participant fumbled with the object (one trial, one participant), performed the incorrect instruction (three trials, two participants), or contaminated the plan phase data by slightly moving their limb or eyes or by performing the action before the go cue (eight trials, four participants) – were identified off-line from the videos recorded during the session, and were excluded from analysis by assigning these trials predictors of no interest.

General linear models

For the localiser experiment data, we used a general linear model with predictors created from boxcar functions convolved with the Boynton (Boynton *et al.*, 1996) hemodynamic response function. A boxcar function was aligned to the onset of each stimulus block, with its duration dependent on stimulus block length. The baseline/fixation epochs were excluded from the model; therefore, all regression coefficients (betas) were defined relative to the activity during these time points. In addition, the time-course for each voxel was converted to percentage signal change before application of the random effects general linear model.

ROIs

For each subject, the following procedure was used to select each ROI. The most significantly active voxel, or peak, was first identified on the basis of a particular contrast (or conjunction), constrained by the anatomical location expected from previous reports (see below for details). Statistical thresholds were then set to a determined minimum ($t = 3$, $P < 0.005$, cluster threshold corrected at $P < 0.05$), and the activity up to $(15 \text{ mm})^3 = 3375 \text{ mm}^3$ around the peak was selected. This approach ensured that regions were selected objectively, that a similar number of voxels were included within each ROI, and that regions could be reliably segregated from adjacent activations. We chose these maximum ROI sizes (i.e. up to 3375 mm^3 of activity) to ensure the inclusion of a sufficient number of functional voxels (up to 5^3 functional voxels) for pattern classification (an important consideration).

For each individual, when possible, we defined the LO, pFs, OFA, FFA and PPA in both hemispheres by using the FPO localiser data, and the EBA and FBA in both hemispheres by using the BOT localiser data.

Object-selective activity (LO and pFs) was localised on the basis of the contrast Objects > Scrambled. LO was defined around the peak voxel of activity near the lateral occipital sulcus (Malach *et al.*, 1995; Grill-Spector *et al.*, 1999, 2001). The pFs was defined around the peak voxel of activity in the posterior aspect of the fusiform gyrus, extending into the occipitotemporal sulcus (Grill-Spector *et al.*, 1999, 2001). Importantly, the activity selected for LO and the pFs did not overlap.

Face-selective activity (OFA and FFA) was localised on the basis of the conjunction contrast [(Faces > Scrambled) AND (Faces > Scenes) AND (Faces > Objects)]. The OFA was defined around the peak voxel of activity in the inferior occipital gyrus (Puce *et al.*, 1996; Gauthier *et al.*, 2000; Haxby *et al.*, 2000; Calder & Young, 2005), inferior and more medially located than the activity for LO. The FFA was defined around the peak voxel of activity in the fusiform gyrus (Kanwisher *et al.*, 1997), generally anterior to the pFs region.

Scene-selective activity (PPA) was localised on the basis of the conjunction contrast [(Scenes > Scrambled) AND (Scenes > Faces) AND (Scenes > Objects)]. The PPA was defined around the peak voxel of activity located medially along the collateral sulcus and parahippocampal gyrus (Epstein & Kanwisher, 1998).

Body-selective activity (EBA and FBA) was selected on the basis of the conjunction contrast [(Bodies > Scrambled) AND (Bodies > Tools) AND (Bodies > Objects)]. The EBA was defined around the peak voxel of activity in the posterior inferior temporal sulcus/middle temporal gyrus (Downing *et al.*, 2001; Peelen & Downing, 2005c), superior to LO and OFA. The FBA was defined around the peak voxel of activity in the fusiform gyrus (Peelen &

Downing, 2005a; Schwarzlose *et al.*, 2005), generally anterior to the pFs region.

Note that, here, we define a conjunction contrast as a Boolean AND, such that, for any one voxel to be flagged as significant, it must show a significant difference for each of the constituent contrasts. As such, the stringency of this selection procedure ensured that adjacent ROIs – when defined by use of the same localiser data – did not overlap (for example, the activities selected for area LO and the OFA did not overlap). Note, however, that because the FBA was defined by use of the BOT localiser and the FFA and the pFs were defined by use of the FPO localiser, we did find some degree of overlap between their activations, consistent with previous investigations at this same imaging resolution [3-mm iso-voxels; for an example of FBA and FFA overlap, see Peelen & Downing (2005a); but see Schwarzlose *et al.* (2005) for their separation at higher imaging resolutions]. In particular, we found that FBA voxels overlapped with FFA voxels in the left hemispheres of two (of nine) subjects (the FBA shared the following percentages of its total voxels with the FFA: subject 2, 25.00%; and subject 7, 16.67%) and in the right hemispheres of eight (of nine) subjects (the FBA shared the following percentages of its total voxels with the FFA: subject 1, 42.86%; subject 2, 23.08%; subject 3, 25.00%; subject 4, 55.00%; subject 5, 36.54%; subject 6, 56.00%; subject 7, 35.29%; and subject 8, 64.10%). We also found that FBA voxels overlapped with pFs voxels in the left hemispheres of two (of nine) subjects (the FBA shared the following percentages of its total voxels with the pFs: subject 1, 20.00%; and subject 6, 59.09%) and in the right hemispheres of four (of nine) subjects (the FBA shared the following percentages of its total voxels with the pFs: subject 3, 28.57%; subject 6, 32.00%; subject 8, 15.38%; and subject 9, 13.33%).

For details about Talairach coordinates (when ROIs were localised at the group level), cluster sizes (averaged across subjects), and the total number of subjects in which each of these category-selective ROIs could be localised, see Table 1.

TABLE 1. ROIs with corresponding Talairach coordinates (mean x , y and z centre of mass)

ROI name	Tailarach coordinates			ROI size		No. of subjects
	x	y	z	mm ³	No. of voxels	
Object-selective ROIs						
Left LO	-43	-67	-9	2699	100	11/11
Right LO	43	-65	-3	2361	87	11/11
Left pFs	-31	-45	-15	1942	72	11/11
Right pFs	28	-45	-15	1506	56	11/11
Face-selective ROIs						
Left OFA	-39	-78	-7	802	30	9/11
Right OFA	36	-72	-9	617	23	10/11
Left FFA	-38	-39	-21	512	19	10/11
Right FFA	39	-43	-18	618	23	11/11
Scene-selective ROIs						
Left PPA	-22	-39	-7	632	23	11/11
Right PPA	24	-41	-7	878	33	11/11
Body-selective ROIs						
Left EBA	-51	-73	7	888	33	9/9
Right EBA	46	-71	6	1296	48	8/9
Left FBA	-42	-41	-21	390	14	7/9
Right FBA	33	-43	-17	648	24	9/9

Mean ROI sizes across subjects from ACPC data (in mm³ and functional voxels). No. of subjects denotes the number of subjects in which the corresponding ROI could be localised.

MVPA

Support vector machine (SVM) single-trial classification

MVPA was performed with a combination of in-house software (using MATLAB) and the Princeton MVPA Toolbox for MATLAB (<http://code.google.com/p/princeton-mvpa-toolbox/>), with an SVM binary classifier (libSVM, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>). The SVM model used a linear kernel function and a constant cost parameter ($C = 1$). We used a 'leave-one-trial-pair-out' N -fold cross-validation to test the accuracy of the SVM classifiers. We performed this $N - 1$ cross-validation procedure until all trial pairs were tested, and then averaged across N iterations in order to produce a classification accuracy measure for each ROI, trial phase, pair-wise discrimination, and subject (Duda *et al.*, 2001). We statistically assessed decoding significance across participants with a two-tailed t -test vs. 50% chance decoding. To control for multiple comparisons, a false discovery rate (FDR) correction of $q \leq 0.05$ was applied, based on all t -tests performed within each time phase (Benjamini & Yekutieli, 2001).

Classifier inputs

To provide inputs for the SVM classifier, the percentage signal change activity was extracted from a windowed average of the time-course at a time point of interest (i.e. preview, plan or execute phase) with respect to an average of the time-course activity at a common baseline, for each voxel in the ROI (a procedure similar to that used for extracting event-related average time-courses). The baseline window was defined as volume -1, a time point prior to the current trial and avoiding contamination from responses of the previous trial. For the preview phase time points, we extracted the mean of volumes 3 and 4; time points corresponding to the peak of the visual transient response (see percentage signal change time-courses in Figs 3–6). For the plan phase – the time points of critical interest – we extracted the average of volumes 8 and 9, the final two volumes of the plan phase, and, importantly, a two-volume window prior to the subject hearing the auditory cue to initiate a movement. Finally, for the execute phase time points, we extracted the average of volumes 12 and 13, time points generally corresponding to the peak (and activity prior to the peak) of the transient movement response, following the subject's action (see percentage signal change time-courses in Figs 3–6). These time points extracted for pattern classification directly follow from our previous work (Gallivan *et al.*, 2011a,b, 2013). Following the extraction of each trial's percentage signal change activity, these values were rescaled between -1 and +1 across all trials for each individual voxel within an ROI (Misaki *et al.*, 2010).

Cross-decoding

In order to test whether an SVM classifier trained to discriminate between two trial types could then be used to accurately predict trial identity when tested on a different set of trials (i.e. cross-decode), instead of using the $N - 1$ cross-validation procedure (implemented above), we used all of the available single-trial data for both classifier training and testing (i.e. one train-and-test iteration) (Smith & Muckli, 2010; Gallivan *et al.*, 2011a). Cross-decoding accuracies for each subject were computed by averaging together the two accuracies generated by using each pair of trial types for classifier training and testing (for example, when testing for action-specific, limb-independent effects, right-hand trials were used to train the classifier in one analysis when left-hand trials were used for testing, and then

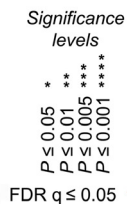


Fig. 3. Movement plan decoding from object-selective ROIs (LO and pFs). Each ROI is associated with three plots of data. Top: percentage signal change (% SC) time-courses. The activity in each plot is averaged across all voxels within each ROI and across subjects. Vertical lines correspond to the onset of the preview, plan and execute phases of each trial (from left to right). MVPA was performed on single trials, based on the windowed average (4 s) of the percentage signal change activity corresponding to the three different time points denoted by each of the gray shaded bars. Bottom left: corresponding decoding accuracies are shown for each time phase of the trial. Note that accurate classification is largely attributable to the spatial activity patterns of different planned movement types, and not to differences in the overall signal amplitude responses within each ROI (that is, time-courses overlap substantially during the plan phase). Bottom right: cross-decoding accuracies are shown for each time phase (preview, plan, and execute). For details on the cross-decoding approach, see Materials and methods and Results. Error bars represent standard error of the mean across subjects. Solid black lines represent chance accuracy level (50%). Black asterisks indicate statistical significance with two-tailed t -tests across subjects with respect to 50%. Red asterisks indicate statistical significance based on an FDR correction of $q \leq 0.05$, based on all t -tests performed within each time phase.

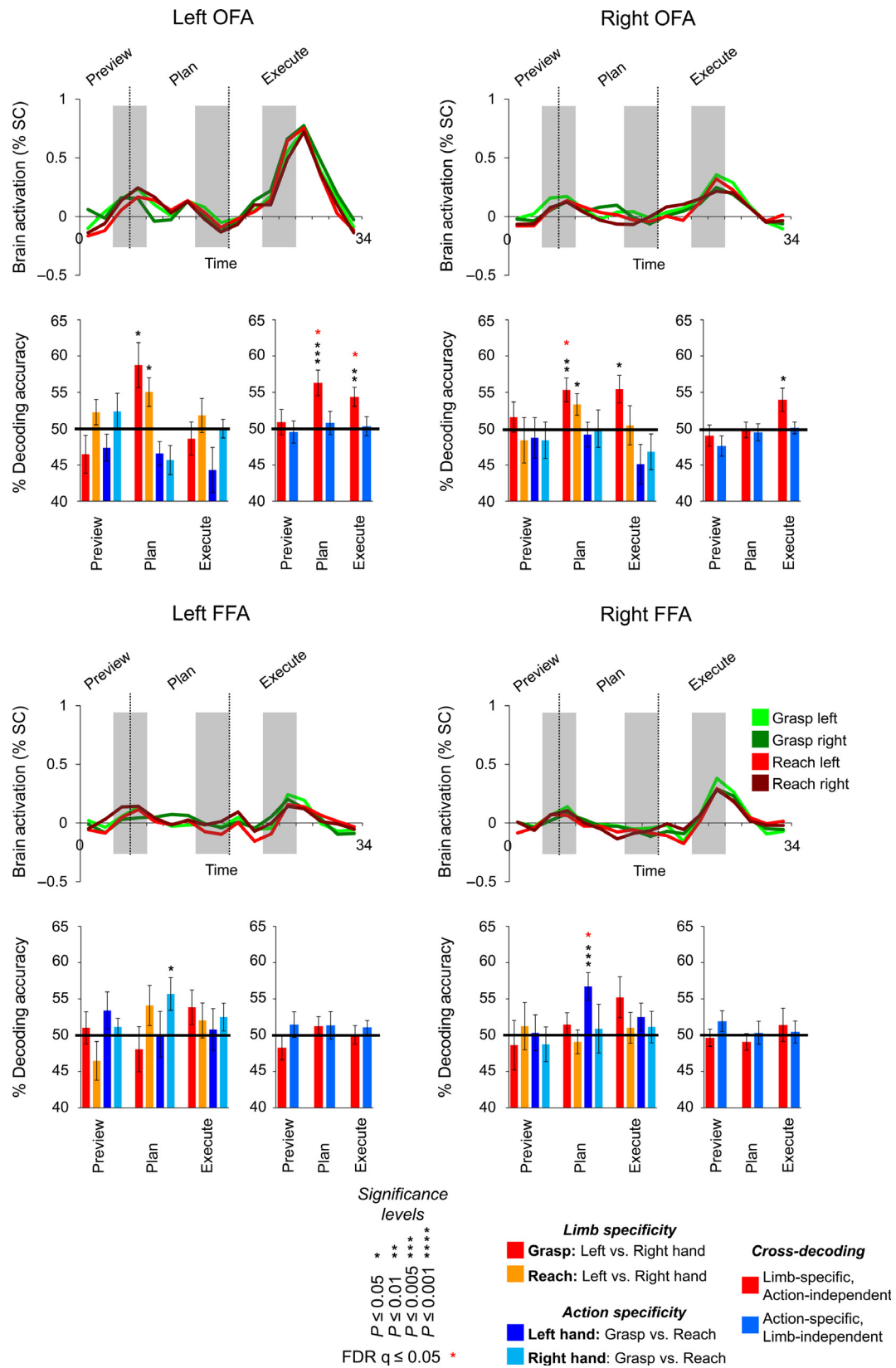


FIG. 4. Movement plan decoding from face-selective ROIs (OFA and FFA). Percentage signal change (% SC) time-courses and decoding accuracies are plotted and computed in the same way as in Fig. 3.

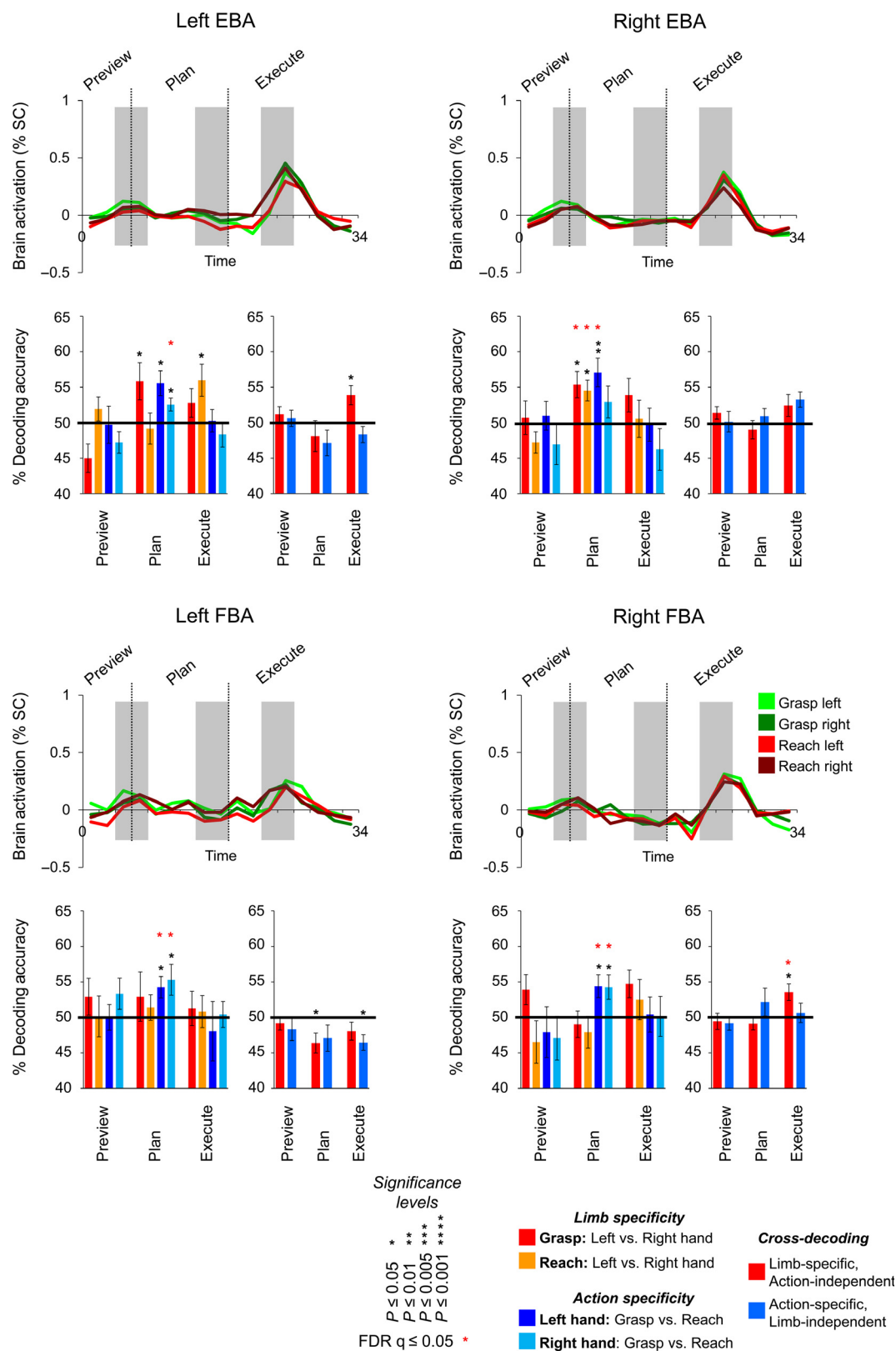


FIG. 5. Movement plan decoding from body-selective ROIs (EBA and FBA). Percentage signal change (% SC) time-courses and decoding accuracies are plotted and computed in the same way as in Fig. 3.

they were used to test the classifier in the other analysis when the left-hand trials were used for classifier training; see Results for more details). The means across participants of this cross-decoding proce-

dures are reported in Figs 3–6. Similar to above, we statistically assessed cross-decoding significance with a two-tailed *t*-test vs. 50% chance decoding, and an FDR correction of $q \leq 0.05$ was applied.

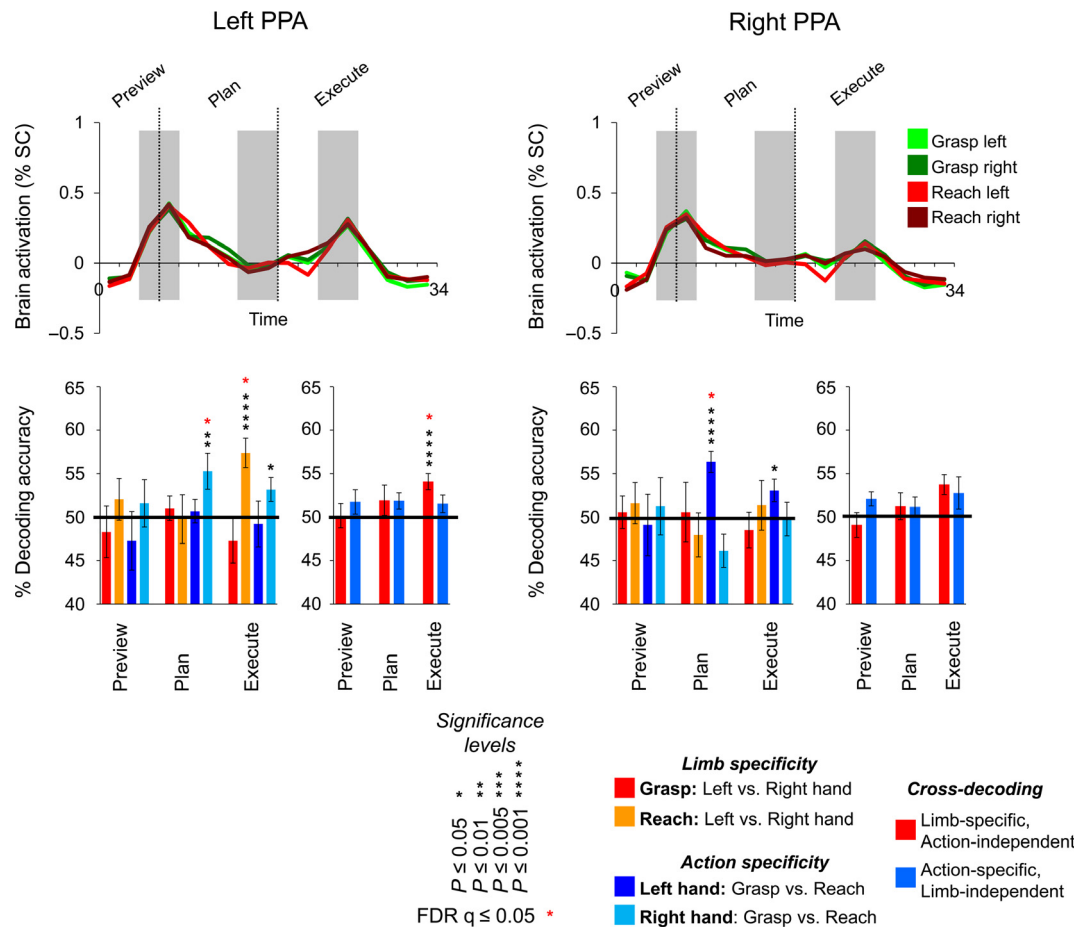


FIG. 6. Movement plan decoding from scene-selective ROIs (PPA). Percentage signal change (% SC) time-courses and decoding accuracies are plotted and computed in the same way as in Fig. 3.

Permutation tests

In addition to the *t*-test, we separately assessed statistical significance with non-parametric randomisation tests for the plan phase decoding accuracies. For specific details pertaining to this test, see our recent work (Gallivan *et al.*, 2011a,b, 2013). The maximum decoding accuracies obtained from the permuted distributions (each $n = 1000$) varied with the phase (i.e. preview, plan, or execute) and pair-wise comparison being considered, ranging from 51.82% (taken from the preview phase/ReachL vs. ReachR permuted distribution of 1000 accuracies) to 55.10% (taken from the execute phase/GraspR vs. ReachR permuted distribution of 1000 accuracies). The important finding highlighted by these tests is that the decoding accuracies (from correctly labeled trials) of pair-wise comparisons deemed statistically significant with the one-sample parametric *t*-tests (vs. 50% at $P < 0.05$) were found to be of even greater significance (at $P < 0.001$) with the empirical non-parametric permutation tests (that is, the correctly labeled trials generated accuracies greater than the maximum accuracy obtained from the corresponding permutation distribution).

Results

Across the OTC, we found a wide range of decoding profiles during movement preparation (these results are discussed in detail below, organised according to the category selectivity of each OTC area). The extent to which we were able to extract this planning-related

information from the spatial voxel patterns of activity was initially surprising, not only from a theoretical perspective (that is, the OTC is not typically associated with representations related to sensorimotor processes) but also from a methodological perspective; when we plotted the pre-movement signal amplitude responses from each ROI (see time-courses in Figs 3–6), we found that, for the most part, they were not only highly overlapping, but that they either: (i) did not substantially differ from baseline activity levels; or (ii) were actually below baseline activity levels [the latter was particularly the case in the more ventral temporal regions; see Figs 3–6; see also Harrison & Tong (2009) for findings of a similar nature]. Indeed, on the basis of the fMRI signal amplitude responses alone, a rather straightforward intuition would be that the OTC should selectively represent only two phases within our action-related trials; the first being visual events associated with presentation of the 3D object (i.e. corresponding to the peak of activity elicited by the preview phase), and the second being visual events associated with movement execution (i.e. corresponding to the peak of activity elicited by the execute phase; see the time-courses in Figs 3–6). Our pattern classification results below suggest that the information coded in the OTC is far richer than would be suggested by this simple intuition.

Preparatory decoding from object-selective ROIs (LO and pFs)

When we extracted the motor experiment planning-related activity from area LO and the pFs, we observed a notable dissociation.

Whereas the activity in area LO represented upcoming hand actions (grasping vs. reaching) with the contralateral limb only, the activity in the pFs represented upcoming hand actions with both the contralateral and ipsilateral limbs (Fig. 3). With the exception of the right LO, however, none of the object-selective areas discriminated the limb to be used in the movement, suggesting that the general level of processing in these object-selective regions was more related to how the object was to be interacted with (grasping vs. reaching) than to the proximal arm musculature required.

For each area (as well as those analysed below), we also examined whether the patterns of brain activity for one set of movements could be used to predict the patterns of brain activity for a different set of movements. To do this, we trained pattern classifiers by using one set of trials, and tested the accuracy of the trained classifiers by using a different set of trials (cross-decoding) [see Dinstein *et al.* (2008), Formisano *et al.* (2008), Harrison & Tong (2009), and Gallivan *et al.* (2011a), and further details in Materials and methods and the caption for Fig. 3]. Above chance cross-decoding would suggest that the representation of information in a ROI is, to some extent, shared across the separate conditions being tested. Here, we used this cross-decoding approach to characterise the limb-specific/action-independent and action-specific/limb-independent nature of the underlying representations in area LO and the pFs (at least to the extent that accurate cross-classification could be achieved). Limb-specific/action-independent accuracies were computed from training classifiers on GraspL vs. GraspR trials; the performance of the classifiers on ReachL vs. ReachR trials was then tested, and the resulting accuracies were averaged with those obtained from the opposite train-and-test ordering, within each subject. Action-specific/limb-independent accuracies were computed from training classifiers on GraspL vs. ReachL trials, and then tested on GraspR vs. ReachR trials (again, averaging these resulting accuracies with those obtained from the opposite train-and-test ordering, within each subject).

We found that the planning-related activity patterns in the right pFs showed significant cross-decoding for the hand actions (blue bar in cross-decoding plots in Fig. 3), suggesting that some portion of the signals in the area reflect more abstract, limb-invariant representations of the upcoming movement. Likewise, in both the left LO and the left pFs during the execution phase of the movement, we found significant cross-decoding for the limb used, suggesting a limb-specific/action-independent representation. We suspect, however, that much of this latter effect might be driven by the similar visual stimulation created when the left versus the right limb performs actions, regardless of the actual hand action being performed (i.e. grasping vs. reaching). For these and other reasons, our results focus on the decoding found during the plan phase.

Preparatory decoding from face-selective ROIs (OFA and FFA)

In the OFA and the FFA, we found evidence for dissociable representations during movement planning. The OFA coded the intended limb only, whereas the FFA coded intended hand actions for the contralateral limb only (Fig. 4). This action-related description of activity in the face-selective ROIs differs appreciably from that typically encountered in the literature. For instance, whereas several investigations have examined the specific attributes of face processing that either differ between, or are shared by, the OFA and the FFA (for example, see Atkinson & Adolphs, 2011 for a review), here we report their dissociation on the basis of neural responses prior to the generation of object-directed actions. Interestingly, in

the left OFA, we also found evidence for limb-specific/action-independent representations during planning (red bar in cross-decoding plots in Fig. 4).

Preparatory decoding from body-selective ROIs (EBA and FBA)

As with the face-selective ROIs, we found evidence for dissociable planning-related signals in the body-selective areas. The EBA primarily discriminated the intended limb and hand actions for the contralateral limb only, whereas the FBA represented hand actions for both the contralateral and ipsilateral limbs, but failed to discriminate the limb to be used (Fig. 5). In cases where action-related activity has been observed in the OTC, previous work has largely implicated body-selective areas, and the EBA in particular, as providing the primary locus (Astafiev *et al.*, 2004; Orlov *et al.*, 2010). This has led to the provocative suggestion that a shared topographic representation might exist between seen and moved body parts (although see Peelen & Downing, 2005b). The present findings provide further evidence in support of this general contention, although with the clarification that this shared topography may instead relate to seen and prepared movements of the body. In addition, we also report that movement representations related to intended hand actions can also be found more ventro-anteriorly in the FBA. It is worth acknowledging, however, that any shared topography for seen and prepared movements would be limited to body-selective areas within the OTC, given that a strict interpretation of this account would have difficulty in explaining why face-selective (and scene-selective) areas, noted above (and below), also represent minor components of the planned actions. Another interesting possibility, based on recent suggestions that the left middle temporal gyrus contains abstract representations of verb meanings (Bedny *et al.*, 2012), is that the EBA, located close to the middle temporal gyrus (e.g. Weiner & Grill-Spector, 2011), may, to some extent, represent the meanings associated with the auditory commands 'grasp' and 'reach' rather than sensorimotor processes related to those intended hand actions.

Preparatory decoding from scene-selective ROIs (PPA)

In the PPA, we found that planning-related representations were linked to hand actions of the contralateral limb (Fig. 6). Although the human medial temporal lobe is most prominently implicated in scene processing and spatial navigation (Epstein, 2008), neural recordings from human patients performing cued hand movements indicate that it may also serve as an important junction in the process of planning and executing motor actions (Tankus & Fried, 2012). In line with this suggestion, here we provide evidence that preparatory signals in the PPA may also code grasp vs. reach hand actions. Given the evidence suggesting that the FFA and the PPA contain inherent retinotopic biases for the center and peripheral visual fields, respectively (Levy *et al.*, 2001), one possibility is that the pre-movement activity patterns in the FFA and the PPA reflect an anticipation of the visual consequences of performing upcoming grasp vs. reach actions in the central and peripheral contralateral visual fields (i.e. with respect to fixation).

Effector-to-action-dependent preparatory gradient in the OTC

To quantify the extent to which the coding of action plans differed across the OTC, we compared decoding accuracies for each pairwise comparison (e.g. GraspL vs. GraspR) across the regions. We reasoned that, if certain category-selective brain areas coded specific

features of a movement (e.g. intended effector or hand action performed) more strongly than others, then decoding accuracies might be significantly higher for those features in those regions. A 2 (hemisphere) \times 7 (number of different ROIs) \times 4 (number of different pairwise comparisons performed per ROI) omnibus repeated measures ANOVA of the plan phase decoding accuracies revealed significant interactions between hemisphere and pairwise comparison ($F_{2,306} = 3.498$, $P = 0.046$) and between ROI and pairwise comparison ($F_{4,567} = 3.144$, $P = 0.021$). However, the hemisphere \times ROI interaction ($F_{3,579} = 1.105$, $P = 0.370$) and the three-way interaction ($F_{5,117} = 1.863$, $P = 0.121$) were non-significant (all statistics are Greenhouse–Geisser corrected). The significant hemisphere \times pairwise comparison interaction indicates differences in decoding between the two hemispheres, which is not surprising, given the strong contralaterality of the effects reported across areas (that is, the LO, EBA, FFA and PPA predominantly represent movements of the contralateral limb; Fig. 2B). The significant ROI \times pairwise comparison interaction indicates general differences in the pattern of decoding across the ROIs, which, again, is not particularly unexpected, given the general gradient of effector-to-action-dependent movement representations observed along the posterior–anterior axis of the OTC (Fig. 2B). To further investigate the differences in decoding across regions, we performed a series of planned comparisons (using paired sample *t*-tests), directly examining decoding between the lateral occipital and ventro-anterior ROIs for each category-selective set of areas. That is, for the object-selective areas, area LO and the pFs, we investigated, in each cortical hemisphere, whether there were significant differences in decoding for each pairwise comparison (e.g. GraspL vs. GraspR), and so forth for each set of the face-selective and body-selective ROIs (note that this analysis did not include the scene-selective PPA, given that our contrasts did not consistently reveal a lateral place-selective region in individuals) (but see Hasson *et al.*, 2003). For the sake of completeness, we report both the significant effects ($P \leq 0.05$) and trends towards significance ($P \leq 0.1$) from paired *t*-tests for differences between regions in the accuracy of classifying the effector or the hand action. We found higher decoding accuracies for the effector during grasping trials (GraspL vs. GraspR) in the following comparisons: left OFA > left FFA (at $P = 0.002$); right OFA > right FFA (at $P = 0.003$); and right EBA > right FBA (at $P = 0.10$). Moreover, these differences in decoding accuracy for the effector were very similar during reaching trials (ReachL vs. ReachR): right LO > right pFs (at $P = 0.066$); right OFA > right FFA (at $P = 0.072$); and right EBA > right FBA (at $P = 0.023$). Next, we found higher decoding accuracies for hand actions with the left limb (GraspL vs. ReachL) in the following comparisons: left pFs > left LO (at $P = 0.052$); and right FFA > right OFA (at $P = 0.007$). For the face-selective areas, these differences in decoding accuracies for hand actions were similar for the right limb (GraspR vs. ReachR): left FFA > left OFA (at $P = 0.015$). Although not all tests reached significance, the trends showed notable consistency between similar contrasts (especially the contrasts of GraspL vs. GraspR and ReachL vs. ReachR) and between comparisons of lateral and ventral foci. Taken together, these direct contrasts support the qualitative impression (Fig. 2B) that, during planning, effector-specific signals can be generally found in the lateral occipital cortex, whereas hand action-specific signals can be reliably found in the ventro-anterior visual cortex.

In general, these findings caution against ascribing any single OTC region a specific modular function on the basis of the category selectivity of its visual–perceptual responses alone (see also Haxby *et al.*, 2001). For instance, whereas the FFA and the PPA are commonly dissociated on the basis of face vs. scene processing (and

often localised accordingly, as done here), we show here that they represent similar aspects of an upcoming movement (i.e. hand actions for the contralateral limb; Figs 4 and 6). Likewise, whereas the OFA and the FFA are commonly co-activated and grouped together as components of a general face-processing network (Haxby *et al.*, 2000), here we show that they represent rather distinct aspects of upcoming hand movements (Fig. 4).

Caveats to interpreting movement-related activity in the OTC

Although, in several OTC areas, we did in fact observe significant decoding during movement execution (i.e. execute phase), we believe that caution should be applied when interpreting these findings. First, when subjects initiate movement, in addition to the activity related to motor commands being generated (e.g. contraction of arm and hand muscles), accompanying visual and proprioceptive responses are evoked. When ventral-stream visual areas in particular are considered, the visual responses elicited by viewing the limb move within the visual field are expected to explain a significant component of the execution-related activity patterns. In this light, recall that a major objective of the current study was to dissociate the proprioceptive and preparatory components of movement generation, so as to clarify the nature of previously reported action-related activations in the OTC. Here, we achieved this dissociation by examining the preparatory activity patterns that evolve prior to movement onset, thus avoiding any proprioceptive, tactile and/or visual confounds associated with the movement itself. Second, from a methodological standpoint, it is a frequent concern that motion artefacts may be transiently introduced into the data when the limb perturbs the magnetic field of the scanner (e.g. Culham, 2004). At present, it remains unclear how movement-related artefacts may influence spatial patterns of activity (if at all), and, in the current study, we wished to avoid this potentially confounding factor. For these and other reasons, the focus of the current analyses (and related interpretations) was to examine the patterns of activity that form prior to movement onset.

Evidence for the decoding of action plans from OTC structures

Critical to the interpretations provided above is the notion that the activity patterns decoded prior to action onset actually reflect components of a planned movement. Several aspects of both the methods and our data strongly support this claim. First, recall that, throughout the full length of the motor-related experiment, the central position of the target object was never changed, and thus differences in its retinal position cannot account for differences in plan phase decoding. Second, consistent with a role in movement planning, in most of the areas examined, plan phase decoding revealed representations of the limb to be used in an upcoming movement (left vs. right) and/or the hand actions (grasping vs. reaching) to be performed with the contralateral limb. With respect to the former, decoding of the movement effector (in this case, the limb) to be employed in an upcoming action towards a single target location is frequently used in the neurophysiological literature to argue for the responses being tied to movement intention (i.e. motor plan) rather than being a general effect of attention to target position (e.g., Snyder *et al.*, 1997; Andersen & Buneo, 2002; Chang *et al.*, 2008; Andersen & Cui, 2009) (see below for a further discussion of this important issue). With respect to the latter, contralateral representations of this sort (i.e. the coding of movements with respect to the intended limb) are commonly associated with the activity of

sensorimotor circuits in the brain (e.g. Cisek *et al.*, 2003; Beurze *et al.*, 2007; Chang *et al.*, 2008; Stark & Zohary, 2008). [Note that because, in the majority of areas examined, we find contralateral preparatory responses (e.g. LO, OFA, FFA, EBA, and PPA), this may further suggest that a more abstract, higher level of processing occurs in the pFs and the FBA, which both show hand action-related decoding for both the contralateral and ipsilateral limbs]. Third, and perhaps most convincingly, in each individual ROI examined, we were unable to decode intended actions during the preview phase portion of the trial, prior to the subject being aware of which action to perform. This indicates that statistically significant decoding during the plan phase is unlikely to arise simply by chance (otherwise, we would have also expected it during the preview phase). (Note also that we have separately validated our findings with a non-parametric permutation test and, for more cautious interpretations, we have also applied an FDR correction).

Discussion

Understanding the flow of visual information in the human brain is fundamental to ultimately understanding how purposeful, goal-directed actions are planned and implemented by the body. Here, we used an object-directed delayed movement task and fMRI MVPA to show that specific hand actions can be predicted from category-selective areas in the OTC prior to the movements being initiated. These findings challenge the widely held view that signals related to visual processing for action are restricted to the dorsal visual stream, by showing that action intentions can be revealed in ventral-stream structures, which are largely believed to exclusively support object-related perception.

Possible sources of preparatory signals in the OTC

What structures might provide the origin for the action-specific and effector-specific signals found in the OTC? One obvious possibility is that they reflect ongoing pre-movement sensorimotor computations in the parietal cortex. For example, the AIP area, an important region involved in sensorimotor transformations for grasping (Taira *et al.*, 1990; Baumann *et al.*, 2009), along with other areas of the dorsal visual stream, provides a significant source of reciprocal connections to the ventral visual stream (Rushworth *et al.*, 2006; Borra *et al.*, 2008, 2010; Kravitz *et al.*, 2011). This connectivity may help explain how signals related to object identity come to interact with sensorimotor signals related to hand pre-shaping at both the level of the OTC (as reported here) and the level of the parietal cortex (e.g. Murata *et al.*, 2000; Shmuelof & Zohary, 2005), and why the OTC and the parietal cortex are often co-activated in grasping tasks (e.g. Cavina-Pratesi *et al.*, 2010). Another possible source of the planning-related signals in the OTC is the prefrontal cortex. Reciprocal connections between ventral visual areas and the prefrontal cortex are well documented in macaques (Webster *et al.*, 1994; Borra *et al.*, 2010; Gerbella *et al.*, 2010), and the prefrontal cortex probably provides biasing signals to the OTC in accordance with higher-level goals and intended actions (Miller & Cohen, 2001; Miller *et al.*, 2002). An alternative possibility is that the findings here reflect feed-forward signals from retinotopically organised early visual structures (e.g. V1 and V2). Whereas, from the standpoint of spatial attention-related processing, feedback interactions between frontoparietal areas and the early visual cortex are becoming increasingly recognised (Desimone & Duncan, 1995; Lauritzen *et al.*, 2009; Borra & Rockland, 2011; Greenberg *et al.*, 2012), it remains largely unexplored whether sensorimotor signals from the frontopari-

etal cortex (e.g. discriminating the limb to be used) are also fed back onto early visual areas. If this is so, it would provide an alternative feed-forward route for the emergence of action-related signals in the OTC. It remains to be seen, however – even when considered outside the temporal constraints imposed by fMRI – whether these different possible sources of planning-related signals can be dissociated within the OTC.

Role of attention in movement preparation

What role might object-related spatial or feature-based attention play in either accounting for or contributing to the discriminatory activity patterns observed in the OTC prior to movement onset? This question seems particularly pertinent, considering the traditional role ascribed to the OTC in visual-perceptual processing and object recognition (e.g. Goodale & Milner, 1992; Grill-Spector & Malach, 2004), and its well-documented modulation by visual attention (e.g. Murray & Wojciulik, 2004; Murray & He, 2006). Multiple lines of behavioral and neural evidence indicate that attention is often directed to the location of planned movements (e.g. Deubel & Schneider, 1996; Moore & Fallah, 2001; Bekkering & Neggers, 2002; Bisley & Goldberg, 2003; Jovancevic *et al.*, 2006; Baldauf & Deubel, 2009). The tight linkage between attention and action planning has even led to suggestions that both processes are subserved by common or shared neural mechanisms (Rizzolatti *et al.*, 1987; for reviews, see Moore *et al.*, 2003; Awh *et al.*, 2006; Baldauf & Deubel, 2010), a notion that has received some empirical support (e.g. Moore & Armstrong, 2003). Several attempts have been made in studies of monkey neurophysiology to further test this notion by examining whether the neural mechanisms of action planning (i.e. intention) and attention can be dissociated. Although any detailed discussion of this important work is beyond the scope of this article, it still remains unclear to what extent preparatory neural responses in visuomotor structures, such as the parietal cortex, reflect attention-related coding for stimulus location and/or its behavioral relevance (Gottlieb *et al.*, 1998; Colby & Goldberg, 1999; Gottlieb & Goldberg, 1999; Bisley & Goldberg, 2003) or intention-related coding for the direction of the movement and/or movement effector (Snyder *et al.*, 1997; Zhang & Barash, 2000, 2004). In the case of the latter, it has been well argued that, if the neuronal responses that form prior to movement were only related to spatial attention or stimulus-related processing, then they should fail to discriminate the effector to be used in an upcoming action (eye vs. limb, left limb vs. right limb, etc.) – a notion considerably at odds with the observation of effector-specific response properties across multiple investigations (for reviews, see Andersen & Buneo, 2002; Andersen & Cui, 2009). Despite the diversity of observations at the single-neuron level, it is generally agreed that the planning of goal-directed behavior should be supported by the multiplexing of attention-related and motor-specific signals at the level of neural populations (Andersen & Buneo, 2002; Awh *et al.*, 2006; Andersen & Cui, 2009; Bisley & Goldberg, 2010), an idea consistent with recent reports showing the prevalence of both signals at the level of single parietal (Liu *et al.*, 2010; Premereur *et al.*, 2011) and frontal (Gregoriou *et al.*, 2012) brain areas.

Regardless of whether or not the current OTC results arise from intention or attention – and, indeed, if one thinks of attention as enhanced object-related processing for upcoming action, then any dichotomy seems artificial – the point to emphasise here is that the intention to act (and perhaps attention to action-relevant object attributes) enhances coding within the ventral stream. According to the traditional view of the two visual streams (Goodale & Milner,

1992), the processing of object information for action occurs in the dorsal stream but not the ventral stream. Indeed, earlier univariate fMRI comparisons from our laboratory (Culham *et al.*, 2003; Cavina-Pratesi *et al.*, 2007) showed no difference in fMRI activation levels in area LO for grasping vs. reaching, as would be predicted if subjects were paying more attention to certain object features (e.g. orientation) during grasping or, alternatively, visually imagining one type of action vs. the other. In contrast, our results here suggest that, when multivariate patterns are taken into account, the intention to act, and perhaps corresponding changes in the spatial processing of the relevant object property, do indeed affect coding in the ventral stream. Moreover, action intentions affect not just the processing in area LO, which is the area most strongly associated with object shape processing, but processing throughout the ventral stream [including areas such as the FFA and the PPA, which, by strong adherence to notions of category-selective modularity in OTC (Kanwisher, 2010), would not necessarily be expected], consistent with other distributed processing accounts (Haxby *et al.*, 2001). Notably, recent evidence suggests that areas within the ventral occipitotemporal (VOT) cortex, such as the PPA, process both object material and texture information (Cant & Goodale, 2011) – a set of object properties that is highly relevant to the planning of hand actions. Thus, areas within the VOT cortex may also discriminate differences in how the object is to be interacted with by the hand. That is, the ability to discriminate spatial activity patterns for grasping vs. reaching may result not just from enhanced processing of object form (e.g. features such as shape and size, which are thought to be processed within the lateral regions of the ventral stream), but also from object material and/or texture properties relevant for selecting appropriate grip force (such as object compliance, fragility, and surface friction). Furthermore, although subjects did not actually lift the object in grasp trials in this experiment, other object properties (such as density/weight) are relevant for programming lifting forces (see Flanagan *et al.*, 2009; Johansson & Flanagan, 2009), and may nevertheless be programmed (automatically) during grasping. Note that these object properties are far less relevant in reach trials, in which the object is simply touched. Together, these factors may account for why the PPA (and perhaps also why the FFA, also located in the VOT cortex) appear to discriminate hand actions of the contralateral limb – a type of representation that is often associated with the organisation of cortical sensorimotor networks (e.g. Cisek *et al.*, 2003).

In summary, we believe that a visual attention-related account of the current findings does not undermine our interpretation of the discriminatory preparatory signals in the OTC as being linked to the coding of plans for action. Indeed, if attending to different features of the object fully accounts for the pattern of decoding found here in the OTC, then, we would argue, it is only because the planned actions differ.

What is being represented in the OTC?

Given the traditional visual-perceptual function ascribed to category-selective areas in the OTC (e.g., Grill-Spector & Malach, 2004), the fact that we can decode features of upcoming hand actions from several of these areas has multiple implications regarding the nature of processing in these regions. The first is a clarification concerning previous reports of fMRI activity in the OTC during seen and unseen hand/arm actions. Prior to the current findings, it was suggested that motor representations in the OTC, in line with a purely perceptual/sensory role, might simply reflect the proprioceptive, tactile and/or visual feedback responses that accompany move-

ment (e.g. Astafiev *et al.*, 2004; Cavina-Pratesi *et al.*, 2010; Orlov *et al.*, 2010). Critically, the fact that we can decode upcoming hand movements (i.e. prior to them being performed) argues against this purely sensory feedback interpretation. One possibility is that the action-related preparatory responses in the OTC actually indicate some underlying role in movement planning, perhaps by using motor efference copies to anticipate the sensory consequences of moving certain body parts (e.g. Haarmeier *et al.*, 1997; Keyser & Perrett, 2004). Given the delay of incoming visual and proprioceptive signals (see Andersen & Cui, 2009; Johansson & Flanagan, 2009 for reviews), these types of forward-state estimations are thought to be critical in motor control (Wolpert & Flanagan, 2001), as well as in distinguishing movements of the body from movements of the world (von Helmholtz, 1866; Haarmeier *et al.*, 2001; Shergill *et al.*, 2003). Another possibility, which is more consistent with the traditional role of the OTC in visual processing, is that the preparatory responses in the OTC instead reflect a perceptual representation of the target object that is dynamically shaped and updated in the context of upcoming actions. Given that the target object remained in a constant location, this cannot be attributable to coarse spatial attention; nevertheless, this alternative explanation could correspond to attention to specific object features or spatial locations relevant to action planning (such as those where the digits will be placed during grasping). As noted above, the current data do not allow us to disentangle these two possibilities.

Other related possibilities, although not mutually exclusive, are that OTC activity during movement planning may be linked to the recall of particular movements from memory (Gelbard-Sagiv *et al.*, 2008; Bar, 2009), associations between auditory instructions and executed actions (Wise & Murray, 2000), or the kinematic properties of executed movements (Tankus & Fried, 2012). It is possible that our findings also resonate with the broader concept of a 'salience' or 'priority' map (Gottlieb *et al.*, 1998; Bisley & Goldberg, 2003; Fecteau & Munoz, 2006 for review), which is often evoked to describe the activity within frontoparietal circuits and brainstem structures, whereby items of interest in the environment are represented by activity in accordance with their behavioral relevance. In this framework, it is possible that attributes that are more relevant for grasping than for reaching (e.g. object shape, object mass distribution, contact points, and hand precision required) or for one limb vs. the other (e.g. selective processing of the object surface closest to the hand to be used) are being represented in several OTC areas prior to movement. Indeed, monkey neurophysiology has shown neural selectivity for object size, shape, surface and orientation in the AIP area (Murata *et al.*, 2000; Srivastava *et al.*, 2009; Theys *et al.*, 2012), and it seems plausible that structures in the ventral visual cortex interconnected with the AIP area (Borra *et al.*, 2008) may also represent similar motor-relevant object-related information. We cannot, however, rule out an alternative possibility that some component of the observed decoding in the OTC prior to movement could be attributable to the processes of motor imagery (see Downing & Peelen, 2011), which is known to recruit similar circuits to those engaged in actual movement execution (Filimon *et al.*, 2007). However, it is worth noting that the discriminatory responses within these regions largely occur during movement preparation rather than during execution, and if the current results were completely attributable to imagery – or other covert processes (visual, motor, or otherwise) – then it follows that we might expect similar (and probably even greater) levels of decoding in the same regions during the actual movement itself (when the visual and/or motor stimulation associated with the movement actually arises). In addition, it is worth noting that the responses in the OTC prior to movement

remain either action-specific (i.e. differentiating upcoming grasp from reach actions), effector-specific (i.e. differentiating upcoming actions of the left and right limb), or some conjunction of both (e.g. responses consistent with a contralateral representation of the movement), making a pure imagery-as-epiphenomenon account an unlikely explanation for the full pattern of results [see also Orlov *et al.* (2010) for findings arguing against a pure imagery account of execution-related fMRI responses in the OTC]. It remains to be seen, however, whether similar OTC results to those of the current study would be obtained if participants simply anticipated observing videos of hand actions (e.g. grasping and reaching with the left or right hand), without the instruction to perform the actions themselves. If this were the case, it might suggest that the action-related representations in the OTC are primarily linked to the sensory consequences of movement (whether from oneself or those of another individual) rather than movement preparation as such.

One point concerning the role of the OTC in sensorimotor processing may require clarification: If the OTC, at some level, is involved in movement planning, then why do transcranial magnetic stimulation (e.g. Cohen *et al.*, 2009) or lesions (e.g. Goodale *et al.*, 1994) to the OTC not impair immediate hand actions (as does disruption to frontoparietal structures)? First, such disruptions may, in fact, lead to impairments that are more subtle than frontoparietal impairments. For example, neuropsychological work has emphasised the preservation of functions (e.g. preserved grasping despite ventral-stream damage in famous patient DF), but this does not necessarily mean that patients perform as well as a group of healthy controls (Himmelbach *et al.*, 2012). Second, according to most accounts, ventral visual-stream structures play only a supportive role in action selection (e.g. Goodale & Milner, 1992; Passingham & Toni, 2001). This has raised the ecological notion that the OTC may not have originally evolved for the purpose of pure perceptual processing as such, but rather for an earlier, more basic role in collecting information (e.g. object shape) from the environment for the purposes of movement preparation (Cisek, 2007; Cisek & Kalaska, 2010). From an evolutionary standpoint, this corresponds well with the intimate link in ethology between stimulus recognition (such as 'sign stimuli') and the performance of specific behaviors (Tinbergen, 1950; Ewert, 1997). One intriguing possibility is that the current findings reveal aspects of this more ancient behaviorally oriented architecture in the OTC.

Probing an action-oriented organisation of the OTC

Given our characterisation of OTC areas according to the features of upcoming movements that they can predict, one obvious question is the extent to which this pattern of results compares with the well-documented organisation of action-related signals across the frontoparietal cortex. Cumulative evidence from neurophysiological studies in humans and monkeys suggests that both the parietal and the frontal cortex contain coarse and continuous activation gradients, in which stronger representations of the effector (e.g. limb) tend to correspond with weaker representations of the action goal (e.g. target location), and *vice versa*. For instance, moving rostrally to caudally in the frontal cortex, an orderly transition from space-related to effector-related signals can be seen, and, likewise, the same spatial-to-effector gradient appears to be mirrored (posteriorly to anteriorly) in the parietal cortex (e.g. Cisek *et al.*, 2003; Beurze *et al.*, 2007). Intriguingly, here we show that, whereas only hand action-related signals can be decoded from ventral temporal regions (pFs, FFA, FBA, and PPA), some mixture of both effector-related and hand

action-related signals can be decoded from lateral occipital regions (LO, OFA, and EBA; Fig. 2B). Compellingly, this gradient of OTC activity loosely corresponds to the general goal-related to effector-related gradient of activity seen in the frontal cortex, as well as the mirrored gradient of activity found in the parietal cortex (see Fillion, 2010 for a review). Just as a considerable amount of work has focused on examining the category selectivity of visual-perceptual responses throughout the OTC (e.g. for faces, scenes, and bodies) (for example, see Levy *et al.*, 2001; Kanwisher, 2010; Mahon & Caramazza, 2011), considerable work will be required to shed light on how action-related responses are represented throughout the OTC.

Acknowledgements

This work was supported by an operating grant from the Canadian Institutes of Health Research (MOP84293) awarded to J. C. Culham, and by a Banting Postdoctoral fellowship, an Ontario Ministry of Research and Innovation Postdoctoral fellowship and Natural Sciences and Engineering Research Council Collaborative Research and Training Excellence funding awarded to J. P. Gallivan. We thank Derek Quinlan for his assistance in collecting the eye movement data.

Abbreviations

3D, three-dimensional; ACPC, anterior and posterior commissure; AIP, anterior intraparietal; BOT, bodies, objects and tools; EBA, extrastriate body area; FBA, fusiform body area; FDR, false discovery rate; FFA, fusiform face area; fMRI, functional magnetic resonance imaging; FPO, faces, places and objects; GraspL, 'grasp left' trial; GraspR, 'grasp right' trial; ITI, intertrial interval; LED, light-emitting diode; LO, lateral occipital; MRI, magnetic resonance imaging; MVPA, multi-voxel pattern analysis; OFA, occipital face area; OTC, occipitotemporal cortex; pFs, posterior fusiform sulcus; PPA, parahippocampal place area; ReachL, 'reach left' trial; ReachR, 'reach right' trial; ROI, region of interest; SVM, support vector machine; VOT, ventral occipitotemporal.

References

- Andersen, R.A. & Buneo, C.A. (2002) Intentional maps in posterior parietal cortex. *Annu. Rev. Neurosci.*, **25**, 189–220.
- Andersen, R.A. & Cui, H. (2009) Intention, action planning, and decision making in parietal-frontal circuits. *Neuron*, **63**, 568–583.
- Astafiev, S.V., Stanley, C.M., Shulman, G.L. & Corbetta, M. (2004) Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nat. Neurosci.*, **7**, 542–548.
- Atkinson, A.P. & Adolphs, R. (2011) The neuropsychology of face perception: beyond simple dissociations and functional selectivity. *Philos. T. Roy. Soc. B.*, **366**, 1726–1738.
- Awh, E., Armstrong, K.M. & Moore, T. (2006) Visual and oculomotor selection: links, causes and implications for spatial attention. *Trends Cogn. Sci.*, **10**, 124–130.
- Baldauf, D. & Deubel, H. (2009) Attentional selection of multiple goal positions before rapid hand movement sequences: an event-related potential study. *J. Cognitive Neurosci.*, **21**, 18–29.
- Baldauf, D. & Deubel, H. (2010) Attentional landscapes in reaching and grasping. *Vision Res.*, **50**, 999–1013.
- Bar, M. (2009) The proactive brain: memory for predictions. *Philos. T. Roy. Soc. B.*, **364**, 1235–1243.
- Baumann, M.A., Fluet, M.C. & Scherberger, H. (2009) Context-specific grasp movement representation in the macaque anterior intraparietal area. *J. Neurosci.*, **29**, 6436–6448.
- Bedny, M., Caramazza, A., Pascual-Leone, A. & Saxe, R. (2012) Typical neural representations of action verbs develop without vision. *Cereb. Cortex*, **22**, 286–293.
- Bekkering, H. & Neggers, S.F. (2002) Visual search is modulated by action intentions. *Psychol. Sci.*, **13**, 370–374.
- Benjamini, Y. & Yekutieli, D. (2001) The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.*, **29**, 1165–1188.
- Beurze, S.M., de Lange, F.P., Toni, I. & Medendorp, W.P. (2007) Integration of target and effector information in the human brain during reach planning. *J. Neurophysiol.*, **97**, 188–199.

- Bisley, J.W. & Goldberg, M.E. (2003) Neuronal activity in the lateral intraparietal area and spatial attention. *Science*, **299**, 81–86.
- Bisley, J.W. & Goldberg, M.E. (2010) Attention, intention, and priority in the parietal lobe. *Annu. Rev. Neurosci.*, **33**, 1–21.
- Borra, E. & Rockland, K.S. (2011) Projections to early visual areas v1 and v2 in the calcarine fissure from parietal association areas in the macaque. *Front. Neuroanat.*, **5**, 35.
- Borra, E., Belmalih, A., Calzavara, R., Gerbella, M., Murata, A., Rozzi, S. & Luppino, G. (2008) Cortical connections of the macaque anterior intraparietal (AIP) area. *Cereb. Cortex*, **18**, 1094–1111.
- Borra, E., Ichinohe, N., Sato, T., Tanifuji, M. & Rockland, K.S. (2010) Cortical connections to area TE in monkey: hybrid modular and distributed organization. *Cereb. Cortex*, **20**, 257–270.
- Boynton, G.M., Engel, S.A., Glover, G.H. & Heeger, D.J. (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J. Neurosci.*, **16**, 4207–4221.
- Calder, A.J. & Young, A.W. (2005) Understanding the recognition of facial identity and facial expression. *Nat. Rev. Neurosci.*, **6**, 641–651.
- Cant, J.S. & Goodale, M.A. (2007) Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cereb. Cortex*, **17**, 713–731.
- Cant, J.S. & Goodale, M.A. (2011) Scratching beneath the surface: new insights into the functional properties of the lateral occipital area and parahippocampal place area. *J. Neurosci.*, **31**, 8248–8258.
- Cavina-Pratesi, C., Goodale, M. & Culham, J.C. (2007) fMRI reveals a dissociation between grasping and perceiving the size of real 3D objects. *PLoS ONE*, **9**, 1–14.
- Cavina-Pratesi, C., Monaco, S., Fattori, P., Galletti, C., McAdam, T.D., Quinlan, D.J., Goodale, M.A. & Culham, J.C. (2010) Functional magnetic resonance imaging reveals the neural substrates of arm transport and grip formation in reach-to-grasp actions in humans. *J. Neurosci.*, **30**, 10306–10323.
- Chang, S.W., Dickinson, A.R. & Snyder, L.H. (2008) Limb-specific representation for reaching in the posterior parietal cortex. *J. Neurosci.*, **28**, 6128–6140.
- Cisek, P. (2007) Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos. T. Roy. Soc. B.*, **362**, 1585–1599.
- Cisek, P. & Kalaska, J.F. (2010) Neural mechanisms for interacting with a world full of action choices. *Annu. Rev. Neurosci.*, **33**, 269–298.
- Cisek, P., Crammond, D.J. & Kalaska, J.F. (2003) Neural activity in primary motor and dorsal premotor cortex in reaching tasks with the contralateral versus ipsilateral arm. *J. Neurophysiol.*, **89**, 922–942.
- Cohen, N.R., Cross, E.S., Tunik, E., Grafton, S.T. & Culham, J.C. (2009) Ventral and dorsal stream contributions to the online control of immediate and delayed grasping: a TMS approach. *Neuropsychologia*, **47**, 1553–1562.
- Colby, C.L. & Goldberg, M.E. (1999) Space and attention in parietal cortex. *Annu. Rev. Neurosci.*, **22**, 319–349.
- Culham, J.C. (2004) Human brain imaging reveals a parietal area specialized for grasping. In Kanwisher, N. & Duncan, J. (Eds), *Attention and Performance XX: Functional Brain Imaging of Human Cognition*. Oxford University Press, Oxford, pp. 417–438.
- Culham, J.C., Danckert, S.L., DeSouza, J.F., Gati, J.S., Menon, R.S. & Goodale, M.A. (2003) Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Exp. Brain Res.*, **153**, 180–189.
- David, N., Cohen, M.X., Newen, A., Bewernick, B.H., Shah, N.J., Fink, G.R. & Vokeley, K. (2007) The extrastriate cortex distinguishes between the consequences of one's own and others' behavior. *NeuroImage*, **36**, 1004–1014.
- Desimone, R. & Duncan, J. (1995) Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.*, **18**, 193–222.
- Deubel, H. & Schneider, W.X. (1996) Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vision Res.*, **36**, 1827–1837.
- Dinstein, I., Hasson, U., Rubin, N. & Heeger, D.J. (2007) Brain areas selective for both observed and executed movements. *J. Neurophysiol.*, **98**, 1415–1427.
- Dinstein, I., Gardner, J.L., Jazayeri, M. & Heeger, D.J. (2008) Executed and observed movements have different distributed representations in human aIPS. *J. Neurosci.*, **28**, 11231–11239.
- Downing, P.E. & Peelen, M.V. (2011) The role of occipitotemporal body-selective regions in person perception. *Cogn. Neurosci.*, **2**, 186–226.
- Downing, P.E., Jiang, Y., Shuman, M. & Kanwisher, N. (2001) A cortical area selective for visual processing of the human body. *Science*, **293**, 2470–2473.
- Duda, R.O., Hart, P.E. & Stork, D.G. (2001) *Pattern Classification*, 2nd Edn. Wiley, New York.
- Epstein, R. & Kanwisher, N. (1998) A cortical representation of the local visual environment. *Nature*, **392**, 598–601.
- Epstein, R.A. (2008) Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn. Sci.*, **12**, 388–396.
- Ewert, J.P. (1997) Neural correlates of key stimulus and releasing mechanism: a case study and two concepts. *Trends Neurosci.*, **20**, 332–339.
- Fecteau, J.H. & Munoz, D.P. (2006) Saliency, relevance, and firing: a priority map for target selection. *Trends Cogn. Sci.*, **10**, 382–390.
- Filimon, F. (2010) Human cortical control of hand movements: parietofrontal networks for reaching, grasping, and pointing. *Neuroscientist*, **16**, 388–407.
- Filimon, F., Nelson, J.D., Hagler, D.J. & Sereno, M.I. (2007) Human cortical representations for reaching: mirror neurons for execution, observation, and imagery. *NeuroImage*, **37**, 1315–1328.
- Filimon, F., Nelson, J.D., Huang, R.S. & Sereno, M.I. (2009) Multiple parietal reach regions in humans: cortical representations for visual and proprioceptive feedback during on-line reaching. *J. Neurosci.*, **29**, 2961–2971.
- Flanagan, J.R., Merritt, K. & Johansson, R.S. (2009) Predictive mechanisms and object representations used in object manipulation. In Hermsdörfer, J. & Nowak, D.A. (Eds), *Sensorimotor Control of Grasping: Physiology and Pathophysiology*. Cambridge University Press, Cambridge, pp. 161–177.
- Formisano, E., De Martino, F., Bonte, M. & Goebel, R. (2008) 'Who' is saying 'what'? Brain-based decoding of human voice and speech. *Science*, **322**, 970–973.
- Gallivan, J.P., McLean, D.A., Smith, F.W. & Culham, J.C. (2011a) Decoding effector-dependent and effector-independent movement intentions from human parieto-frontal brain activity. *J. Neurosci.*, **31**, 17149–17168.
- Gallivan, J.P., McLean, D.A., Valyear, K.F., Pettypiece, C.E. & Culham, J.C. (2011b) Decoding action intentions from preparatory brain activity in human parieto-frontal networks. *J. Neurosci.*, **31**, 9599–9610.
- Gallivan, J.P., McLean, D.A., Flanagan, J.R. & Culham, J.C. (2013) Where one hand meets the other: limb-specific and action-dependent movement plans decoded from preparatory signals in single human frontoparietal brain areas. *J. Neurosci.*, **33**, 1991–2008.
- Gauthier, I., Tarr, M.J., Moylan, J., Skudlarski, P., Gore, J.C. & Anderson, A.W. (2000) The fusiform 'face area' is part of a network that processes faces at the individual level. *J. Cognitive Neurosci.*, **12**, 495–504.
- Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R. & Fried, I. (2008) Internally generated reactivation of single neurons in human hippocampus during free recall. *Science*, **322**, 96–101.
- Gerbella, M., Belmalih, A., Borra, E., Rozzi, S. & Luppino, G. (2010) Cortical connections of the macaque caudal ventrolateral prefrontal areas 45A and 45B. *Cereb. Cortex*, **20**, 141–168.
- Goodale, M.A. & Milner, A.D. (1992) Separate visual pathways for perception and action. *Trends Neurosci.*, **15**, 20–25.
- Goodale, M.A., Jakobson, L.S. & Keillor, J.M. (1994) Differences in the visual control of pantomimed and natural grasping movements. *Neuropsychologia*, **32**, 1159–1178.
- Gottlieb, J. & Goldberg, M.E. (1999) Activity of neurons in the lateral intraparietal area of the monkey during an antisaccade task. *Nat. Neurosci.*, **2**, 906–912.
- Gottlieb, J.P., Kusunoki, M. & Goldberg, M.E. (1998) The representation of visual salience in monkey parietal cortex. *Nature*, **391**, 481–484.
- Greenberg, A.S., Verstynen, T., Chiu, Y.C., Yantis, S., Schneider, W. & Behrmann, M. (2012) Visuotopic cortical connectivity underlying attention revealed with white-matter tractography. *J. Neurosci.*, **32**, 2773–2782.
- Gregoriou, G.G., Gotts, S.J. & Desimone, R. (2012) Cell-type-specific synchronization of neural activity in FEF with V4 during attention. *Neuron*, **73**, 581–594.
- Grill-Spector, K. & Malach, R. (2004) The human visual cortex. *Annu. Rev. Neurosci.*, **27**, 649–677.
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y. & Malach, R. (1999) Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, **24**, 187–203.
- Grill-Spector, K., Kourtzi, Z. & Kanwisher, N. (2001) The lateral occipital complex and its role in object recognition. *Vision Res.*, **41**, 1409–1422.
- Haarmeier, T., Thier, P., Repnow, M. & Petersen, D. (1997) False perception of motion in a patient who cannot compensate for eye movements. *Nature*, **389**, 849–852.
- Haarmeier, T., Bunjes, F., Lindner, A., Berret, E. & Thier, P. (2001) Optimizing visual motion perception during eye movements. *Neuron*, **32**, 527–535.
- Harrison, S.A. & Tong, F. (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature*, **458**, 632–635.

- Hasson, U., Harel, M., Levy, I. & Malach, R. (2003) Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron*, **37**, 1027–1041.
- Haxby, J.V., Hoffman, E.A. & Gobbini, M.I. (2000) The distributed human neural system for face perception. *Trends Cogn. Sci.*, **4**, 223–233.
- Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L. & Pietrini, P. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, **293**, 2425–2430.
- von Helmholtz, H. (1866) *Handbook of Physiological Optics*, 3rd Edn. Dover Publications, New York.
- Himmelbach, M., Boehme, R. & Karnath, H.O. (2012) 20 years later: a second look on DF's motor behaviour. *Neuropsychologia*, **50**, 139–144.
- Johansson, R.S. & Flanagan, J.R. (2009) Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nat. Rev. Neurosci.*, **10**, 345–359.
- Jovancevic, J., Sullivan, B. & Hayhoe, M. (2006) Control of attention and gaze in complex environments. *J. Vision*, **6**, 1431–1450.
- Kanwisher, N. (2010) Functional specificity in the human brain: a window into the functional architecture of the mind. *Proc. Natl. Acad. Sci. USA*, **107**, 11163–11170.
- Kanwisher, N., McDermott, J. & Chun, M.M. (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.*, **17**, 4302–4311.
- Keysers, C. & Perrett, D.I. (2004) Demystifying social cognition: a Hebbian perspective. *Trends Cogn. Sci.*, **8**, 501–507.
- Kravitz, D.J., Saleem, K.S., Baker, C.I. & Mishkin, M. (2011) A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.*, **12**, 217–230.
- Lauritzen, T.Z., D'Esposito, M., Heeger, D.J. & Silver, M.A. (2009) Top-down flow of visual spatial attention signals from parietal to occipital cortex. *J. Vision*, **9**, 11–14.
- Levy, I., Hasson, U., Avidan, G., Hendler, T. & Malach, R. (2001) Center-periphery organization of human object areas. *Nat. Neurosci.*, **4**, 533–539.
- Liu, Y., Yttri, E.A. & Snyder, L.H. (2010) Intention and attention: different functional roles for LIPd and LIPv. *Nat. Neurosci.*, **13**, 495–500.
- Mahon, B.Z. & Caramazza, A. (2011) What drives the organization of object knowledge in the brain? *Trends Cogn. Sci.*, **15**, 97–103.
- Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R. & Tootell, R.B. (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci. USA*, **92**, 8135–8139.
- Miller, E.K. & Cohen, J.D. (2001) An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.*, **24**, 167–202.
- Miller, E.K., Freedman, D.J. & Wallis, J.D. (2002) The prefrontal cortex: categories, concepts and cognition. *Philos. T. Roy. Soc. B.*, **357**, 1123–1136.
- Misaki, M., Kim, Y., Bandettini, P.A. & Kriegeskorte, N. (2010) Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *NeuroImage*, **53**, 103–118.
- Moore, T. & Armstrong, K.M. (2003) Selective gating of visual signals by microstimulation of frontal cortex. *Nature*, **421**, 370–373.
- Moore, T. & Fallah, M. (2001) Control of eye movements and spatial attention. *Proc. Natl. Acad. Sci. USA*, **98**, 1273–1276.
- Moore, T., Armstrong, K.M. & Fallah, M. (2003) Visuomotor origins of covert spatial attention. *Neuron*, **40**, 671–683.
- Murata, A., Gallese, V., Luppino, G., Kaseda, M. & Sakata, H. (2000) Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J. Neurophysiol.*, **83**, 2580–2601.
- Murray, S.O. & He, S. (2006) Contrast invariance in the human lateral occipital complex depends on attention. *Curr. Biol.*, **16**, 606–611.
- Murray, S.O. & Wojciulik, E. (2004) Attention increases neural selectivity in the human lateral occipital complex. *Nat. Neurosci.*, **7**, 70–74.
- Orlov, T., Makin, T.R. & Zohary, E. (2010) Topographic representation of the human body in the occipitotemporal cortex. *Neuron*, **68**, 586–600.
- Passingham, R.E. & Toni, I. (2001) Contrasting the dorsal and ventral visual systems: guidance of movement versus decision making. *NeuroImage*, **14**, S125–S131.
- Peelen, M.V. & Downing, P.E. (2005a) Selectivity for the human body in the fusiform gyrus. *J. Neurophysiol.*, **93**, 603–608.
- Peelen, M.V. & Downing, P.E. (2005b) Is the extrastriate body area involved in motor actions? *Nat. Neurosci.*, **8**, 125; author reply 125–126.
- Peelen, M.V. & Downing, P.E. (2005c) Within-subject reproducibility of category-specific visual activation with functional MRI. *Hum. Brain Mapp.*, **25**, 402–408.
- Premereur, E., Vanduffel, W. & Janssen, P. (2011) Functional heterogeneity of macaque lateral intraparietal neurons. *J. Neurosci.*, **31**, 12307–12317.
- Puce, A., Allison, T., Asgari, M., Gore, J.C. & McCarthy, G. (1996) Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J. Neurosci.*, **16**, 5205–5215.
- Rizzolatti, G., Riggio, L., Dascola, I. & Umiltà, C. (1987) Reorienting attention across the horizontal and vertical meridians: evidence in favor of premotor theory of attention. *Neuropsychologia*, **25**, 31–40.
- Rushworth, M.F., Behrens, T.E. & Johansen-Berg, H. (2006) Connection patterns distinguish 3 regions of human parietal cortex. *Cereb. Cortex*, **16**, 1418–1430.
- Schwarzlose, R.F., Baker, C.I. & Kanwisher, N. (2005) Separate face and body selectivity on the fusiform gyrus. *J. Neurosci.*, **25**, 11055–11059.
- Shergill, S.S., Bays, P.M., Frith, C.D. & Wolpert, D.M. (2003) Two eyes for an eye: the neuroscience of force escalation. *Science*, **301**, 187.
- Shmuelof, L. & Zohary, E. (2005) Dissociation between ventral and dorsal fMRI activation during object and action recognition. *Neuron*, **47**, 457–470.
- Smith, F.W. & Muckli, L. (2010) Non-stimulated early visual areas carry information about surrounding context. *Proc. Natl. Acad. Sci. USA*, **107**, 20099–20103.
- Snyder, L.H., Batista, A.P. & Andersen, R.A. (1997) Coding of intention in the posterior parietal cortex. *Nature*, **386**, 167–170.
- Srivastava, S., Orban, G.A., De Maziere, P.A. & Janssen, P. (2009) A distinct representation of three-dimensional shape in macaque anterior intraparietal area: fast, metric, and coarse. *J. Neurosci.*, **29**, 10613–10626.
- Stark, A. & Zohary, E. (2008) Parietal mapping of visuomotor transformations during human tool grasping. *Cereb. Cortex*, **18**, 2358–2368.
- Taira, M., Mine, S., Georgopoulos, A.P., Murata, A. & Sakata, H. (1990) Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Exp. Brain Res.*, **83**, 29–36.
- Tankus, A. & Fried, I. (2012) Visuomotor coordination and motor representation by human temporal lobe neurons. *J. Cognitive Neurosci.*, **24**, 600–610.
- Theys, T., Srivastava, S., van Loon, J., Goffin, J. & Janssen, P. (2012) Selectivity for three-dimensional contours and surfaces in the anterior intraparietal area. *J. Neurophysiol.*, **107**, 995–1008.
- Tinbergen, N. (1950) The hierarchical organisation of nervous mechanisms underlying instinctive behavior. *Symp. Soc. Exp. Biol.*, **4**, 305–312.
- Valyear, K.F. & Culham, J.C. (2010) Observing learned object-specific functional grasps preferentially activates the ventral stream. *J. Cognitive Neurosci.*, **22**, 970–984.
- Webster, M.J., Bachevalier, J. & Ungerleider, L.G. (1994) Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cereb. Cortex*, **4**, 470–483.
- Weiner, K.S. & Grill-Spector, K. (2011) Not one extrastriate body area: using anatomical landmarks, hMT+, and visual field maps to parcellate limb-selective activations in human lateral occipitotemporal cortex. *NeuroImage*, **56**, 2183–2199.
- Wise, S.P. & Murray, E.A. (2000) Arbitrary associations between antecedents and actions. *Trends Neurosci.*, **23**, 271–276.
- Wolpert, D.M. & Flanagan, J.R. (2001) Motor prediction. *Curr. Biol.*, **11**, R729–R732.
- Wolpert, D.M., Diedrichsen, J. & Flanagan, J.R. (2011) Principles of sensorimotor learning. *Nat. Rev.*, **12**, 739–751.
- Zhang, M. & Barash, S. (2000) Neuronal switching of sensorimotor transformations for antisaccades. *Nature*, **408**, 971–975.
- Zhang, M. & Barash, S. (2004) Persistent LIP activity in memory antisaccades: working memory for a sensorimotor transformation. *J. Neurophysiol.*, **91**, 1424–1441.