Feature-selective Attention in Frontoparietal Cortex: Multivoxel Codes Adjust to Prioritize Task-relevant Information

Jade Jackson, Anina N. Rich, Mark A. Williams, and Alexandra Woolgar

Abstract

■ Human cognition is characterized by astounding flexibility, enabling us to select appropriate information according to the objectives of our current task. A circuit of frontal and parietal brain regions, often referred to as the frontoparietal attention network or multiple-demand (MD) regions, are believed to play a fundamental role in this flexibility. There is evidence that these regions dynamically adjust their responses to selectively process information that is currently relevant for behavior, as proposed by the "adaptive coding hypothesis" [Duncan, J. An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience, 2,* 820–829, 2001]. Could this provide a neural mechanism for feature-selective attention, the process by which we preferentially process one feature of a stimulus over another? We used multivariate pattern analysis of fMRI data during a perceptually challenging categorization task to investigate whether the representation of visual object features in the MD regions flexibly adjusts according to task relevance. Participants were trained to categorize visually similar novel objects along two orthogonal stimulus dimensions (length/orientation) and performed short alternating blocks in which only one of these dimensions was relevant. We found that multivoxel patterns of activation in the MD regions encoded the task-relevant distinctions more strongly than the task-irrelevant distinctions: The MD regions discriminated between stimuli of different lengths when length was relevant and between the same objects according to orientation when orientation was relevant. The data suggest a flexible neural system that adjusts its representation of visual objects to preferentially encode stimulus features that are currently relevant for behavior, providing a neural mechanism for feature-selective attention. ■

INTRODUCTION

We live in a complex dynamic environment where the behavioral relevance of different sensory input changes rapidly. To function successfully, we need a cognitive system that can select what is currently relevant, ignore distraction, and update its responses in accordance with events in the world. The selection of relevant information can be specific to different features of visual objects depending on the current goal. For example, if I am looking for my blue cup among other cups, color is the relevant dimension. When I find my cup and reach to pick it up, other features of the cup are now relevant (e.g., orientation). Following Chen, Hoffmann, Albright, and Thiele (2012), we refer to the process of attending to and making a decision about one feature of an object, while ignoring other features of that object, as feature-selective attention. We use this nomenclature to differentiate it from feature-based attention, in which a relevant feature is used to select what object or location to attend to (e.g., attend to the red object). In feature-based attention, attention is directed toward objects and/or locations matching a cued value (e.g., red), while objects of a different color are ignored. In feature-selective attention, attention is instead directed toward a particular stimulus dimension (e.g., color), in preference to other dimensions (e.g., shape), to make a judgment about the relevant feature of a stimulus (Chen et al., 2012).

The adaptive coding hypothesis (Duncan, 2001) offers a possible neural mechanism for feature-selective attention. It holds that, within higher cortical regions, the response properties of single neurons are highly adaptable such that, in any particular task context, many cells become tuned to code information that is currently relevant. Evidence of such "adaptive coding" comes primarily from single-unit work with non-human primates in which neurons in higher cortical regions have been shown to alter coding as needed for behavior (Stokes et al., 2013; Cromer, Roy, & Miller, 2010; Roy, Riesenhuber, Poggio, & Miller, 2010; Freedman & Assad, 2006; Freedman, 2001; Sakagami & Niki, 1994). For example, in a go/no-go discrimination task where the relevant feature of a cue changed between three task contexts (Sakagami & Niki, 1994), 72% of PFC neurons showed different responses during the cue period for each task condition. In another example, Roy et al. (2010) demonstrated that 24% of PFC neurons had a distinct firing rate in response to one category of visual stimuli over another. These neurons responded to the relevant

Macquarie University, Sydney, Australia

Journal of Cognitive Neuroscience 29:2, pp. 310–321
doi:10.1162/jocn_a_01039

category membership of the stimuli, rather than to their simple visual properties. Moreover, when the task changed so that these same stimuli were recategorized in orthogonal categories, these neurons changed their firing rate to reflect the new task (for a similar result, see also Cromer et al., 2010). Similar results have been found for the lateral intraparietal cortex, where neural firing rates reflect learned category boundaries and change to reflect orthogonal category boundaries on retraining (Freedman & Assad, 2006).

In the human brain, candidate regions for adaptive coding are a set of frontal and parietal brain regions known as multiple-demand (MD) regions (Duncan, 2001, 2010). The MD regions incorporate the anterior insula/frontal operculum (AI/FO), the inferior frontal sulcus (IFS), the dorsal ACC/pre-SMA, and the intraparietal sulcus (IPS). They are characterized by their response to a wide range of task demands (Dosenbach et al., 2006; Nyberg et al., 2003; Duncan & Owen, 2000), even at the level of single participants (Fedorenko, Duncan, & Kanwisher, 2013). These regions have also been referred to as the "task-positive network" (Fox et al., 2005) or "frontoparietal control system" (Vincent, Kahn, Snyder, Raichle, & Buckner, 2008) and have been described as "flexible hubs" that adjust their connectivity patterns along with task demands (Cole et al., 2013).

In addition to being active for a range of tasks, human imaging data suggest that the MD regions are capable of coding a range of task features. Evidence for this comes from a multivariate pattern analysis (MVPA) of fMRI data, in which information coding is inferred if patterns of activation across voxels reliably discriminate between task events (e.g., Haynes & Rees, 2005; Kamitani & Tong, 2005; Haxby et al., 2001). Such studies suggest that the MD regions can code several different types of task-relevant information such as rules, stimuli, and motor responses (Woolgar, Williams, & Rich, 2015; Nee & Brown, 2012; Reverberi, Gorgen, & Haynes, 2011; Woolgar, Hampshire, Thompson, & Duncan, 2011; Woolgar, Thompson, Bor, & Duncan, 2011; Bode & Haynes, 2009; Haynes et al., 2007; Li, Ostwald, Giese, & Kourtzi, 2007). Moreover, in response to changes in task demands, the MD regions adjust their representation of perceptual (Woolgar, Williams, et al., 2015; Woolgar, Hampshire, et al., 2011) and rule information (Woolgar, Afshar, Williams, & Rich, 2015). For example, in Woolgar, Hampshire, et al.'s (2011) study, participants had to identify the spatial position of a visual stimulus. When the positions were close and overlapping, such that they were more difficult to discriminate, they were more strongly represented in MD regions, compared with when they were spaced far apart and perceptually easier, despite a weaker representation of the difficult stimuli in the visual cortex. However, in this work, the stimuli were always discriminated according to their spatial position, and this stimulus feature was always task relevant. Here, we examine the complementary question of whether flexibility of the MD system also underpins our capacity to attend to different features of the same object, depending on what is currently relevant. Suggestive evidence comes from a recent study in which object representations in lateral PFC (lPFC) could be decoded more strongly within a single task than between tasks (Harel, Kravitz, & Baker, 2014), raising the possibility that the same objects may be represented differently as task contexts change.

In this study, we examined the responses of the MD regions when different features of the same visual objects were made relevant. We presented a set of novel objects that varied along two dimensions (length of one of the spikes and orientation of that same spike; Figure 1). In separate blocks of trials, participants categorized the stimuli on the basis of one of the feature dimensions (length or orientation). Thus, at any one time, participants were required to discriminate objects according to one (relevant) dimension and ignore the other (irrelevant) dimension. We used MVPA of fMRI data to test whether the patterns of activation in the MD regions discriminate objects according to the externally imposed decision boundary and whether this categorization changes when an orthogonal decision boundary is used. If the representations in the MD regions are driven by physical stimulus characteristics, the same information should be present irrespective of the task. However, the adaptive coding hypothesis predicts that neural populations in the MD regions adjust their responses to adaptively code the currently relevant information; in which case, we should see stronger coding for the task-relevant feature distinction than the task-irrelevant distinction. We also examined responses in ROIs in the lateral occipital complex (LOC) and early visual cortex, for comparison. We found that MD coding of the relevant feature distinction was significantly stronger than discrimination along the irrelevant dimension. LOC followed this pattern, albeit more weakly, and the early visual cortex did not hold information that distinguished the physically identical stimuli depending on task context. Our results suggest that the frontoparietal MD network adjusts its representation of individual objects to make the specific discrimination that is needed for the current task. We suggest that this process supports selective attention to task-relevant object features.

METHODS

Participants

Twenty-six healthy adult volunteers (17 women; mean age = 24.3 years, *SD* = 5.27 years) participated. All participants were right-handed with normal or corrected-to-normal vision and no history of neurological or psychiatric disorder. Participants gave written informed consent, and the study was approved by the human research ethics committee of Macquarie University, Sydney, Australia. The participants received $50. There were 27 participants initially, but one (male) participant was excluded because he did not complete the task.

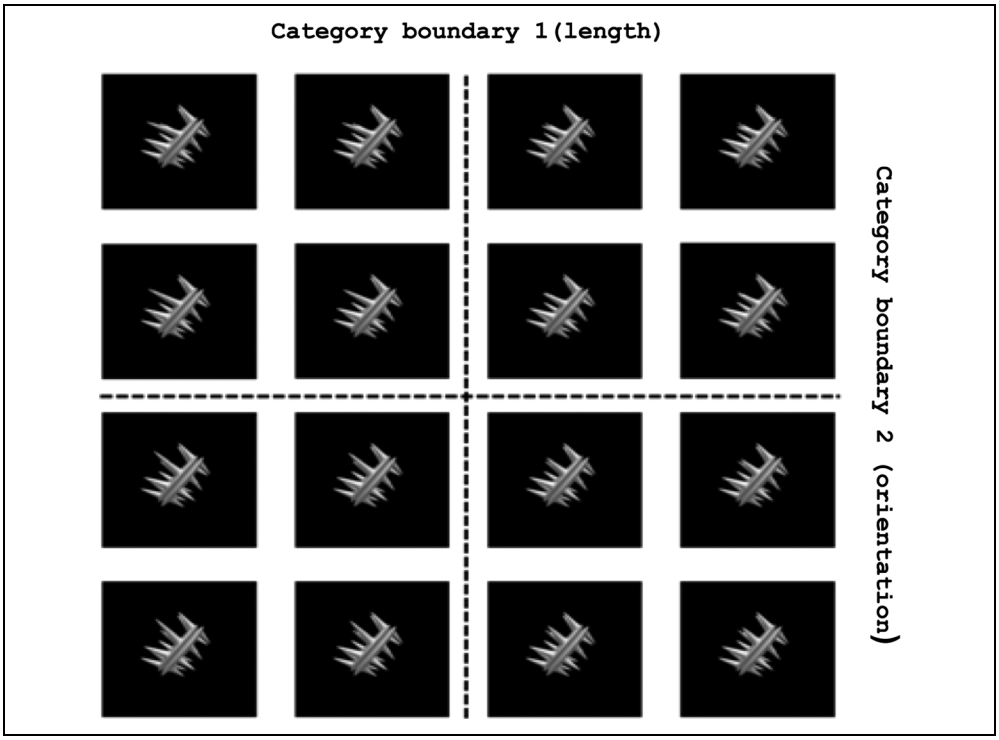

Figure 1. The stimulus set consisted of 16 spiky objects with eight objects on either side of two orthogonal category boundaries. Category boundaries were defined based on the length (horizontal axis) or orientation (vertical axis) of the third spike from the bottom on the left side of the object. Spike length and orientation were titrated on an individual-participant basis to equate the difficulty of categorization across the two dimensions.

Stimuli

Stimuli were abstract novel "spiky" objects created using custom MATLAB scripts (Op de Beeck, Baker, DiCarlo, & Kanwisher, 2006). The stimulus set consisted of 16 objects (Figure 1) in which one spike varied along two dimensions (its length and orientation). Participants learnt to discriminate between the 16 objects across two orthogonal decision boundaries (task contexts) based on the length and orientation dimensions. The relevant visual feature of the stimuli therefore varied depending on the current decision boundary. The stimuli were aligned so that the main stem of the objects appeared at an angle of +37° from vertical. During both the training and scanning sessions, the visual angle (VA) of the spiky object's length along its main axis was 8.07°. Stimulus presentation was controlled by a PC running the Psychophysics Toolbox 3 package (Brainard, 1997) in MATLAB (The MathWorks, Inc, Natick, MA). Stimuli were presented at central fixation on a screen and viewed through a mirror mounted on the head coil in the scanner.

Titrating Task Difficulty

Less than 1 week (1–7 days) before scanning, participants completed a behavioral testing session in which we titrated the discriminability of the stimuli to ensure that the length and orientation tasks were of comparable difficulty on an individual participant level. We started with a difference of 1.26° VA between the shortest and longest lengths of the spike and a maximum difference of 27.02° in the angle of orientation. After 192 trials, the range of lengths or orientations in the stimulus set was adjusted if there was a significant difference in the participant's RTs between the two tasks ($p < .05$). For this, we increased the difficulty of the task that had the lower average RT. This procedure was repeated until there was no difference in RT between the two task contexts, as assessed with Bayes factor (BF) analysis (Love et al., 2015) in each participant separately (BF < 1 taken as evidence for no difference between conditions; Dienes 2011). The difficulty of the orientation context was increased for 11 participants (the maximum angle decreased to 21.92° for six participants and 13.86° for five participants), and the difficulty of the length context was increased for five participants (the maximum difference in VA was decreased to 1.06° for two participants and 0.95° for three participants).

Procedure

Before titrating the stimuli, participants completed six blocks of practice trials to learn the task. Stimuli were initially presented for 400 msec until participants achieved >80% correct, after which objects were presented for 216 msec. Feedback (correct/incorrect) was presented after each response until participants achieved >80% performance, after which feedback (percent correct) was only given at the end of each block. Once participants reached a high performance level (>80% correct) in both task contexts, we then titrated the stimuli to ensure equal performance in RTs (as described above). During titration, participants only received feedback at the end of each block. Immediately before entering the scanner, participants

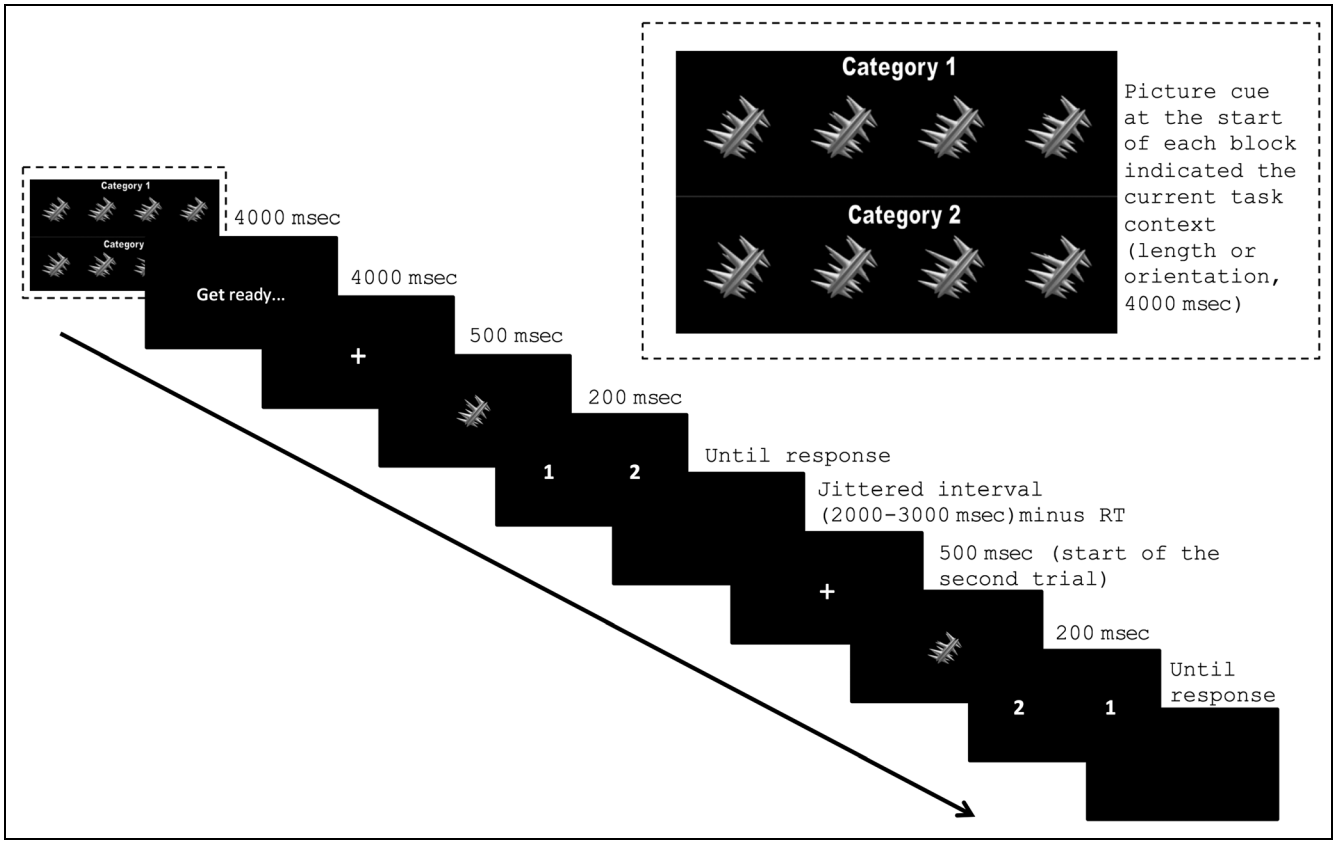

Figure 2. Stimulus categorization task: A picture cue at the start of each block indicated the current task context for categorization (orientation or length; inset shows cue display for the length task). On each trial, a fixation cross was presented for 500 msec followed by a spiky object for 200 msec. Finally, a response mapping screen appeared, which indicated the appropriate response button. In the example shown, the current context is length. For the first trial, the stimulus is Category 1 on the length dimension, and, therefore, the correct response was the left button.

completed a further two practice blocks of each task context to remind them of the task and to avoid initial practice effects in the scanner. These two practice blocks also introduced a response mapping screen to be used in the scanning task, which randomly assigned the button to be pressed for each category (short or long spikes in the length task and clockwise or anticlockwise spikes in the orientation task) on a trial-by-trial basis (Figure 2). This allowed separate estimation of the BOLD response associated with perceptual information about each category from that associated with each button press. Participants also performed an additional two practice blocks in the scanner during the structural scan, before commencing the main experiment, to familiarize them with the button-response box in the scanner.

Participants were scanned while performing the categorization task shown in Figure 2. Each participant completed four acquisition runs (8.09 min each) consisting of four blocks (32 trials/block), totaling 128 trials per acquisition run (2.02 min/block). At the start of each block, a picture cue (4000 msec) indicated the current task (length or orientation) and which attribute was Category 1 and 2 (e.g., whether short spikes/long spikes were Category 1 or 2; counterbalanced across participants). The order of task contexts was counterbalanced across participants as well as within participants across runs. The picture cue depicted spiky objects from the extremes of the currently relevant dimension (see Figure 2, inset). The stimulus set was identical across the two contexts, but the relevant feature was either the length of the same spike relative to the category boundary or the orientation of a particular spike relative to the category boundary (rotated clockwise vs. anticlockwise from the boundary) in the different contexts. Thus, participants were attending to the same part of the object, but different features of that object part, in the two conditions. On each trial, participants saw a white central fixation cross (500 msec), after which the spiky object was presented at fixation for 216 msec. Finally, participants saw a response mapping screen, which indicated the category-to-button response mapping on this trial, and responded regarding the category membership of the stimulus. The response mapping screen randomly assigned Category 1 and 2 decisions to either the left or right response button, operated by the index or middle finger of the participant's right hand. The response mapping screen was visible until a button press was made or until the jittered time interval timed out (2000–3000 msec). If a response was made before the end of the intertrial

interval, a blank black screen was shown for the remainder of the trial time. Feedback (accuracy score) was presented at the end of each block for 6000 msec, after which there was a delay of 4000 msec before the start of the next block. At the end of each run, a blank black screen was shown for 4000 msec.

After completion of the main task during the scanning session, we ran a localizer task to functionally identify the LOC as an a priori ROI. Participants viewed central intact and scrambled versions of black and white common objects in 16.8-sec blocks of 16 trials (1100 msec/trial), while fixating on a central cross. Participants had to indicate via a button response when the fixation cross changed color to remind them to fixate centrally. There were 21 blocks consisting of alternating blocks of whole objects, scrambled objects, and rest blocks (counterbalanced across participants). The EPI (acquisition time) for the localizer task was 6.25 min.

Data Acquisition

The data were collected using a 3-T Verio Siemens (Erlangen, Germany) MRI scanner at Macquarie Medical Imaging, Macquarie University Hospital, Sydney, Australia. We used a sequential descending T2*-weighted EPI acquisition sequence with the following parameters: acquisition time = 2000 msec, echo time = 30 msec, 34 oblique axial slices with a slice thickness of 3.0 mm and a 0.70-mm interslice gap, in-plane resolution = 3.0 × 3.0 mm, matrix = 64 × 64, field of view = 210 mm, and flip angle = 78°. T1-weighted magnetization prepared rapid gradient echo structural images were also acquired for all participants (slice thickness = 1.0 mm, resolution = 1.0 × 1.0 mm).

Preprocessing

MRI data were preprocessed using SPM 5 (Wellcome Department of Imaging Neuroscience, London, UK; www.fil.ion.ucl.ac.uk/spm) in MATLAB 2011b. Functional MRI data were converted from DICOM to NIFTII format, spatially realigned to the first functional scan and slice timing corrected, and structural images were co-registered to the mean EPI. EPIs from the main experiment were smoothed slightly (4-mm FWHM Gaussian kernel) to improve signal-to-noise ratio. LOC localizer EPIs were also smoothed (8-mm FWHM Gaussian kernel), and, in all cases, the data were high-pass filtered (128 sec). Structural scans were additionally normalized, using the segment and normalize routine of SPM5, to derive the individual participant normalization parameters needed for ROI definition (below).

ROIs

MD ROIs were defined using coordinates from a previous review of activity associated with a diverse set of cognitive demands (Duncan & Owen, 2000) using the kernel method described in Cusack, Mitchell, and Duncan (2010), as in our previous work (Woolgar, Williams, et al., 2015; Woolgar, Hampshire, et al., 2011; Woolgar, Thompson, et al., 2011). The procedure yielded seven ROIs: left and right IFS (center of mass = ±38, 26, 24; volume = 17 cm^3), left and right AI/FO (±35, 19, 3; 3 cm^3), left and right IPS (±35, −58, 41; 7 cm^3), and ACC/pre-SMA (0, 23, 39; 21 cm^3).

Left and right visual cortex ROIs were derived from the Brodmann's template provided with MRIcro (Rorden & Brett, 2000; Brodmann's area [BA] 17; center of mass = −13, −81, 3/16, −79, 3; volume = 54 cm^3). All coordinates are given in MNI152 space (McConnell Brain Imaging Centre, Montreal Neurological Institute). MD and BA 17 ROIs were deformed for each participant by applying the inverse of the participant's normalization parameters. This allowed analyses to be carried out using native space (i.e., nonnormalized) EPI data.

Using the functional localizer scan data, we defined LOC for each participant as the brain area that responded more strongly to whole objects than to scrambled versions of the same objects. We used the standard multiple regression approach of SPM5 (Wellcome Department of Imaging Neuroscience, London, UK; www.fil.ion.ucl.ac.uk) to estimate values pertaining to the whole and scrambled object conditions (block design). Blocks were modeled using a box car function lasting 16 sec convolved with the hemodynamic response of SPM5. The run mean was included in the model as a covariate of no interest. Whole-brain analyses (paired *t* tests) compared voxelwise BOLD response in the two conditions (whole objects minus scrambled objects). The resulting map was thresholded such that there was at least one cluster with a minimum size of 20 voxels. These clusters were then imported into MarsBaR (Brett, Anton, Valabregue, & Poline, 2002), and those active voxel clusters close to anatomical LOC coordinates from previous studies (Grill-Spector, Kushnir, Hendler, & Malach, 2000; Grill-Spector et al., 1999) were selected to form the ROI.

First-level Model

To obtain estimated activation patterns for multivariate analysis, a general linear model was estimated for each participant using the realigned, slice-time-corrected, and smoothed native space EPIs using SPM5 (Wellcome Department of Imaging Neuroscience, London, UK; www.fil.ion.ucl.ac.uk). We classified each stimulus as either "short" or "long" on the length dimension and "rotated clockwise" or "rotated anticlockwise" relative to the category boundary on the orientation dimension. Trials were modeled as events of zero duration convolved with the hemodynamic response of SPM5. Each trial contributed to the estimation of two beta values, the relevant feature (short or long for length context blocks and clockwise or anticlockwise for orientation context blocks) and the irrelevant feature (short or long for orientation context blocks and clockwise

or anticlockwise for length context blocks). We derived the estimates for each feature in each block separately. The two run means were included in the model as covariates of no interest. Error trials were excluded from the analysis.

MVPA

We used MVPA to examine the representation of relevant and irrelevant stimulus features. Of central interest was whether the MD regions adapted to code length and orientation information more strongly when it was relevant for the task than when it was task irrelevant. We also examined the same stimulus feature distinctions in the LOC and early visual cortex (BA 17).

We implemented MVPA using the Decoding Toolbox (Hebart, Görgen, & Haynes, 2015), which wraps the LIBSVM library (Chang & Lin, 2011). We examined coding of orientation when orientation was relevant, orientation when orientation was irrelevant, length when length was relevant, and length when length was irrelevant. For each participant and ROI, a linear support vector machine was trained to decode the relevant (clockwise/anticlockwise in orientation blocks and short/long in length blocks) and irrelevant (clockwise/anticlockwise in length blocks and short/long in orientation blocks) stimulus features for each task context separately. In total, there were 16 blocks for each participant: eight with relevant length and eight with relevant orientation. For each classification, we used a leave-one-out eight-fold splitter whereby the classifier was trained using the data from seven of the eight blocks and subsequently tested on its accuracy at classifying the unseen data from the remaining block, iterating over all possible combinations of training and testing blocks. For example, to yield a classification accuracy score for the task-relevant length distinctions, we took the eight blocks in which participants performed the length task and trained the classifier to distinguish between patterns of activation representing short and long spikes in seven of these eight blocks and then tested generalization to the remaining unseen block. The accuracies were then averaged to give a mean accuracy score for task-relevant length coding. This was repeated for each condition, participant, and ROI separately.

The mean classification accuracy for each participant in each ROI and in each condition was then entered into a second level analysis. Of central interest was whether the MD network would code task-relevant stimulus features more strongly than task-irrelevant ones. To address this, we conducted a three-factor ANOVA on classifier accuracy with the factors Relevancy, Feature, and MD region. To explore any hemispheric effects, we ran an additional ANOVA with factors Relevancy, Feature, MD region, and Hemisphere.

Because a difference in coding between relevant and irrelevant conditions is only interpretable if coding in at least one condition is also significantly above chance, we also conducted one-sample t tests against the classification accuracy expected by chance (50%) in each condition (relevant and irrelevant) separately. One-tailed significance tests were used where appropriate for inference: Tests comparing classification of relevant with irrelevant feature distinctions in the MD regions are one tailed because the direction of the effect is prespecified, and tests comparing classification accuracy with chance are one tailed as below-chance classifications are not interpretable. All other tests are two tailed. Alpha was adjusted for four comparisons using Bonferroni correction (.05 divided by 4).

We also examined whether coding in the visual cortices was stronger for task-relevant than task-irrelevant stimulus features. For this, we used a two-factor ANOVA on classifier accuracy with factors Relevancy (task relevant, task irrelevant) and Feature (orientation, length) for each of the visual cortex ROIs (LOC, early visual cortex) collapsed across hemispheres. Again, we also tested whether coding in the visual cortices was above chance in each condition separately (one-sample t test against the classification accuracy expected by chance, 50%).

We conducted an additional analysis in which the classifier was trained on data representing the category number decisions in one task context (Category 1 or 2) and tested on the category number decisions participants made in the other task context (Category 1 or 2). We included this analysis to explore whether the categorization decision was represented at the level of the stimulus (i.e., short/long) or at the level of the category number (Category 1 or 2).

Finally, in addition to the standard parametric statistics outlined above, we conducted a two-step permutation test (Stelzer, Chen, & Turner, 2013) to compare coding of relevant and irrelevant information to chance and coding of relevant and irrelevant information to each other, in a nonparametric framework. For each decoding analysis (e.g., short/long in length blocks), each ROI, and each person, we exhaustively permuted the class labels within each block (128 combinations) and trained and tested the classifier on each permutation. Next, we built three group-level null distributions, one for relevant information (short/long in length blocks and clockwise/anticlockwise in orientation blocks), one for irrelevant information (short/long in orientation blocks and clockwise/anticlockwise in length blocks), and one for the difference between relevant and irrelevant information (relevant minus irrelevant). Each null distribution was created by sampling (with replacement) 10,000 times from the set of participants × 128 relevant permutation classification results (one sample per participant per iteration, single-participant data collapsed permutation-wise over length and orientation). Finally, we calculated the probability p of observing each decoding accuracy (from the correctly labeled data) given the relevant group null distribution, using the Monte-Carlo approach (Hammersley, 2013), in which $p = (k + 1) / (n + 1)$ where k is the number of permutations in the group null with equal or higher accuracy to the actual value and n is the number

of all permutations in the group null. For the comparison of relevant and irrelevant information, we compared the difference in observed classification accuracy (relevant minus irrelevant) to the null distribution derived from the difference between the relevant and irrelevant classification accuracies on each permutation. For the four MD regions, alpha was again adjusted for four comparisons using Bonferroni correction.

RESULTS

Behavioral Results

Before scanning, participants practiced the task until they scored at least 80% correct in both task conditions. Task difficulty was then titrated to match RTs between the two conditions for each participant separately (assessed with BF analysis for each participant separately, all $BF_{10} < 0.89$).

In the scanning session, participants performed with a high degree of accuracy (mean = 89.6%, *SD* = 9.1%). Accuracy scores were assessed using BF analysis to check for differences in performance between the two task contexts. There were no differences in accuracy score between the two conditions for any participant individually (all $BF_{10} < 0.64$). RT data from the scanning session are not meaningful as the response mapping screen defined the response after the stimulus display.

Decoding Task-relevant and Task-irrelevant Stimulus Features

MD Regions

MVPA was used to differentiate multivoxel patterns pertaining to stimulus feature distinctions (orientation: clockwise/anti-clockwise, length: short/long) when they were relevant to the task (orientation in orientation task blocks and length in length task blocks) and when they were irrelevant to the task (orientation in length task blocks and length in orientation task blocks). The resulting classification accuracy signified the relative strength of coding.

The adaptive coding hypothesis (Duncan, 2001) proposes that neurons dynamically adjust their responses to selectively code information that is currently relevant for our behavior. We asked whether this could provide a basis for feature-selective attention, with preferential coding of currently relevant visual features in the MD system. Our prediction was that the MD regions would change their representation of the visual stimuli between tasks to reflect or emphasize the distinctions that were needed in each task context.

The results are presented in Figure 3. In line with the hypothesis, a three-way ANOVA with factors Relevancy (task-relevant and task-irrelevant stimulus features), Region (AI/FO, IFS, ACC/pre-SMA, and the IPS; collapsed across hemispheres where appropriate), and Feature (orientation and length) revealed a main effect of Relevancy ($F(1, 25) = 1.13, p = .03$). The ANOVA showed no main effect of Feature ($F(1, 25) = 1.01, p = .32$); no main effect of Region ($F(3, 75) = 0.77, p = .51$); no significant interaction between Relevancy and MD region ($F(3, 75) = 2.05, p = .12$), Region and Feature ($F(3, 75) = 1.75, p = .16$), or Relevancy and Feature ($F(3, 75) = 0.21, p = .65$); and no significant three-way interaction ($F(3, 75) = 1.41, p = .25$). These results indicate that the MD regions coded the task-relevant feature distinctions more strongly than the task-irrelevant distinctions, despite these features being actually physically identical. In our additional ANOVA, included to check for hemispheric differences, there was

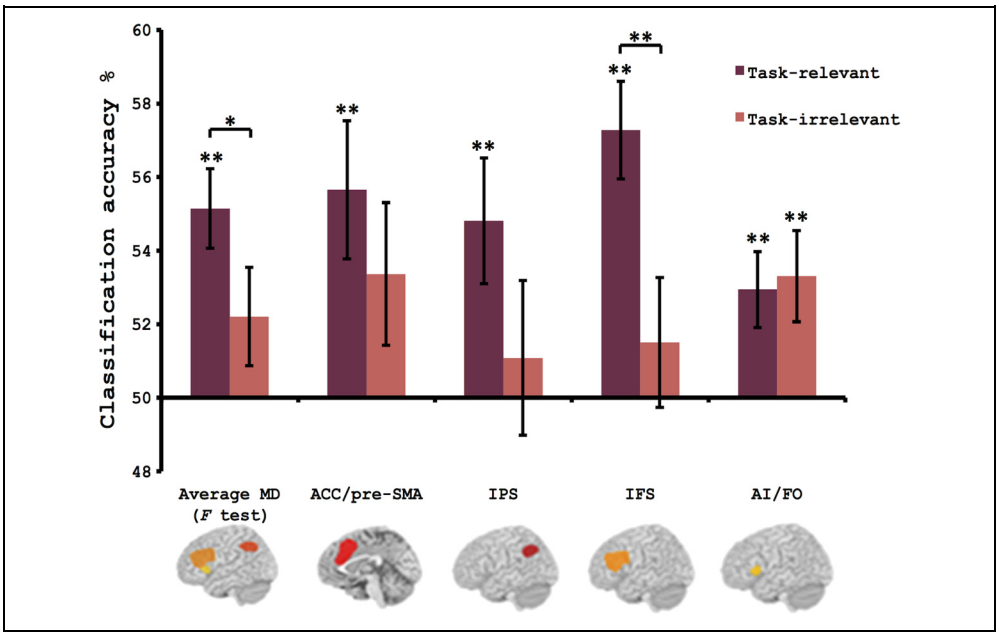

Figure 3. Coding of task-relevant and task-irrelevant stimulus distinctions in the MD regions. Error bars indicate *SE*. Significance markings for individual bars indicate whether coding was significantly greater than chance in each condition separately (one-sample *t* test against chance = 50%); significance markings between bars indicate where coding was significantly greater for relevant compared with irrelevant distinctions (main effect of relevancy/paired *t* test). $p < .05$, $^{**}p < .01$—alpha for the analyses of individual regions corrected for four comparisons using Bonferroni correction. The MD regions coded task-relevant stimulus distinctions more strongly than the physically identical task-irrelevant distinctions.

no main effect of Hemisphere ($F(3, 24) = 2.52, p = .13$), but there was a significant interaction between Relevancy and Hemisphere ($F(1, 24) = 5.78, p = .02$), reflecting a stronger relevancy effect on the left. Overall, then, we see evidence that the MD regions code relevant stimulus features in preference to irrelevant ones.

The coding of relevant over irrelevant features is only interpretable if coding in one or more of the relevancy conditions is also significantly above chance. Therefore, we conducted one-sample t tests against the classification accuracy expected by chance (50%) in each relevancy condition separately. We found that the MD regions coded the task-relevant stimulus features significantly (mean classification accuracy for relevant across all regions = 55.14%; one-sample t test against chance, $t(25) = 4.75$, $p < .001$), whereas classification of the task-irrelevant stimulus distinctions was not significantly different from chance (mean classification accuracy for irrelevant = 52.21%; $t(25) = 1.65, p = .11$). Thus, at the network level, these regions only encoded the task-relevant stimulus distinctions. Considering each MD ROI separately, the relevant stimulus distinctions were coded in all four MD ROIs (ACC/pre-SMA: mean accuracy = 55.64%, $t(25) = 3.01, p = .004$; IPS: mean accuracy = 54.80%, $t(25) = 2.81$, $p = .01$; IFS: mean accuracy = 57.27%, $t(25) = 5.47, p < .001$; AI/FO: mean accuracy = 52.94%, $t(25) = 2.85, p = .01$; Figure 3, dark bars), whereas the irrelevant stimulus distinctions were only coded in the AI/FO (mean classification accuracy = 53.31%, $t(25) = 2.66, p = .01$; other ps $> .05$; Figure 3, light bars).

We also conducted two-step permutation tests (Stelzer et al., 2013) to assess the significance of our results in a nonparametric framework. In this analysis, coding of both the relevant ($p < .001$) and irrelevant ($p = .03$) information was above chance in the MD system overall, but coding of relevant information was still significantly stronger than that of irrelevant information ($p = .03$). Considering each region individually, relevant information was coded in three regions (ACC/pre-SMA: $p = .002$, IPS: $p = .002$, IFS: $p < .001$), whereas irrelevant information did not reach our Bonferroni-corrected significance threshold in any region (all ps $> .02$). The difference between relevant and irrelevant coding only reached significance in the IFS ($p = .006$; all other ps $> .05$).

Recall that the stimulus set was identical across conditions; each feature (length and orientation) was relevant in one condition and irrelevant in the other. The data suggest that the MD system encoded the relevant feature in each case; that is, it adjusted such that it coded the same physical stimulus distinction (e.g., length) more strongly when it was relevant than when it was irrelevant.

LOC

The LOC is known to respond strongly to object features (Grill-Spector, Kourtzi, & Kanwisher, 2001) and has previously been found to show preferential representation of attended relative to distractor objects (Woolgar, Williams, et al., 2015). It therefore seemed a likely candidate for preferential coding of relevant object features in our task. Figure 4 presents the data from the LOC, and visual inspection suggests a trend in this direction. However, an ANOVA with factors Relevancy (relevant, irrelevant) and Feature (orientation, length) showed no significant main effect of Relevancy ($F(1, 25) = 2.83, p = .11$), no main effect of Feature ($F(1, 25) = 0.32, p = .57$), and no Relevancy × Feature interaction ($F(1, 25) = 0.08$, $p = .78$). When we compared coding with chance, the LOC carried significant information about task-relevant distinctions (mean classification accuracy = 56.67%, $t(25) = 3.65, p < .001$) but not about irrelevant distinctions (mean classification accuracy = 52.63%, $t(25) = 1.39, p = .15$). The results of the permutation test mirrored these data, with coding of relevant information significantly above chance ($p < .001$) whereas coding of irrelevant features was not ($p = .09$). There was a nonsignificant trend toward stronger coding of relevant features ($p = .07$). Thus, although the LOC only discriminated objects along the stimulus dimension that was currently relevant, the interaction with relevancy did not reach significance.

Early Visual Cortex

We also tested whether information pertaining to task-relevant and task-irrelevant stimulus distinctions was coded in the early visual cortex (BA 17). An ANOVA with factors Relevancy (relevant, irrelevant) and Feature (orientation, length) showed no main effect of Relevancy ($F(1, 25) = 0.57, p = .81$), no main effect of Feature ($F(1, 25) = 2.1, p = .17$), and no Relevancy × Feature interaction ($F(1, 25) = 2.06, p = .16$). Thus, we found no evidence that context modulates coding of orientation and length in this

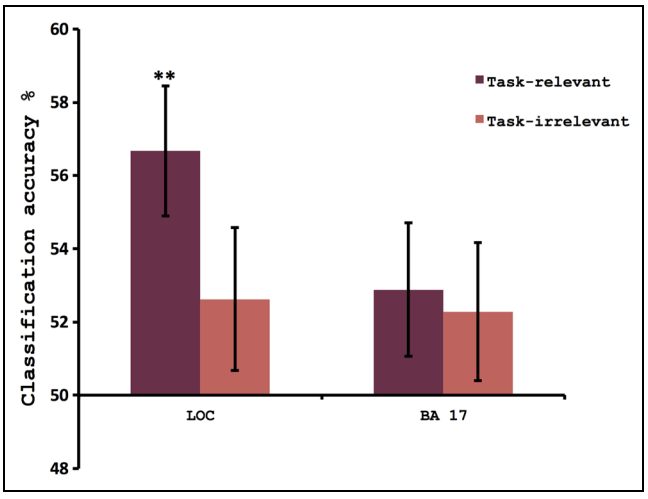

Figure 4. Coding of task-relevant and task-irrelevant stimulus distinctions in LOC and BA 17. Error bars indicate *SE*. The significance markings for individual bars indicate whether coding was significantly greater than chance in each condition separately (one-sample t test against chance = 50%). **$p < .01$.

region. Furthermore, BA 17 did not show categorical discrimination of these visually similar objects according to either the task-relevant (relative to chance; mean classification accuracy = 52.91%, $t(25) = 1.62$, $p = .14$) or task-irrelevant (relative to chance; mean classification accuracy = 52.31%, $t(25) = 1.17$, $p = .25$) stimulus features. The two-step permutation tests similarly showed no evidence that BA 17 coded the task-relevant over the task-irrelevant features ($p = .41$). In addition, coding of relevant ($p = .08$) and irrelevant ($p = .12$) features did not reach significance. As the stimuli were identical in the different task contexts, it is not at all surprising to find that the early visual cortex information does not distinguish between these stimuli, although, as with all null effects, we must be cautious about interpretation.

Coding of Category Placement

On each trial, participants had to categorize the object according to the relevant feature dimension (e.g., short/long on length blocks), associate that decision with the number representing the chosen category (1 or 2, e.g., short = 1 in half of the participants), and then use the response mapping screen to transform their choice into the appropriate button press response (left or right, 1 = left in half of the trials). Therefore, it is possible that, as well as the categorization decision at the level of the stimulus, participants also held a category number in mind on each trial. To test for representation at this higher level of abstraction, we ran an additional analysis where we trained the classifier on the data representing the category number decisions in one task context and tested this on the category number decisions participants made in the other task context. We were not able to decode the category number placement of the objects in the MD system (mean classification accuracy = 50.54%, $t(25) = 1.31$, $p = .21$). To interpret this null effect, we calculated the BF. Coding of category number placement revealed a BF_{10} of 0.44. As this is less than 1 (Dienes, 2011) and approaches the level of 0.33 suggested by Jeffreys (1998) to represent significant evidence for the null hypothesis, we interpret this as evidence that, although the MD regions encode task-relevant stimulus distinctions, the representation is not abstracted to the level of categorical number placement.

DISCUSSION

The adaptive coding hypothesis (Duncan, 2001) proposes that neural populations dynamically adjust their responses to selectively code information that is currently relevant for our behavior. This provides a possible mechanism for feature-selective attention, which allows information about task-relevant stimulus features to be processed in preference to irrelevant attributes. We examined the responses of the MD regions in a difficult visual object categorization task in which the relevant stimulus dimension varied on physically identical stimuli. The MD system adjusted its representation of these novel objects to preferentially encode feature distinctions that were relevant for the task. When the task required participants to categorize the objects based on length, the MD regions coded length information more strongly than orientation information, but when the task was to categorize based on orientation, orientation was encoded in preference to length. Thus, the MD system adjusted its representation of the features of an object to encode the discrimination necessary for the current task. Consistent with the proposal that the cognitive flexibility of these regions underlies their involvement in a wide range of tasks (e.g., Duncan, 2010; Cole & Schneider, 2007; Duncan & Owen, 2000), our data suggest that the coding of this adaptive system adjusts to hold the currently relevant features of a stimulus as needed for behavior.

Electrophysiology studies in nonhuman primates have previously shown that neurons in higher cortical regions adapt their tuning profiles to respond most strongly to the information that is currently relevant (Stokes et al., 2013; Cromer et al., 2010; Roy et al., 2010; Freedman & Assad, 2006; Freedman, 2001; Sakagami & Niki, 1994). The implementation of MVPA for fMRI has shown similar results in humans: Patterns of activation in the MD regions code a range of different types of task-related information (e.g., Woolgar, Williams, et al., 2015; Nee & Brown, 2012; Reverberi et al., 2011; Woolgar, Hampshire, et al., 2011; Woolgar, Thompson, et al., 2011; Bode & Haynes, 2009; Haynes et al., 2007; Li et al., 2007) and adjust their responses when task demands vary (Woolgar, Afshar, et al., 2015; Woolgar, Williams, et al., 2015; Woolgar, Hampshire, et al., 2011; Li et al., 2007). The MD regions also encode attended objects in preference to unattended objects (Woolgar, Williams, et al., 2015), and a previous adaptation study demonstrated that these regions show greater responses to changes in attended stimulus features (color/shape) than to changes in unattended stimulus features (Thompson & Duncan, 2009). Here, we find that these regions can also flexibly adapt their representations of single objects to emphasize task-relevant stimulus distinctions, resulting in preferential coding of attended stimulus features.

Our data align with a recent study in which objects were strongly represented in lPFC in individual task contexts, but the representation did not generalize between task contexts (Harel et al., 2014). Those data suggested that the same set of objects may be represented differently as task contexts change. Here, we tested this possibility directly by specifying the specific stimulus distinctions that an adaptive system should make in each task context. We found that the MD system adjusted its representation of the set of novel objects to make the specific distinctions needed for the task. In a related study (Peelen & Caramazza, 2012), participants responded to one of two semantic dimensions of an object (how the object is used or where the object is found) in a 1-back task. Results from a

whole-brain searchlight revealed several regions, including the right lPFC, which showed coding of the two semantic dimensions. However, no region showed preferential coding of task-relevant over task-irrelevant dimensions. The MD regions are known to be recruited most strongly when tasks are challenging (e.g., Duncan & Owen, 2000). It may be, then, that preferential coding of task-relevant information is only observed when the task is sufficiently difficult (Woolgar, Afshar, et al., 2015; Woolgar, Hampshire, et al., 2011), as in the current experiment.

The observation that stimulus features were coded more strongly when they were relevant than when they were irrelevant suggests that some filtering of information occurs between input (the relevant and irrelevant features were physically identical in our case) and MD representation as recorded with fMRI. In our data, we could not detect MD coding of irrelevant information in three of the four MD regions. However, recent evidence from nonhuman primates suggests that task-irrelevant information can affect firing rates in higher cortical regions such as the FEFs (Mante, Sussillo, Shenoy, & Newsome, 2013). Differences in the tasks, sensitivity of the methods and recording site, may account for the different results. In that study, consistent with our findings, the effect of irrelevant information was weaker than that of relevant information, but interestingly, the size of the difference there was too small to account for the behavioral effect (Mante et al., 2013). Other work suggests a dynamic change in the responses of lateral prefrontal neurons over time (Kadohisa et al., 2013). When presented with a target and distractor object in the left and right visual fields, prefrontal activity was initially dominated by the contralateral object, regardless of its relevance, but over the course of the trial, prefrontal resources were quickly reassigned such that the representation of the target came to dominate in both hemispheres (Kadohisa et al., 2013).

In our data, the LOC held information about the task-relevant feature distinctions, demonstrating that it is sensitive to minimal change in the shape of an object (e.g., to small changes in the length of one spike) when that change is relevant for behavior. The LOC did not make the task-irrelevant distinctions, but the difference in coding between relevant and irrelevant conditions did not reach significance. The trend for relevant coding to be greater than irrelevant coding is in line with previous work, which has emphasized a role for the LOC in responding to attended compared with unattended objects (e.g., Woolgar, Williams, et al., 2015; Konen & Kastner, 2008; Murray & He, 2006; Xu & Chun, 2005; Murray & Wojciulik, 2003). It seems likely that the magnitude of feature-selective attention effects, as in our study, would be considerably smaller than effects of attention allocated on a whole-object level. Although we could not look at it in this study, it is possible that similar patterns might occur in other brain regions such as the FEFs, based on previous findings that they are reliably activated by attentional tasks in humans (e.g., Ester, Sprague, & Serences, 2015; Jerde, Merriam, Riggall, Hedges, & Curtis, 2012; Culham, Cavanagh, & Kanwisher, 2001; Corbetta et al., 1998) and in nonhuman primates (Mante et al., 2013).

We did not observe multivoxel coding of object-category information in the early visual cortex (BA 17) or any interaction with relevancy. Although previous investigations using MVPA with fMRI have reported preferential coding of task-relevant stimuli in this region (Woolgar, Williams, et al., 2015; Kok, Jehee, & de Lange, 2012; Jehee, Brady, & Tong, 2011), there are some marked differences between these and the current study that account for the different results. In a previous work, coding corresponded to discrimination between physically dissimilar stimuli (gratings of 55° and 145° orientation in Jehee et al., 2011; 40° and 135° in Kok et al., 2012; and different objects in Woolgar, Williams, et al., 2015). In our task, coding corresponded to discriminations across an arbitrary boundary on identical sets of stimuli. Objects close to the decision boundary were physically very similar, meaning that stimulus-driven activation patterns would also be very similar and therefore difficult to classify. Moreover, the objects on either side of the category boundary were collapsed in our analyses, meaning that the classifier was required to generalize over physical differences of a similar magnitude to those it needed to discriminate between. This makes our result in the MD regions all the more striking, because the information they encoded was based on such minimal visual differences.

Successful behavior requires a flexible cognitive system. Here, we have demonstrated that the frontoparietal cortex adjusts its representation of visual objects to make the distinctions that are needed for the current task. In this way, visually minimal task-relevant feature distinctions are coded more strongly than the equivalent irrelevant distinctions. This study exemplifies the extent to which the MD network can flexibly emphasize different features of an object, providing a possible neural mechanism for the implementation of feature-selective attention.

Acknowledgments

J. J. was supported by an International Macquarie University Research Excellence Scholarship from Macquarie University, and A. W., M. A. W., and A. N. R. have been recipients of ARC Fellowships (DE120100898, DP0984919, and DP0984494, respectively). We thank Hans Op de Beeck for providing the stimuli. This work was funded by the Australian Research Council's Discovery Project funding scheme (DP12102835) to A. W., A. N. R., and M. A. W. and the Macquarie University Cognitive Science Postgraduate Grant to J. J.

Reprint requests should be sent to Alexandra Woolgar, Level 3, 16 University Avenue, Macquarie University, NSW, 2109, Australia, or via e-mail: alexandra.woolgar@mq.edu.au.

REFERENCES

Bode, S., & Haynes, J.-D. (2009). Decoding sequential stages of task preparation in the human brain. *Neuroimage, 45,* 606–613.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Brett, M., Anton, J., Valabregue, R., & Poline, J. (2002). Region of interest analysis using an SPM toolbox. *Neuroimage, 16,* S497.
Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology, 2,* 27.
Chen, X., Hoffmann, K. P., Albright, T. D., & Thiele, A. (2012). Effect of feature-selective attention on neuronal responses in macaque area MT. *Journal of Neurophysiology, 107,* 1530–1543.
Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., & Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience, 16,* 1348–1355.
Cole, M. W., & Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *Neuroimage, 37,* 343–360.
Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., et al. (1998). A common network of functional areas for attention and eye movements. *Neuron, 21,* 761–773.
Cromer, J. A., Roy, J. E., & Miller, E. K. (2010). Representation of multiple, independent categories in the primate prefrontal cortex. *Neuron, 66,* 796–807.
Culham, J. C., Cavanagh, P., & Kanwisher, N. G. (2001). Attention response functions: Characterizing brain areas using fMRI activation during parametric variations of attentional load. *Neuron, 32,* 737–745.
Cusack, R., Mitchell, D. J., & Duncan, J. (2010). Discrete object representation, attention switching, and task difficulty in the parietal lobe. *Journal of Cognitive Neuroscience, 22,* 32–47.
Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science, 6,* 274–290.
Dosenbach, N. U. F., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., et al. (2006). A core system for the implementation of task sets. *Neuron, 50,* 799–812.
Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience, 2,* 820–829.
Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences, 14,* 172–179.
Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobes recruited by diverse cognitive demands. *Trends in Neurosciences, 23,* 475–483.
Ester, E. F., Sprague, T. C., & Serences, J. T. (2015). Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron, 87,* 893–905.
Fedorenko, E., Duncan, J., & Kanwisher, N. G. (2013). Broad domain-generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 110,* 16616–16621.
Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). From the cover: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 9673–9678.
Freedman, D. J. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science, 291,* 312–316.
Freedman, D. J., & Assad, J. A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature, 443,* 85–88.
Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. G. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 21,* 1409–1422.
Grill-Spector, K., Kushner, T., Edelman, S., Avidan, G., Itzchak, 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., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience, 3,* 837–843.
Hammersley, J. (2013). *Monte Carlo methods*. Berlin, Germany: Springer Science & Business Media.
Harel, A., Kravitz, D., & Baker, C. I. (2014). Task context impacts visual object processing differentially across the cortex. *Proceedings of the National Academy of Sciences, U.S.A., 111,* E962–E971.
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,* 2425–2429.
Haynes, J. D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience, 8,* 686–691.
Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R. E. (2007). Reading hidden intentions in the human brain. *Current Biology, 17,* 323–328.
Hebart, M. N., Görgen, K., & Haynes, J.-D. (2015). The Decoding Toolbox (TDT): A versatile software package for multivariate analyses of functional imaging data. *Frontiers in Neuroinformatics, 8,* 88. doi:10.3389/fninf.2014.00088
Jeffreys, H. (1998). *The theory of probability*. Oxford, UK: OUP Oxford.
Jehee, J. F. M., Brady, D. K., & Tong, F. (2011). Attention improves encoding of task-relevant features in the human visual cortex. *Journal of Neuroscience, 31,* 8210–8219.
Jerde, T. A., Merriam, E. P., Riggall, A. C., Hedges, J. H., & Curtis, C. E. (2012). Prioritized maps of space in human frontoparietal cortex. *Journal of Neuroscience, 32,* 17382–17390.
Kadohisa, M., Petrov, P., Stokes, M., Sigala, N., Buckley, M., Gaffan, D., et al. (2013). Dynamic construction of a coherent attentional state in a prefrontal cell population. *Neuron, 80,* 235–246.
Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience, 8,* 679–685.
Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron, 75,* 265–270.
Konen, C. S., & Kastner, S. (2008). Two hierarchically organized neural systems for object information in human visual cortex. *Nature Neuroscience, 11,* 224–231.
Li, S., Ostwald, D., Giese, M., & Kourtzi, Z. (2007). Flexible coding for categorical decisions in the human brain. *Journal of Neuroscience, 27,* 12321–12330.
Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., et al. (2015). *JASP (Version 0.7) [computer software]*. Amsterdam, The Netherlands: JASP Project.
Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature, 503,* 78–84.
Murray, S. O., & He, S. (2006). Contrast invariance in the human lateral occipital complex depends on attention. *Current Biology, 16,* 606–611.
Murray, S. O., & Wojciulik, E. (2003). Attention increases neural selectivity in the human lateral occipital complex. *Nature Neuroscience, 7,* 70–74.
Nee, D. E., & Brown, J. W. (2012). Rostral–caudal gradients of abstraction revealed by multi-variate pattern analysis of working memory. *Neuroimage, 63,* 1285–1294.

Nyberg, L., Marklund, P., Persson, J., Cabeza, R., Forkstam, C., Petterson, K. M., et al. (2003). Common prefrontal activations during working memory, episodic memory, and semantic memory. *Neuropsychologia, 41,* 371–377.
Op de Beeck, H. P., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *Journal of Neuroscience, 26,* 13025–13036.
Peelen, M. V., & Caramazza, A. (2012). Conceptual object representations in human anterior temporal cortex. *Journal of Neuroscience, 32,* 15728–15736.
Reverberi, C., Gorgen, K., & Haynes, J. D. (2011). Compositionality of rule representations in human prefrontal cortex. *Cerebral Cortex, 22,* 1237–1246.
Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology, 12,* 191–200.
Roy, J. E., Riesenhuber, M., Poggio, T., & Miller, E. K. (2010). Prefrontal cortex activity during flexible categorization. *Journal of Neuroscience, 30,* 8519–8528.
Sakagami, M., & Niki, H. (1994). Encoding of behavioral significance of visual stimuli by primate prefrontal neurons: Relation to relevant task conditions. *Experimental Brain Research, 97,* 423–436.
Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. *Neuroimage, 65,* 69–82.
Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron, 78,* 364–375.
Thompson, R., & Duncan, J. (2009). Attentional modulation of stimulus representation in human frontoparietal cortex. *Neuroimage, 48,* 436–448.
Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology, 100,* 3328–3342.
Woolgar, A., Afshar, S., Williams, M. A., & Rich, A. N. (2015). Flexible coding of task rules in frontoparietal cortex: An adaptive system for flexible cognitive control. *Journal of Cognitive Neuroscience, 27,* 1895–1911.
Woolgar, A., Hampshire, A., Thompson, R., & Duncan, J. (2011). Adaptive coding of task-relevant information in human frontoparietal cortex. *Journal of Neuroscience, 31,* 14592–14599.
Woolgar, A., Thompson, R., Bor, D., & Duncan, J. (2011). Multi-voxel coding of stimuli, rules, and responses in human frontoparietal cortex. *Neuroimage, 56,* 744–752.
Woolgar, A., Williams, M. A., & Rich, A. N. (2015). Attention enhances multi-voxel representation of novel objects in frontal, parietal and visual cortices. *Neuroimage, 109,* 429–437.
Xu, Y., & Chun, M. M. (2005). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature, 440,* 91–95.