Characterizing object- and position-dependent response profiles to uni- and bilateral stimulus configurations in human higher visual cortex

CrossMark

Characterizing object- and position-dependent response profiles to uni- and bilateral stimulus configurations in human higher visual cortex: a 7T fMRI study

Joel Reithler[a,b,c,*,1], Judith C. Peters[a,b,c,1], Rainer Goebel[a,b,c]

[a] *Cognitive Neuroscience Department, Faculty of Psychology and Neuroscience, Maastricht University, 6229 EV Maastricht, The Netherlands*
[b] *Maastricht Brain Imaging Center (M-BIC), Faculty of Psychology and Neuroscience, Maastricht University, 6229 EV Maastricht, The Netherlands*
[c] *Department of Neuroimaging and Neuromodeling, Netherlands Institute for Neuroscience, Royal Netherlands Academy of Arts and Sciences (KNAW), 1105 BA Amsterdam, The Netherlands*

ABSTRACT

Visual scenes are initially processed via segregated neural pathways dedicated to either of the two visual hemifields. Although higher-order visual areas are generally believed to utilize invariant object representations (abstracted away from features such as stimulus position), recent findings suggest they retain more spatial information than previously thought.

Here, we assessed the nature of such higher-order object representations in human cortex using high-resolution fMRI at 7T, supported by corroborative 3T data. We show that multi-voxel activation patterns in both the contra- and ipsilaterally hemisphere can be exploited to successfully classify the object category of unilaterally presented stimuli. Moreover, robustly identified rank order-based response profiles demonstrated a strong contralateral bias which frequently outweighed object category preferences. Finally, we contrasted different combinatorial operations to predict the responses during bilateral stimulation conditions based on responses to their constituent unilateral elements. Results favored a max operation predominantly reflecting the contralateral stimuli. The current findings extend previous work by showing that configuration-dependent modulations in higher-order visual cortex responses as observed in single unit activity have a counterpart in human neural population coding. They furthermore corroborate the emerging view that position coding is a fundamental functional characteristic of ventral visual stream processing.

Introduction

The way in which the primate visual system processes incoming information is remarkably fast, flexible and accurate. The apparent ease with which an observer detects, segments and recognizes different elements of a visual scene is particularly striking when taking into account the vast amount of different objects that can potentially be encountered. Our visual system therefore has to display a high degree of *selectivity* in order to correctly distinguish between classes of objects or individual exemplars within an object category. Various levels of selectivity have indeed been clearly demonstrated in the neural responses of higher-order visual areas (e.g., Desimone et al., 1984; Tsao et al., 2006). However, any given object can be viewed in a wide variety of circumstances, leading to marked variations in the generated retinal input across different encounters (for example, in terms of size,

occlusion, illumination etc.). A certain degree of *invariance* is therefore required to still recognize an object despite changes in its appearance, without confusing it with any of the alternatives. The exact brain mechanisms enabling fast and reliable recognition independent of such 'identity preserving transformations' are still poorly understood (DiCarlo et al., 2012).

A key example of the described flexibility of the visual system relates to the identification of an object irrespective of where it is presented within the visual field. Whereas visual information is initially processed by segregated neural pathways depending on the visual hemifield being stimulated, later processing stages in higher visual cortex show response profiles compatible with a more abstract coding scheme of object identity independent of lower-level stimulus features such as position (Schwartz et al., 1983; Sary et al., 1993; Tovee et al., 1994; Ito et al., 1995). However, although the *relative* stimulus

* Correspondence to: Cognitive Neuroscience Department, Faculty of Psychology & Neuroscience, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands.
E-mail address: j.reithler@maastrichtuniversity.nl (J. Reithler).
[1] These authors contributed equally

preference in higher-order visual cortex is preserved across different retinal positions, *absolute* response amplitudes for the same stimuli can still vary in a position-dependent manner (suggesting varying degrees of "position tolerance", rather than strict "invariance"). One form of position-dependent coding becomes apparent in enhanced responses to isolated stimuli in the contralateral compared to ipsilateral hemifield, observable in monkey single unit (Schwartz et al., 1983, Ito et al., 1995; but see Tovee et al., 1994), as well as monkey (Rajimehr et al., 2014) and human (Hemond et al., 2007) neuroimaging data. These and complementary findings (Schwarzlose et al., 2008; Kravitz et al., 2010; Carlson et al., 2011; Rauschecker et al., 2012; MacEvoy and Yang, 2012) suggest that although fine-grained retinotopic localization is sacrificed to promote flexibility, neural representations in higher visual areas are still sufficiently rich to incorporate position information to some extent (Goebel, 2012). In fact, while earlier studies put emphasis on the large receptive fields of neurons in inferotemporal cortex (spanning 10–30° in both dimensions [Gross et al., 1969; Tanaka et al., 1991]), more recent studies highlight a wide variety of observed receptive field sizes (from 2.8° to 26° [Op de Beeck and Vogels, 2000]), up to the level that small variations in stimulus position (in the range of ± 1.5°) could still be disentangled (DiCarlo and Maunsell, 2003). One question that arises therefore is how such differences in receptive field properties play out when considering population responses reflecting the activity of many individual neurons (Kravitz et al., 2008). More specifically, it is currently still unclear how to best characterize responses within large (assembled) receptive fields spanning both hemifields (possibly underlying position tolerance) when half of the required input is only accessible through callosal fibers after being (pre-)processed by the opposite hemisphere. Moreover, most of what is known about response properties embedded in higher visual activation patterns is based on stimulation conditions in which a single stimulus is shown in isolation. Although some interaction effects have been described (e.g., Rolls and Tovee, 1995; Rolls et al., 2003), it remains largely unknown to what extent object representations in higher-order visual areas reflect the properties of multiple stimuli when they occur simultaneously, as often is the case in more natural visual scenes.

Here, we used functional Magnetic Resonance Imaging (fMRI) at 7T, complemented with additional 3T data, to address these questions. Participants saw stimuli of two different object classes (presented in blocks) while performing a visual detection task. Whereas fixation was kept at the middle of the screen, stimuli were presented at peripheral locations (at 6° eccentricity in the lower visual quadrant). Crucially, a stimulus could either be presented in isolation in one hemifield (unilateral stimulation), or could be combined with a simultaneously presented stimulus in the opposite hemifield (bilateral stimulation). Focusing on the entire ensemble of visually responsive voxels in occipitotemporal cortex, we assessed to what extent object category decoding was possible based on distributed voxel activation patterns in either hemisphere both within and across hemifields, as previously shown for category-selective regions at 3T (e.g., Cichy et al., 2011; 2013). Furthermore, we examined how object position and identity preferences are reflected in (rank ordered) voxel-wise response amplitudes. In addition, we quantified the consistency of the overlap in activations to uni- and bilateral stimulus configurations across subjects. Finally we explored how to optimally predict the response to bilateral stimulus configurations based on the responses to their constituent elements shown in isolation.

Materials and methods

In the sections below, details regarding the included participants, the employed design, the used scanning parameters and the performed analyses are provided for the main experiment at 7T. A number of findings were corroborated by a complementary 3T dataset (n=10) whenever fair comparisons could be drawn. The relevant information related to this 3T dataset is listed in the supplementary online material.

Participants

Six healthy volunteers (3 female; mean age = 31 years [SD = 6.9]) with normal or corrected to normal vision participated in the 7T experiment. All participants gave written informed consent according to a protocol approved by the local research ethics committee (Faculty of Psychology and Neuroscience, Maastricht University) and received a monetary reimbursement for their participation.

Stimuli and task

Two randomized sets of fifteen face (Tübingen database [Troje and Bülthoff, 1996]; http://faces.kyb.tuebingen.mpg.de) or chair (custom database) images (together with their phase-scrambled counterparts) were presented as stimuli. Images were 5.5 × 5.5° and the depicted objects were rotated inwards (i.e., towards the fixation cross) by 30°. All images had the same mean luminance and RMS contrast (see e.g., Goffaux et al., 2012 for details), and were presented on a grey background at 6° eccentricity from fixation in the lower visual quadrant. Stimulus presentation occurred in 6 s blocks (400 ms per stimulus interleaved with 100 ms blank intervals), alternated with resting baseline periods of 10–12 s. During unilateral stimulation, blocks of faces (F), objects (O; i.e., chairs) or scrambled images (S) were presented in only one of the two visual hemifields at a time (resulting in right hemifield conditions 'xF','xO', 'xS' and conditions 'Fx','Ox', 'Sx' for left hemifield stimulation). In contrast, bilateral stimulation conditions entailed that both hemifields were stimulated simultaneously with each hemifield containing either a series of faces, chairs, or scrambled images (i.e., constituting conditions 'FO' (with Faces in the left and Objects in the right hemifield), 'OF', 'FS', 'SF', 'OS', 'SO', 'FF', 'OO'). Each stimulus configuration was presented twice per run in a pseudo-randomized order (amounting to 28 blocks in total per run), with the constraint that no condition was ever directly repeated. Each participant performed 5 or 6 runs in total, while keeping central fixation throughout. Subjects were instructed to detect a small (1.16 × 1.16°) semi-transparent (16% opaque) red square, which briefly appeared once or twice at random times and positions (within the screen partitions where the main stimuli could appear) per block. To ensure the same distribution of attentional resources across conditions, the square appeared in the left or right hemifield with equal probability (i.e., even during the unilateral stimulation conditions).

Stimulus presentation was controlled using the Presentation software package (v. 18; Neurobehavioral Systems, San Francisco, CA) and images were viewed on a back-projection display via a mirror mounted onto the MR head coil.

Scanning parameters

Imaging data were acquired on a Siemens Magnetom 7T scanner (Siemens, Erlangen, Germany) using a 32-channel head coil (NovaMedical Inc.; Wilmington, MA, USA). Anatomical data were obtained with a T1-weighted magnetization prepared 2 rapid gradient echoes (MP2RAGE; Marques et al., 2010) sequence (240 sagittal slices, Repetition Time [TR]=5000 ms; Echo Time [TE]=2.47 ms; Inversion Time [TI1/I2]=900/2750 ms; Flip Angle [FA1/FA2]=5/3°; Field of View [FoV]=224×203 mm^2; generalized autocalibrating partially parallel acquisitions [GRAPPA]=2; voxel size=0.7 mm^3 isotropic). High-resolution functional images were obtained using 2D gradient echo (T2* weighted) echo-planar imaging (60 transversal slices; TR=2000 ms; TE=21 ms; FA=70°; FoV=192×192 mm^2; GRAPPA=3; multi-band factor=2; voxel size=1.1×1.1×0.8 mm^3). Slices were positioned in parallel to the longitudinal axis of the temporal lobe. To enable the correction of geometric distortions later on, five additional functional volumes with a reversed phase encoding direction (i.e., posterior-

anterior) were acquired at the start of each session.

(f)MRI analyses

All imaging data were analyzed using BrainVoyager QX (v2.8; Brain Innovation, Maastricht, the Netherlands). Anatomical datasets underwent brain extraction, followed by inhomogeneity correction and interpolation to a nominal voxel size of 0.55 mm³ (i.e., a factor 2 smaller than the resolution of the functional data). Preprocessing of the functional datasets followed standard procedures including slice scan time correction, (rigid body) motion correction, linear trend removal, temporal high-pass filtering (up to 2 cycles per run), yet no spatial smoothing. Geometric distortions were corrected by estimating the susceptibility-induced off-resonance field based on a set of 5 functional volumes acquired with reversed phase encoding as described by Andersson et al. (2003). The functional data were finally interpolated to a 1.1 mm³ voxel target resolution and aligned to the corresponding anatomical dataset.

Region of Interest definition

Subject-specific, non-contiguous Regions of Interest (ROIs) within the occipito-temporal cortex were separately determined for the left and right hemisphere by selecting all visually-responsive voxels (i.e., voxels in which the contrast of all conditions combined vs. baseline led to a q(FDR) < 0.01 while applying a cluster threshold of 24 voxels) within a cortical mask defined by the white-grey matter segmented anatomical scans. Early visual areas were localized by acquiring individual meridian maps (by presenting short blocks [8 s each interspersed with 8 s resting baseline] of black-and-white checkerboard wedges flickering at 8 Hz and oriented along the horizontal or vertical meridian in an alternating fashion). A mask encompassing the full extent of the retinotopically activated regions was individually defined for each inflated hemisphere. Subsequently, this mask was converted to volume-space (using a liberal cortex sampling to ensure any retinotopically responsive patches at the mask's fringes were included) and was finally excluded from the main ROI. As a fixed statistical threshold was applied throughout, the number of included voxels could vary across hemispheres and subjects. Nevertheless, the number of included voxels generally matched relatively well (on average 7466 ± 990 [s.e.] and 6775 ± 579 [s.e.] for the left and right hemisphere respectively). The resulting hemisphere-specific ROIs for all subjects are depicted in Fig. S1.

The reasons to a priori focus on the entire collection of visually responsive voxels instead of a number of discrete category-selective ROIs were threefold: First, category-selective single unit responses have been consistently found outside of such discrete functionally-defined patches. Neurophysiological non-human primate studies in which electrode penetrations were not guided by fMRI data have for example repeatedly reported subsets of 15–35% of face-preferring neurons (e.g., Desimone et al., 1984; Kiani et al., 2007). Similarly, a study by Bell et al. (2011) in which fMRI and electrophysiological data from the same monkeys were available showed that recordings outside of the functionally defined category-selective fMRI patches yielded comparable proportions of selective neurons for all examined stimulus categories (i.e., faces, body parts, objects and places). A similar mix of modular and distributed representations has been found in human neuroimaging (e.g., Haxby et al., 2001; Weiner and Grill-Spector, 2010) and intracranial EEG (e.g., Jonas et al., 2016) studies, and is furthermore supported by tracer data (Borra et al., 2010). Second, despite showing the strongest responses to their preferred object category, neurons or voxels within fMRI-defined patches frequently also respond to images from other stimulus classes (albeit to a lesser degree; e.g., Mur et al., 2012). Moreover, even though the large majority of neurons in the center of such fMRI patches portray clear stimulus-preference (e.g., ~90% in Tsao et al., 2006), their proportion

appears to drop towards the patches' more peripheral parts (41–69% in Bell et al., 2011). These findings suggest that a given fMRI voxel in a typical functionally-defined ROI can essentially encompass a number of differently selective subpopulations. In addition, defining functional ROIs is not trivial per se (particularly in unsmoothed high-resolution data) as definition criteria and thresholding choices can have substantive effects (e.g., Weiner and Grill-Spector, 2012). Consequently, attaching binary labels to voxels indicating class selectivity could partially obscure certain effects. Lastly, we were interested in the information present throughout the occipito-temporal cortex as a whole as these regions collectively carry the representations which are potentially available to higher-level downstream areas and it is still largely unknown how modular the organization of this information transfer is.

Decoding analysis

Single-trial (i.e., stimulation block) responses were estimated through a General Linear Model (GLM) fitting the expected hemodynamic response (2-gamma hemodynamic response function [Friston et al., 1998]) to the z-normalized trial data (windowed between −1 to 8 TRs relative to trial onset). The GLM included a constant (reflecting the mean activation level) and an additional confound predictor capturing the linear trend in the BOLD time courses. The resulting estimates (i.e., t-values) across the relevant voxels within the defined ROIs then formed the feature vectors used to train a Support Vector Machine (SVM) classifier (LIBSVM [Chang and Lin, 2011] implementation in BrainVoyager v2.8; linear kernel). SVM training/testing followed a 'leave 1 run out' procedure, iteratively leaving out one run (encompassing two stimulation blocks per condition) for testing while training on the remaining runs. For classification analyses involving training and testing on stimuli in opposite hemifields, both training and testing were performed on all single stimulation block instances collectively. Decoding results were evaluated for statistical significance using the non-parametric Wilcoxon signed-rank test (relative to 50% chance level; one-sided).

Rank order and Overlap analysis

As a descriptive means to capture the response profile characteristics across all included ROI voxels, the beta coefficients for the unilateral stimulation conditions vs. baseline were rank-ordered from the highest to the lowest value for each voxel separately. Each voxel was subsequently assigned to 1 out of 720 possible permutations of the 6 unilateral conditions, and the number of assigned voxels per permutation was counted and normalized with respect to the total number of ROI voxels on a subject-by-subject basis. Lastly, these normalized counts were averaged across subjects, separately for the left and right hemisphere.

To quantify the level of overlap in neural responses to different stimulus configurations and its consistency across subjects, we adopted an approach introduced by Cohen et al. (2014). Here, the beta coefficients for all stimulation conditions vs. baseline were separately rank-ordered (from max to min) for each voxel in the hemispheric ROIs. Next, the number of overlapping voxels were calculated for each possible condition pair and converted to a 'percentage overlap' score, while varying the inclusion threshold from 1% to 99% (in 1% steps). That is, the number of shared voxels across condition pairs was determined across a range of thresholds: from only taking into consideration the 1% most active voxels in both conditions to basically including almost all voxels (leading to an overlap score of close to 100% by definition). Subsequently, the area under the curve (AUC) was calculated to integrate over all considered activation thresholds and the pairwise AUCs were aggregated in an overlap matrix summarizing the activation profile across conditions separately for each hemisphere per subject. To assess the consistency of the resulting overlap scores,

pairwise correlations were calculated across subjects (resulting in 15 pairwise correlation values) using Kendall's tau as implemented in Matlab. Significance testing of the individual pairwise correlations between subject pairs was performed through a comparison with correlations computed based on random permutations of the subjects' overlap matrices (1000 iterations per pair). Finally, the resulting correlations were Fisher r-to-z transformed and tested against a null hypothesis of no correlation (z= 0). The different analysis steps are illustrated in Fig. 3 below.

Combinatory prediction analysis

The gathered data allowed the exploration of different ways to combine the unilateral conditions in order to predict the observed response patterns to the bilateral stimulus configurations. After extraction of the beta values (representing the evoked activation for each condition when separately contrasted against the resting baseline), we applied three combinatorial operations to pairs of unilateral condition responses: summation (SUM), averaging (AVG) or a max (MAX) operation (i.e., simply taking the strongest unilateral response). These synthetic predicted responses were subsequently plotted against the actually observed responses to each of the included bilateral stimulus configurations. Quantification of the match between the predicted and observed responses was achieved by calculating the root mean square error (RMSE) for each condition pair, and lastly pooling these values across all conditions to compare the overall fit produced by the different combination methods. Finally, to facilitate comparisons to earlier findings in the extant literature, the abovementioned averaging approach was extended to quantify the contribution of each hemifield's stimulation on the observed bilateral response pattern by computing a range of weighted averages. More specifically, the responses to the bilateral stimulus configurations were predicted based on an average across the responses to the two unilateral conditions such that the contribution of each hemifield varied between 0% (i.e., not playing a role at all) and 100% (i.e., fully determining the predicted response level). Next, the RMSE between the observed and predicted responses were calculated across all voxels, for each stimulus pair, hemisphere and subject. Lastly, the weighting resulting in the lowest RMSE was identified and averaged across conditions and subjects.

Eye movements and task performance

We recorded eye movements in five subjects (of whom 3 participated in the 7T study) performing our detection task offline, to examine the feasibility of keeping central fixation while covertly attending both peripheral stimulus positions. The stimulus design was identical to the 7T study, except that resting baseline periods were shortened to 1 s. Eye tracking (Eyelink 1000; SR Research, Kanata, Ontario, Canada; 500 Hz sampling) results were analyzed with custom Matlab code computing the condition-specific percentages of eye position samples that were recorded within a circle with a 2° visual angle diameter around the fixation cross. The condition-specific mean fixation times recorded throughout the stimulus presentation blocks were entered into a repeated measures analysis of variance (ANOVA) with 'Stimulus condition' as factor. To study task performance, we performed an ANOVA on reaction time (RT) and accuracy data of the detection task (eye tracking and fMRI experiment) with 'Stimulus condition' as factor. All results were checked for violations of the sphericity assumptions in order to perform a Greenhouse-Geisser correction of the degrees of freedom if required.

Results

Maintenance of fixation and behavioral performance during peripheral detection task

Eye movement analyses indicated that fixation was kept centrally

throughout the task as instructed, as on average (across conditions and subjects) 97.4% (s.e.=0.7) of eye position samples were recorded within a circle with a 2° visual angle diameter around the fixation cross. There were no main effects of Stimulus condition ($F_{(13,52)}=0.59$; $p=0.850$), indicating that fixation was steadily maintained throughout the experiment.

Task performance in the 7T scanner was excellent, with a mean RT of 533.28 ms (s.e.=26.8) and accuracy of 85.84% (s.e. = 8.25). Two subjects underwent the task with a less opaque red rectangle (making the rectangle less visible) resulting in relatively low accuracy scores (mean=60%). Accuracy for the other subjects was close to perfect (mean=98.85%). Importantly, RT ($F_{(13,65)}=0.72$; $p=0.74$) and accuracy ($F_{(13,65)}=1.01$; $p=0.41$) did not differ between conditions. Likewise, in the eye tracking experiment performance was close to perfect, with a mean RT of (536.63 ms; s.e.=19.8) and accuracy of (99.8%; s.e.=0.2) with no differences between conditions (RT: $F_{(13,52)}=0.86$; $p=0.600$; accuracy: $F_{(13,52)}=1.26$; $p=0.238$).

Together, these data show that it was feasible to fulfill the task requirements while keeping central fixation, supporting the conclusions drawn based on the observed pattern of fMRI responses described below, which by themselves would have been unlikely to occur in case subjects would have ignored the central fixation instructions.

Ipsilateral responses to unilateral stimuli are informative

The first analyses focused on the unilateral stimuli presented in isolation (i.e., without a stimulus occurring in the opposite hemifield). Note that the following covers the results based on the aforementioned higher-order visual cortex ROIs, while analogous control analyses on early visual cortex responses are presented in the online supplementary materials. When training and testing were both based on stimuli presented in the *contralateral* visual hemifield, the MVPA decoding reached an accuracy of 89.8% (s.e.=3.35) and 91.7% (s.e.=2.44) for the left and right hemispheric Region of Interest (ROI) respectively (green bars in Fig. 1A). Interestingly, the decoding accuracies following training and testing on stimuli shown in the *ipsilateral* visual hemifield were also high, reaching 70.5% (s.e.=2.84) for the left, and 76.6% (s.e.=2.57) for the right hemispheric ROI (red bars in Fig. 1A). The object category decoding therefore was well above chance irrespective of the visual hemifield or cerebral hemisphere being considered (contra: Z=2.21 (p=0.014) [left hemisphere], Z=2.21 (p=0.014) [right hemisphere]; ipsi: Z=2.21 (p = 0.014) [left hemisphere], Z=2.21 (p=0.014) [right hemisphere]). Thus, also the activation patterns in the hemisphere *ipsilateral* to the site of stimulus presentation carried sufficient information to successfully decode the object category of the presented stimuli, even under *bilateral* attention conditions. The relevance of the latter qualification is supported by the complementary 3T data (acquired under *unilateral* attention conditions; see Fig. S2A): Next to similarly high accuracy levels for the purely contralateral classification conditions (96.6% correct, Z=2.87 (p=0.002) [left hemisphere]; 91.4% correct, Z=2.21 (p=0.014) [right hemisphere]), the decoding results for the ipsilaterally trained and tested stimuli were higher than for the 7T data, namely 87.6% (s.e.=4.43) correct for the left hemisphere (Z=2.82; p=0.003) and 95.0% (s.e.=3.34) for the right hemisphere (Z=2.91; p=0.002). Most likely, the fact that both hemifields needed to be monitored throughout the 7T experiment (i.e., even during the unilateral stimulation conditions) led to the somewhat attenuated decoding accuracies, though still allowing above-chance classification in all cases.

Another way to probe the information content present in the recorded activation patterns was to base the classifier training and testing on stimuli in opposite hemifields (see Fig. 1B). For the left hemisphere 7T ROI, the decoding accuracy was 74.92% (s.e.=4.87) when training on the ipsilaterally presented stimuli and performing the classification on the contralaterally presented stimuli (Z=2.20;
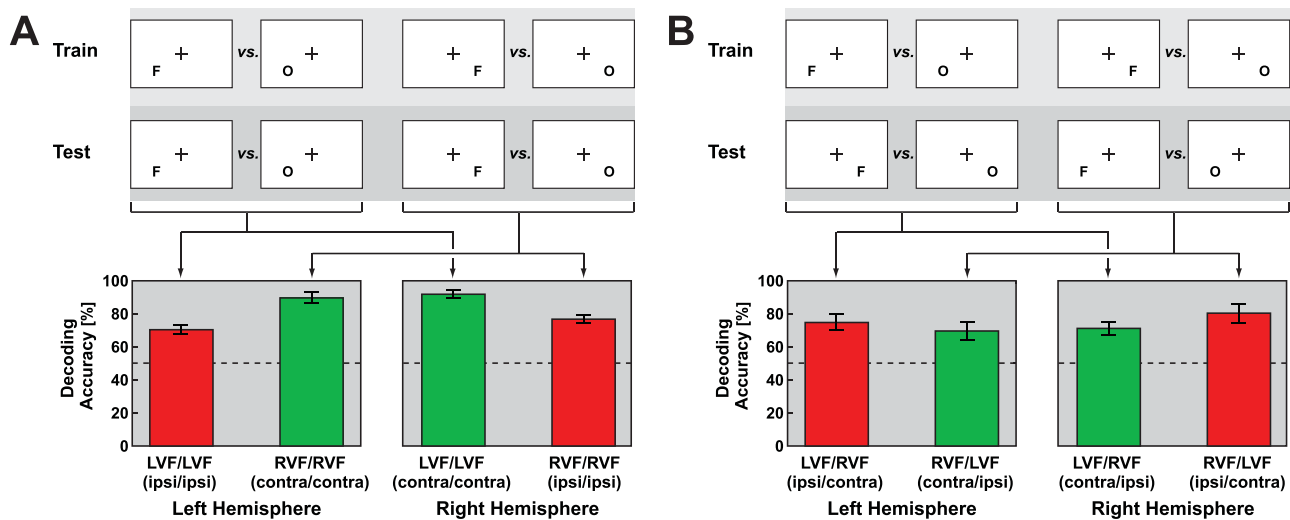
Fig. 1. Decoding results for the unilateral stimulation conditions. An SVM classifier was trained to discriminate Faces (F) from Objects (O) presented in isolation in the lower left or right visual quadrant (i.e., ipsi- or contralateral depending on the cerebral hemisphere under investigation). The upper panels illustrate the different training and test conditions, while the lower panels show the resulting decoding accuracies, with red and green bars referring to training based on ipsilateral and contralateral stimulation conditions, respectively. A) Cases where classifier training and testing were always performed in the same hemifield. B) Cases where classifier training and testing occurred in opposite hemifields. Note that the decoding accuracy was above chance level (indicated by the dotted line) in all cases. Bars represent mean decoding accuracies across subjects; error bars indicate the standard error of the mean (in this and all subsequent figures).

p=0.014). Following the opposite route (i.e., contralateral training and ipsilateral testing), an accuracy of 69.67% (s.e.=5.20) was reached (Z=2.20; p=0.014). Similar results were obtained for the right hemisphere: 80.28% of the trials were correctly classified when moving from the ipsilateral to the contralateral hemifield (s.e.=5.63; Z=2.20; p=0.014), while the accuracy level was 71.03% (s.e.=3.96) when shifting from the contralateral hemifield to the ipsilateral one (Z=2.20; p=0.014). The 3T data corroborate this pattern of results while the overall decoding accuracies again are somewhat higher (see Fig. S2B in the supplementary material, which also includes additional analyses directly contrasting both training/testing schemes). Training on the ipsilaterally presented stimuli while running the classification on the contralateral stimulus conditions led to decoding accuracies of 91.58% (s.e.=3.33; Z=2.84; p=0.002) in the left hemisphere ROI, and 94.13% (s.e.=3.18; Z=2.87; p=0.002) correct in the right hemisphere ROI. With the opposite training vs. testing scheme, accuracies of 80.50% (s.e.=5.76; Z=2.67; p=0.004) and 86.00% (s.e.=3.09; Z=2.812; p=0.003) were achieved for the left and right hemisphere, respectively. Overall, these results suggest the activation patterns in both the contra- and ipsilateral hemisphere relative to the stimulated visual hemifield carry ample information on the semantic category of the presented unilateral stimuli.

Contralateral bias frequently outweighs category-preference

To get more insight into the way in which responses to unilaterally presented stimuli in the ipsi- or contralateral hemifield relate to each other, a ranking (in descending order) of the beta values for all included unilateral conditions was performed throughout the left and right hemispheric ROIs. Based on these rank ordered response profiles, all ROI voxels were subsequently categorized as belonging to 1 out of 720 possible permutations when including all 6 conditions.

The lower panels of Fig. 2 show the number of voxels (expressed as a percentage of all voxels in each hemispheric ROI) for each of the 720 possible response profiles, averaged across participants. When reviewing the distribution of voxel counts across all permutations, it becomes apparent that the majority of voxels (82% [xF+xO+xS] and 83% [Fx +Ox+Sx] on average for the left and right hemisphere) responded most strongly to one of the contralateral stimulation conditions (i.e., had one of these conditions at their top-ranked position). Interestingly, closer

inspection of these subsets of voxels (further illustrated in the upper panels of Fig. 2) revealed a consistent pattern in both hemispheres: a large majority of voxels also placed a contralateral stimulus condition at the 2nd position in their ranking. More specifically, 51.7% of the left hemisphere voxels which responded most vigorously to condition 'xF' had condition 'xO' at the second rank order position, while 41.5% of the voxels with 'xO' as their top-ranked condition had 'xF' being ranked in second position. For the right hemisphere (shown in panel B), the corresponding percentages were 46.5% ('Fx' followed by 'Ox') and 37.3% ('Ox' followed by 'Fx'). In other words, when a voxel maximally responds to a stimulus from a given category presented in the contralateral hemifield, the likelihood of being ranked in second position is 2- to 3- times higher for a stimulus from *another* semantic category in the same *contralateral* hemifield, then for a stimulus from the *same* category presented *ipsilaterally*. These skewed proportions highlight the preponderance of position-related coding in higher-order visual cortex.

An interesting extension of these results relates to the rank orders obtained when zooming into subsets of voxels showing stimulus category preferences. More specifically, when selecting such subsets (through a conjunction contrasting the homogenous bilateral stimulation conditions to baseline and each other, e.g. conj(['FF' > baseline] & ('FF' > 'OO')] for face-preferring voxels at p < 0.01 [uncorr.]), the following pattern emerged: The contralaterally presented preferred stimulus class elicited the strongest response in 92.5% and 84.5% (faces) and 86.1% and 77.3% (objects) of cases in the left and right hemisphere, respectively. Interestingly, 56.1% (left hemisphere) and 53.8% (right hemisphere) of the face preferring voxels now featured the ipsilaterally presented faces at the second rank order position (while the contralateral objects were ranked in second place in 34.1% and 30.5% of these voxels). So here, category preferences do dominate the response profile in about half of the voxels. In contrast, for the object-preferring voxels, the second rank position was most frequently taken by the contralateral scrambled images, namely in 56.8% (left hemisphere) and 68.9% (right hemisphere) of the cases (whereas the ipsilateral objects came in second place in only 11.3% and 10.9% of voxels). These findings suggest that overall, a larger percentage of face-preferring voxels portray a response profile in line with position-tolerant coding (in the sense that ipsilaterally presented faces produce stronger responses than contralaterally presented objects or scrambled
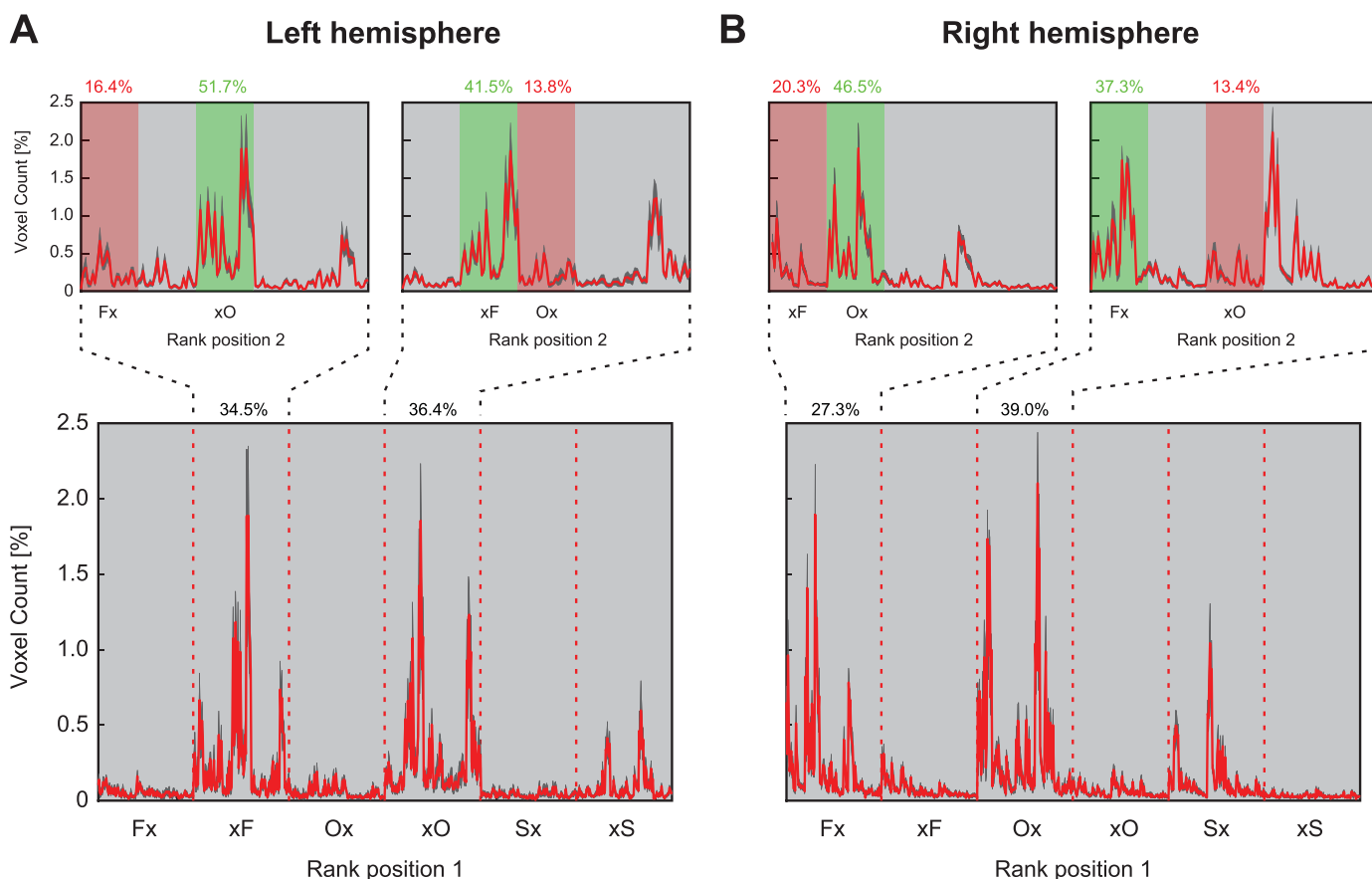
Fig. 2. Rank order results for the unilateral stimulation conditions. Overview of the rank ordering of voxel-wise response amplitudes (as indexed by beta values) for all unilateral conditions. The lower panels show the voxel counts (as percentage of the total nr of ROI voxels) for all possible (i.e., 720) permutations of the 6 included unilateral conditions for both the left and right hemisphere (depicted in panels A and B respectively; shaded regions corresponding to the standard error of the mean across all subjects). The vertical dotted lines represent the transitions between the different subsets of response profiles based on the highest ranking conditions (as indicated on the x-axis). The distributions clearly show a preference for the contralateral hemifield (i.e., xF, xO and xS for the left hemisphere and Fx, Ox, and Sx for the right hemisphere). Notably, when zooming into the subsets in which the first-ranked condition is the contralaterally presented face or object (as depicted in the upper panels), the following pattern is discernable in both hemispheres: the most frequently occurring 2nd-ranked condition always represents the other meaningful stimulus category presented contralaterally. For example, in the left hemisphere, 51.7% of the voxels which responded most strongly to the contralaterally presented Faces (xF) have the contralaterally presented Objects (xO) at rank position 2, whereas ipsilaterally presented faces (Fx) occur at rank position 2 only in 16.4% of voxels.

images in ~55% of the cases) compared to when indiscriminately taking all visually responsive voxels or the subset of object-preferring voxels in a given hemisphere into account.

Response profiles across uni- and bilateral stimulus configurations are highly consistent across subjects

In previous sections, the responses to the different unilateral stimulation conditions were described both in terms of their information content and their rank ordering when taking into account visual stimulation in either the ipsi- or contralateral hemifield. However, next to these unilateral conditions, the experiment also included a number of bilateral stimulus configurations in which stimuli were simultaneously presented (and attended) in both hemifields. Consequently, it was interesting to explore how the response profiles elicited by uni- and bilateral stimulation conditions relate to each other. More specifically, we set out to determine how much overlap in terms of neural response patterns there was between the different uni- and bilateral stimulus conditions and how consistent these patterns were across subjects. To quantify the amount of overlap, an approach introduced by Cohen et al. (2014) was applied (see Methods for details).

The different panels of Fig. 3 illustrate the analysis steps and the generated results. First, the beta values representing the contrasts between each visual stimulation condition versus resting baseline were sorted (from max to min) within each ROI voxel (separately for the left

and right hemisphere in each subject; Fig. 3A). Second, the percentage overlap in maximally responsive voxels was determined for each possible condition pair (i.e., 91 pairs across 14 conditions) per subject, while varying the inclusion threshold from 1 to 99% of maximally responsive voxels (Fig. 3B). In other words, the inclusion threshold was varied from a very conservative (i.e., only including the 1% most responsive voxels for each condition in a pair) to a very liberal (i.e., basically including all ROI voxels) setting, instead of choosing a fixed (but arbitrary) threshold. For each subject and each hemisphere, the overlap between two stimulation conditions was then summarized by integrating across all thresholds and calculating the area under the curve for each condition pair. The resulting pairwise overlap measures were subsequently entered into an 'overlap matrix' where each cell represents one condition pair (Fig. 3C). The range of the pairwise overlap percentages across conditions was highly consistent within our subject sample, with average minima of 58.92% (s.e.=1.53) and 56.26% (s.e.=1.42) and maxima of 82.52% (s.e.=0.80) and 82.04% (s.e.=0.54) for the left and right hemispheric ROI, respectively.

To further assess the consistency of the response profiles across participants, we calculated the correlation (Kendall's tau) between the overlap matrices of each possible subject pair (i.e., 15 pairs). As illustrated in Fig. 3D/F, the resulting pairwise correlations ranged from 0.49 to 0.84 (left hemisphere) and 0.61 to 0.84 (right hemisphere), and all were significant based on permutation testing (1000 iterations; all p's < 0.001). Finally, to go beyond pairwise testing, all
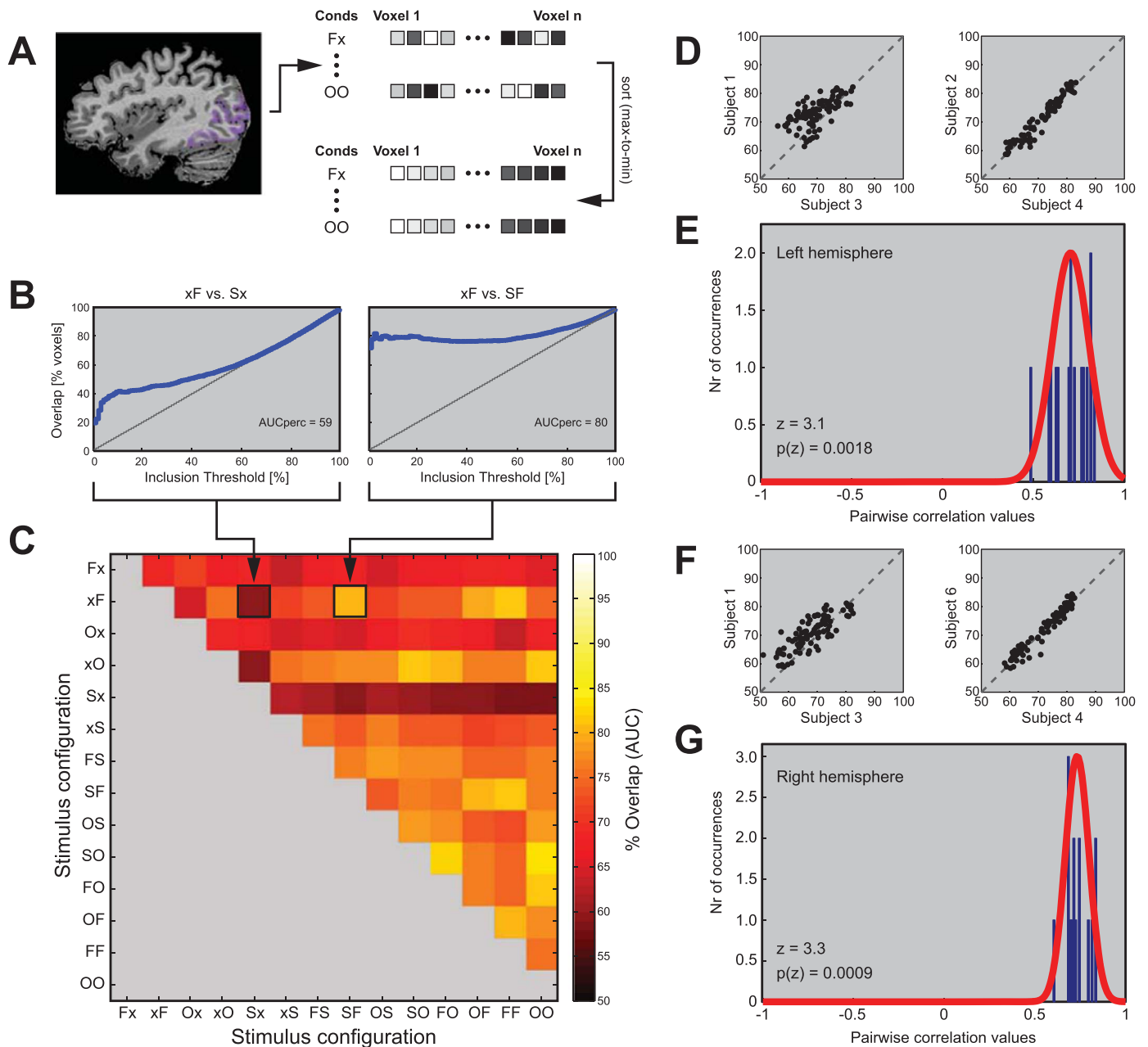
Fig. 3. Processing steps and results of overlap analysis across both uni- and bilateral stimulation conditions. A) For each ROI voxel, the beta values were extracted for all stimulation conditions and rank-ordered from the strongest to the weakest response (for each hemisphere and subject separately). B) For each possible condition pair, the overlap (expressed in % of voxels) was calculated across a range of inclusion thresholds (1-99%; depicted on the x-axis) and the area under the curve was determined as a summary measure (e.g., 59 for 'xF vs. Sx' and 80 for 'xF vs. SF' in the dataset shown here). Note that the grey curve serves as a reference showing the expected overlap when randomly shuffling the rank-ordered voxel lists. C) The overlap measures across all condition pairs were aggregated in an overlap matrix as depicted here for one subject's left hemisphere ROI. To determine how consistent these response patterns were across subjects, pairwise correlations were subsequently performed across the acquired overlap matrices for all subjects. In panel D, the lowest (left) and highest (right) pairwise correlations are shown for the left hemisphere ROIs, yielding correlation values of 0.49 and 0.84 respectively. The distribution of the observed pairwise correlations across all subjects is depicted in panel E. Panels F/G provide the lowest (0.61) and highest (0.84) pairwise correlations and their distribution for the right hemisphere ROI.

correlations across subject pairs were Fisher r-to-z transformed and their average was tested against the null hypothesis of tau equaling zero (Fig. 3E/G). For both hemispheres, the correlation in representational overlap across subjects was highly significant (left hemisphere: z=3.1, p=0.002; right hemisphere: z=3.3, p=0.001). In sum, the way in which response patterns to different visual stimulus configurations (in terms of stimulated hemifield and object category content) correspond to each other is highly consistent across subjects.

Responses to bilateral stimulus configurations are well approximated by a max operation on unilateral stimulation responses

Since all bilateral stimulation conditions constituted different combinations of the included unilateral conditions, it was possible to investigate whether the responses to the unilateral stimuli could be used to predict their counterpart during bilateral stimulation. For each voxel in the hemispheric ROIs, a predicted response was calculated for each bilateral stimulation condition (i.e., 8 conditions in total) based on a combination of responses to its constituent unilateral parts. Three different types of combinations were tested: taking the sum of both
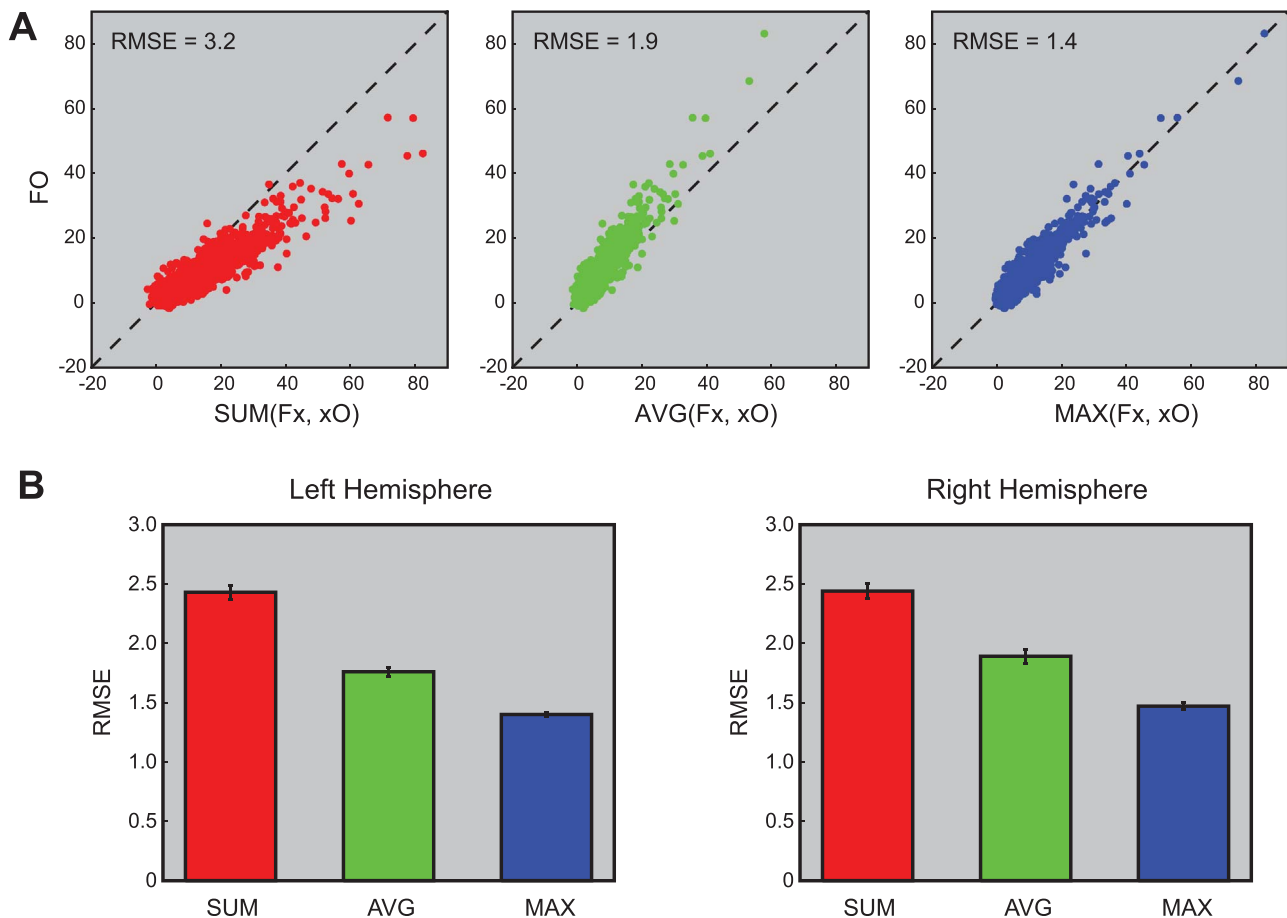
Fig. 4. Predicting the observed responses to bilateral stimulus configurations through various combinations of the unilateral stimulus responses. A) Exemplary results from a single participant where the response (in terms of voxel beta values) to the bilateral 'FO' condition (Faces in the LVF and Objects in the RVF) is predicted by the response to the corresponding unilateral conditions 'Fx' and 'xO' through summation ('sum'), averaging ('avg') or a max operation ('max'). B) RMSE results for the different combination options collapsed across stimulation conditions for both the left and right hemisphere ROI across all subjects.

unilateral building blocks ('sum'), taking their average ('avg'), or using whichever of the responses during unilateral stimulation was highest ('max'). The degree in which each prediction matched the actual data was determined by calculating the root mean square error (RMSE) across all ROI voxels.

Fig. 4A shows the match between the different predictions and the actually observed responses across all voxels in the left hemispheric ROI for a representative participant (the dotted lines represent the unity line, i.e. the line a given data point would fall on in case its predicted value would exactly match its observed value). The plots suggest that summing the unilateral responses leads to an overestimation, averaging to an underestimation, and the max operation resulted in the closest correspondence to the observed bilateral responses. On the group level, for each of the considered cases, the 'max' operation led to the lowest RMSE and therefore predicted the response to the bilateral conditions best in both hemispheres. There was a significant linear trend (from lowest RMSE for 'max' to highest RMSE for 'sum') across the three combination types for all bilateral conditions (all p's < 0.02 for the left and all p's < 0.009 for the right hemisphere respectively). A direct comparison between the 'avg' and 'max' RMSEs was significant for 10 bilateral conditions (p < 0.05), while 2 showed a clear trend (p < 0.07) and 4 were not significant (SF and FF in the left hemisphere; SF and SO in the right hemisphere). Notably though, the 'max' operation was not outperformed once when directly contrasting it to a prediction based on the average of the two unilateral responses. Moreover, when averaging across conditions, the obtained RMSEs were 2.43 (s.e.=0.06), 1.76 (s.e.=0.04) and 1.40 (s.e.=0.02) [left hemisphere], and 2.44 (s.e.=0.06), 1.89 (s.e.=0.06) and 1.47 (s.e.=0.03)

[right hemisphere] for the sum, average, and max prediction respectively (as depicted in Fig. 4B). Pooled as such, the overall linear trend was significant (left hemisphere: p=0.003; right hemisphere: p=0.0004) and the RMSE for the 'max' operation was significantly lower than the 'avg' RMSE (left hemisphere: t(5)=4.2; p=0.009; right hemisphere: t(5)=4.3; p=0.008).

One might argue that the averaging that was applied in the abovementioned analyses is just one (somewhat arbitrarily chosen) instance of a number of ways in which the responses to the unilateral stimuli could be combined through averaging. So next to the fixed 50%-50% hemifield contribution described above, we also explored different weighted average schemes by systematically varying each hemifield's contribution from 0–100% (in steps of 1%) and calculating the RMSEs for each weighting and bilateral stimulation condition. The optimal weighting (i.e., the one resulting in the lowest RMSE) was selected for each stimulus pair, pooled across conditions and averaged across subjects. The obtained outcome indicated that the weighting which most closely approximated the observed responses to the bilateral stimulation conditions was obtained when the contralateral hemifield was allowed to have a contribution of 87% (s.e. = 0.02) for both the left and right hemisphere with corresponding average RMSE values of 1.39 (s.e. = 0.03) and 1.45 (s.e. = 0.04), respectively. The weighting leading to the smallest prediction errors therefore was the one which predominantly incorporated the stimulus in the contralateral hemifield and resulted in RMSEs equivalent to the ones obtained through the max operation reported above.

The same pattern of results was found in the corresponding conditions of the 3T experiment (shown in Fig. S3). Also here, the

prediction based on the 'max' operation came closest to the actual responses evoked by the bilateral stimulation conditions in both hemispheres (all p's < 0.0005 for the linear fit from lowest RMSE for 'max', intermediate RMSE for 'avg' and highest RMSE for 'sum' when considering 'FO' and 'OF' separately; both p's < 0.0002 when conditions are combined). Paired t-tests revealed that the RMSE for the 'max' operation was lower than the 'avg' RMSE in both the left (t(9) =4.1; p=0.03) and right (t(9)=3.02; p=0.014) hemisphere. When additionally exploring the total range of possible weightings between the hemifield contributions, the best approximations were on par with the max operation results in terms of the RMSE. Again, the optimal weighting was heavily skewed towards the stimulus presented in the contralateral hemifield, with contributions of 88% (s.e.=0.6) and 80% (s.e.=6.2) for the left and right hemisphere, respectively.

Discussion

In the current study, we investigated the response profile characteristics of human higher visual cortex to the isolated unilateral or combined bilateral presentation of real-world object stimuli. First, we asked to what extent information in the recorded fMRI activation patterns could be used to decode the object category of the shown stimuli when being presented either contra- or ipsilaterally to the hemisphere under consideration. Second, we explored the different response profiles that emerged when rank-ordering the voxel responses to all unilateral stimulus conditions, with a particular interest in how the top-ranked responses were structured (i.e., based on object category or stimulus position?). Third, we assessed the overlap in responses to all possible condition pairs and quantified the consistency of the revealed patterns across participants. Finally, different strategies to predict the observed responses to the bilateral stimulus configurations based on the responses to the unilateral stimuli were contrasted. Each of these aspects will be discussed in more detail in the different sections below.

Accurate decoding of unilateral hemifield content in both hemispheres

The presented 7T results show that the multi-voxel patterns within the defined hemispheric ROIs could be used to correctly decode the shown unilateral stimulus' object category in 91% of cases when testing and training on *contralateral* responses. When performing the same analyses based on purely *ipsilateral* response patterns, significant decoding accuracies of 74% could be achieved. As a test of equivalence, we also assessed whether it would be possible to train the classifier based on the contralateral stimuli and then perform the testing on the data acquired during the ipsilateral stimulation conditions (and vice versa). Across both hemispheres, decoding performance reached a level of 70% correct in the former case, and 78% in the latter. These significant decoding results show that hemispheric response patterns to either of the included stimulation conditions could thus to a considerable degree function as an informative template for the other.

These findings were corroborated by the 3T dataset: Training and testing on contralateral responses allowed decoding accuracies of 94% correct, while in its ipsilateral counterpart a correct classification on average was achieved 91% of the time. When crossing the training and testing sets, accuracies of 83% (for the case in which the training occurred contralaterally while testing ipsilaterally) and 93% (when changing from the ipsi- to the contralateral hemifield between training and testing) were reached. The higher decoding accuracies (particularly in the cases involving the ipsilateral responses) for the 3T compared to the 7T data most likely reflect a difference in attentional state: in the 3T experiment, the events the subjects were instructed to detect could only occur in the hemifield in which the object stimuli were shown. Therefore, covered attention could be fully deployed to the currently relevant hemifield, presumably allowing more effective information transfer from the contralateral to the ipsilateral hemisphere, leading to

more pronounced activation patterns ipsilaterally. In contrast, the to-be detected events in the 7T experiment could occur in either hemifield (even when the object stimuli were only presented unilaterally), ensuring a consistent divided attentional state across all stimulation conditions.

The presented decoding results are in agreement with recent MVPA decoding studies at 3T based on predefined category-selective ROIs ipsilateral to the presented stimuli (Cichy et al., 2011; 2013). However, when contrasted with our 3T data (which are more comparable in terms of spatial attention demands), the ipsilateral classification accuracies in these studies were somewhat lower than in our case (ranging from ~59% for the Fusiform Body Area to ~82% in area LO [lateral occipital cortex]), while the contralateral bias in terms of decoding performance appeared to be more pronounced (e.g., compare Fig. 2D in Cichy et al., 2013 with current Fig. S2). This discrepancy might imply that the distributed responses to non-preferred stimulus categories (which were taken into account in our analyses) meaningfully complement the processing in discrete category-selective brain areas (see also e.g., Haxby et al., 2001).

Rank-order based response profiles for unilateral stimuli: preponderance of position information

The voxel-wise rank ordering of the responses to the unilateral stimulation conditions revealed that across our subject sample, on average 640 out of the 720 possible permutations of the 6 included stimulus conditions (equalling 89%) occurred at least once in either hemisphere. As shown in Fig. 2, subdividing the observed permutations in terms of the preferred (top-ranked) condition clearly indicated that a large majority of voxels in both hemispheres responded most strongly to one of the contralateral stimulation conditions (adding up to ~82%). In itself, this might not be unexpected. However, the distribution of condition labels for the responses at the next-to highest ranks revealed an interesting pattern: the observed preference for the majority of voxels was 2 to 3 times more likely to be linked to the stimuli's location (i.e., representing a switch in terms of object category compared to the top-ranked condition). That is, a voxel which showed the strongest response to a stimulus belonging to a particular object category (e.g., a face) presented in its contralateral hemifield, was 2 to 3 times more likely to respond more vigorously (or at least at a comparable level) to a *contralaterally* presented stimulus from a *different* object category (e.g. a chair) than to an *ipsilaterally* presented stimulus from the *same* category (e.g., another face). Additional analyses on category-preferring subsets of voxels per hemisphere revealed that position tolerance was more pronounced across face-preferring compared to object-preferring or the broader class of visually-responsive voxels (i.e., responses to ipsilaterally presented faces were ranked in second position more frequently here). However, even across this subset of face-preferring voxels, the tolerance to changes in hemifield by no means was absolute (being apparent in only ~55% of cases). Admittedly, the used voxel selection procedure did not allow to further distinguish between different face-preferring areas (as they were treated collectively here), so future extensions of our work which specifically target such circumscribed regions might lead to varying estimates of the proportion of voxels in which stimulus position trumps object class, depending on the different degrees of face-selectivity of the probed regions.

Overall, the described rank order results are in line with human (Niemeier et al., 2005; Hemond et al., 2007) and monkey (Rajimehr et al., 2014) fMRI studies showing a strong contralateral bias in higher visual cortex for unilateral object stimuli shown in isolation. A similar contralateral bias was also found when examining multi-voxel pattern classification results in category-selective regions (e.g., Cichy et al., 2013). Such an emphasis can be seen as a natural consequence of the underlying anatomy: as visual information is initially processed through segregated neural pathways depending on the hemifield being

stimulated, higher visual areas are (almost) exclusively dependent on the input provided by the intrahemispheric striate cortex. Ablation of unilateral striate cortex has been shown to eliminate IT responses in either hemisphere to visual stimuli in the contralesional hemifield (Rocha-Miranda et al., 1975). Moreover, transection of the callosal splenium and/or the anterior commissure precludes IT responses to stimuli presented in the ipsilateral hemifield (Gross et al., 1974), whereas unilateral inferotemporal lesions cause perceptual deficits in the contralesional hemifield while sparing the ipsilesional hemifield (Merigan and Saunders, 2004). It is this form of 'inheritance' of information across hemispheres (see also e.g., Gazzaniga et al., 1996) which allows any of the MVPA decoding results relying on the ipsilateral hemisphere reported above. The rank order results extend these findings by providing a description of the response profile distribution across all higher-order visually responsive voxels per hemisphere (which was remarkably consistent across subjects) and importantly quantifying the predominance of hemifield over object category preferences. The latter finding could be taken to suggest that labelling voxels as preferring a given visual object category solely based on e.g. foveal stimulation might be misrepresenting the employed neural coding scheme. Together with complementary findings on e.g. category structure for non-preferred object classes in category-preferring regions (Cichy et al., 2013), these results underline where our current understanding of neural processing mechanisms still falls short. In addition, future studies could attempt to quantify how the observed response profiles and the portrayed contralateral bias therein relate to previously proposed parcellation schemes of occipitotemporal cortex in terms of identified field maps (Wandell and Winawer, 2011), eccentricity bias (Malach, Levy, Hasson, 2002), real-world object size (Konkle and Oliva, 2012) or associations to (macro-)anatomical features (Grill-Spector and Weiner, 2014). Particular attention could be paid to assessing differences between within- vs. across-hemifield object pairings in relation to object category preferences. In addition, future research could compare various object categories and see how these relate to the current results which are solely based on chairs as instances of the 'objects' category. This could further minimize the dependency of any of the observed effects on more subtle low-level feature differences across categories (which are close to impossible to fully match across stimuli when using real world object images). Specifically, approaches that incorporate the use of tailored stimuli (based on preferences for different object categories and visual feature conjunctions) into model-driven retinotopic mapping techniques (e.g., Kay et al., 2015; Wandell and Winawer, 2015) will likely prove to be very fruitful in this endeavour.

The overlap in terms of voxel activation patterns across conditions is highly consistent across subjects

In the previous sections, an attempt was made to capture the similarity in neural responses to the different unilateral stimulation conditions by assessing the decodable information content present in multivoxel activation patterns, as well as examining voxelwise rank-order profiles. In order to describe the full breadth of evoked responses while incorporating all presented stimulus configurations, the rank-ordered activation profiles were further exploited to calculate the amount of overlap in activations induced across all possible condition pairs. The range of pairwise overlap percentages across conditions was strikingly similar, varying between ~58% (i.e., slightly above what would be expected by chance for conditions with little in common) and ~82% (for conditions which are more similar) in both hemispheres of each participant. Additionally, the rank correlations across all stimulation conditions when considering all possible subject pairs took average values of ~0.72 for both hemispheres. These findings indicate that the mosaic of overlap estimates that emerges when quantifying how similar the response patterns to each of the different visual stimulation conditions are to each other, is highly reproducible both across

hemispheres and subjects.

Previous work has shown that the degree to which simultaneously presented stimuli in a change detection task could be successfully processed depended on the amount of neural overlap in occipitotemporal response patterns to the included stimulus categories (Cohen et al., 2014). Similarly, response times in a classification task have been shown to correlate with the representational distance (as indexed by fMRI activation patterns) between a given exemplar and the neurally-derived decision boundary separating animate from inanimate objects (Carlson et al., 2014). These findings demonstrate that the applied overlap analysis can tap into meaningful neural representations which are behaviourally relevant. Methods summarizing the level of similarity in neural activation patterns across different experimental conditions are increasingly appreciated as valuable tools to provide an intermediate level description of the underlying neural representations (Kriegeskorte and Kievit, 2013). The results presented here exemplify how such techniques can be used in a straightforward manner to confirm the consistency of evoked response patterns across hemispheres and participants.

Max operation on responses to unilateral stimuli approximates bilateral conditions best

The final question we tackled was the following: How could the acquired fMRI responses to the various unilateral stimulation conditions be optimally combined to best approximate the observed responses to the bilateral stimulation conditions? The presented results suggest that overall a 'max' operation (i.e., using the stronger of the two unilateral responses to predict a voxel's response) led to the prediction which approximated the observed bilateral responses best.

This conclusion is in line with earlier single unit recordings in non-human primates. Sato (1989) reported that the majority of measured units in area TE showed stronger responses to unilateral stimuli presented in the contralateral hemifield, and that the responses to bilateral stimulation configurations were similar in amplitude (i.e., as if the weaker ipsilateral stimulus was largely neglected). This was the case both during a passive fixation and a visual discrimination task. Comparable results were found by Chelazzi et al. (1998) on a memory-guided visual search task in which stimuli were presented on opposite sides of the vertical meridian. In "target-absent" trials in which a preselected "good" stimulus was presented in the contralateral hemifield and a "poor" stimulus was shown ipsilaterally, the population of recorded IT cells responded as if only the contralateral stimulus was shown. Conversely, when this configuration was reversed (i.e., the good stimulus now occurred ipsilaterally while the poor stimulus was placed in the contralateral hemifield), the cell population only responded weakly as if merely reflecting the presence of the poor contralateral stimulus. In either case the predominance of the stimulus in the contralateral hemifield could not be overcome by placing a search target ipsilaterally (i.e., attention could not override the contralateral bias in the 2-stimulus configuration). Moreover, only 6/58 cells responded more vigorously to the good stimulus placed in the contralateral compared to the ipsilateral hemifield when shown in isolation (with 4/58 cells showing the opposite pattern). Contrastingly, half of the cells (29/58) showed a stronger response when the good stimulus appeared in the contralateral hemifield as part of a bilateral stimulus configuration.

At face value, these findings appear to be at odds with more recent neurophysiological and neuroimaging data, pointing to simple (Zoccolan, Cox and DiCarlo, 2005; MacEvoy and Epstein, 2009) or weighted (Agam et al., 2010; Baeck, Wagemans, and Op de Baeck, 2013) averaging as the optimal approximation when trying to predict the response to simultaneously presented objects based on the response to each of the constituent stimuli in isolation. However, in all of these studies stimuli were presented in a vertical arrangement at positions along the vertical meridian, positioned relatively close to the

point of fixation. Estimates of IT receptive field sizes and locations have identified a strong foveal bias (Op de Beeck and Vogels, 2000), suggesting that the reported averaging mainly occurs in situations where both stimuli are bound to fall within the receptive field of the majority of considered neurons (and as such are treated equivalently to within-hemifield interactions). Furthermore, each hemisphere is believed to receive input covering an ipsilateral strip of the visual field surrounding the vertical meridian either via ganglion cells in the temporal retina with crossed projections (Bunt, Minckler, and Johanson, 1977) or via callosal fibers (Payne, 1994), potentially minimizing hemifield-dependent effects which do occur at larger eccentricities. Interestingly, this line of reasoning is consistent with human behavioural data showing that visual processing of multiple objects is easier when they are presented in different hemifields, suggesting both hemispheres can rely on independent attentional resources (e.g., Störmer, Alvarez, Cavanagh, 2014; Alvarez and Cavanagh, 2005). Within-hemifield presentation leads to increased competition, similar to the competitive interactions observed in single unit responses when two stimuli fall within the same receptive field, leading to evoked responses resembling a weighted average of the responses observed when either of the stimuli is shown in isolation (Chelazzi et al., 1998). Of note, 'max'-like responses have been reported in a substantial fraction of macaque V4 cells when explicitly trying to maximize the spatial separation of stimuli shown simultaneously in the recorded neurons' receptive field (Gawne and Martin, 2002). Moreover, applying more lenient selection criteria for inclusion, a wide range of combinatorial responses have been reported in monkey IT (with the degree of clutter-tolerance being inversely related to the recorded units' shape-selectivity [Zoccolan et al., 2007]). Finally, whenever a weighted average has been put forward as the best approximation (Agam et al., 2010; Baeck, Wagemans, and Op de Baeck, 2013), a larger weight was consistently attributed to the stimulus in the pair which by itself evoked the stronger response. This is in line with the weighted average variant presented for our data above, although in our case the contribution of the stronger constituent stimulus was much more pronounced (on average ~87%).

The current results suggest a max (or heavily skewed weighted average) operation is in place when stimulus configurations span both hemifields. Previous findings that vertical stimulus arrangements close to the midline or within one hemifield produce (weighted) average responses might simply represent another part of the same underlying continuum. Future work might attempt to describe this continuum more fully by probing different sets of configurations consisting of identical stimuli within the same subject (e.g. by systematically varying the eccentricity at which stimulus pairs are shown on opposite sides from fixation). By employing such a parametric design, one might be able to capture the wide range of response modulations within a unifying model (e.g. analogous to the contrast normalization at play in V1 population responses to superimposed gratings with varying contrasts [Busse et al., 2009]). Furthermore, although a max operation has been incorporated as a critical processing step in influential artificial object recognition models (e.g., Serre et al., 2007) and might produce more straightforward read-out signals for downstream areas (Cox and Riesenhuber, 2015), it currently remains unclear how the components of a rich visual scene (which appear to be processed in a segregated fashion) are ultimately joined to form a unified percept.

To conclude, the presented 7T and corroborative 3T fMRI data indicate that 1) both contra- and ipsilateral occipitotemporal response patterns contain ample information allowing object category decoding based on unilateral stimulation, even when classifier training and testing are based on visual stimulation in opposite hemifields; 2) rank order-based response profiles across unilateral stimulation conditions reflect a clear contralateral bias and suggest that the portrayed preference for the contralateral hemifield frequently outweighs any object category preference; 3) the overlap in neural representation across all included condition pairs resulted in systematic variations

forming a consistent pattern that could be robustly identified in all subjects; and 4) the responses to the bilateral stimulus configurations could be approximated well by a max operation on the responses to the constituent elements when shown in isolation.

In sum, these findings further strengthen the claim that position (and not only object identity) information is coded along the ventral visual processing stream (Kravitz et al., 2013). Additionally, the presented results highlight that stimulus position frequently overrules object category preferences, and as such provide important constraints for ongoing computational modelling efforts.

Acknowledgments

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neuroimage.2017.03.038.

References

Agam, Y., Liu, H., Papanastassiou, A., Buia, C., Golby, A.J., Madsen, J.R., Kreiman, G., 2010. Robust selectivity to two-object images in human visual cortex. Curr. Biol. 20, 872–879.
Alvarez, G.A., Cavanagh, P., 2005. Independent resources for attentional tracking in the left and right visual hemifields. Psychol. Sci. 16, 637–643.
Andersson, J.L.R., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. NeuroImage 20, 870–888.
Baeck, A., Wagemans, J., Op de Beeck, H., 2013. The distributed representation of random and meaningful object pairs in human occipitotemporal cortex: the weighted average as a general rule. NeuroImage 70, 37–47.
Bell, A.H., Malecek, N.J., Morin, E.L., Hadj-Bouziane, F., Tootell, R.B., Ungerleider, L.G., 2011. Relationship between functional magnetic resonance imaging-identified regions and neuronal category selectivity. J. Neurosci. 31, 12229–12240.
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.
Bunt, A.H., Minckler, D.S., Johanson, G.W., 1977. Demonstration of bilateral projection of the central retina of the monkey with horseradish peroxidase neuronography. J. Comp. Neurol. 171, 619–630.
Busse, L., Wade, A.R., Carandini, M., 2009. Representation of Concurrent Stimuli by Population Activity in Visual Cortex. Neuron 64, 931–942.
Carlson, T., Hogendoorn, H., Fonteijn, H., Verstraten, F.A.J., 2011. Spatial coding and invariance in object-selective cortex. Cortex 47, 14–22.
Carlson, T.A., Ritchie, J.B., Kriegeskorte, N., Durvasula, S., Ma, J., 2014. RT for object categorization is predicted by representational distance. J. Cogn. Neurosci. 26, 132–142.
Chang, C.C., Lin, C.J., 2011. LIBSVM: a library for support vector machines. ACM TIST 2 (27), 1–27, (27).
Chelazzi, L., Duncan, J., Miller, E.K., Desimone, R., 1998. Responses of neurons in inferior temporal cortex during memory-guided visual search. J. Neurophysiol. 80, 2918–2940.
Cichy, R.M., Chen, Y., Haynes, J.-D., 2011. Encoding the identity and location of objects in human LOC. Neuroimage. 54, 2297–2307.
Cichy, R.M., Sterzer, P., Heinzle, J., Elliott, L.T., Ramirez, F., Haynes, J.-D., 2013. Probing principles of large-scale object representation: category preference and location encoding. Hum. Brain Map. 34, 1636–1651.
Cohen, M.A., Konkle, T., Rhee, J.Y., Nakayama, K., Alvarez, G.A., 2014. Processing multiple visual objects is limited by overlap in neural channels. Proc. Natl. Acad. Sci. USA 111, 8955–8960.
Cox, P.H., Riesenhuber, M., 2015. There Is a "U" in Clutter: evidence for Robust Sparse Codes Underlying Clutter Tolerance in Human Vision. J. Neurosci. 35, 14148–14159.
Desimone, R., Albright, T.D., Gross, C.G., Bruce, C., 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. J. Neurosci. 4, 2051–2062.
DiCarlo, J.J., Maunsell, J.H.R., 2003. Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. J. Neurophysiol. 89, 3264–3278.
DiCarlo, J.J., Zoccolan, D., Rust, N.C., 2012. How does the brain solve visual object

recognition? Neuron 73, 415–434.

Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., Turner, R., 1998. Event-related fMRI: characterizing differential responses. NeuroImage 7, 30–40.

Gawne, T.J., Martin, J.M., 2002. Responses of primate visual cortical V4 neurons to simultaneously presented stimuli. J. Neurophysiol. 88, 1128–1135.

Gazzaniga, M.S., Eliassen, J.C., Nisenson, L., Wessinger, C.M., Fendrich, R., Baynes, K., 1996. Collaboration between the hemispheres of a callosotomy patient. Brain 119, 1255–1262.

Goebel, R., 2012. Position coding in the visual word form area. Proc. Natl. Acad. Sci. USA 109, 9226–9227.

Goffaux, V., Peters, J., Haubrechts, J., Schiltz, C., Jansma, B., Goebel, R., 2012. From coarse to fine? Spatial and temporal dynamics of cortical face processing. Cereb. Cortex 21, 467–476.

Grill-Spector, K., Weiner, K.S., 2014. The functional architecture of the ventral temporal cortex and its role in categorization. Nat. Rev. Neurosci. 15, 536–548.

Gross, C.G., Bender, D.B., Rocha-Miranda, C.E., 1969. Visual receptive fields of neurons in inferotemporal cortex of the monkey. Science 166, 1303–1306.

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.

Hemond, C.C., Kanwisher, N.G., Op de Beeck, H.P., 2007. A preference for contralateral stimuli in human object- and face-selective cortex. PLoS One, e574. http://dx.doi.org/10.1371/journal.pone.0000574.

Ito, M., Tamura, H., Fujita, I., Tanaka, K., 1995. Size and position invariance of neuronal responses in monkey inferotemporal cortex. J. Neurophysiol. 73, 218–226.

Jonas, J., Jacques, C., Liu-Shuang, J., Brissart, H., Colnat-Coulbois, S., Maillard, L., Rossion, B., 2016. A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. Proc. Natl. Acad. Sci. USA 113, E4088–E4097.

Kay, K., Weiner, K.S., Grill-Spector, K., 2015. Attention reduces spatial uncertainty in human ventral temporal cortex. Curr. Biol. 25, 595–600.

Kiani, R., Esteky, H., Mirpour, K., Tanaka, K., 2007. Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. J. Neurophys. 97, 4296–4309.

Konkle, T., Oliva, A., 2012. A real-world size organization of object responses in occipitotemporal cortex. Neuron 74, 1114–1124.

Kravitz, D.J., Vinson, L.D., Baker, C.I., 2008. How position dependent is visual object recognition? Trends Cogn. Sci. 12, 114–122.

Kravitz, D.J., Kriegeskorte, N., Baker, C.I., 2010. High-level visual object representations are constrained by position. Cereb. Cortex 20, 2916–2925.

Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G., Mishkin, M., 2013. The ventral visual pathway: an expanded neural framework for the processing of object quality. Trends Cogn. Sci. 17, 26–49.

Kriegeskorte, N., Kievit, R.A., 2013. Representational geometry: integrating cognition, computation, and the brain. Trends Cogn. Sci. 17, 401412.

MacEvoy, S.P., Epstein, R.A., 2009. Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. Curr. Biol. 19, 943–947.

MacEvoy, S.P., Yang, Z., 2012. Joint neuronal tuning for object form and position in the human lateral occipital complex. Neuroimage. 63, 1901–1908.

Malach, R., Levy, I., Hasson, U., 2002. The topography of high-order human object areas. Trends Cogn. Sci. 6 (4), 176–184.

Marques, J.P., Kober, T., Krueger, G., van der Zwaag, W., van de Moortele, P.F., Gruetter, R., 2010. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. NeuroImage 49, 1271–1281.

Merigan, W.H., Saunders, R.C., 2004. Unilateral deficits in visual perception and learning after unilateral inferotemporal cortex lesions in macaques. Cereb. Cortex 14, 863–871.

Mur, M., Ruff, D.A., Bodurka, J., De Weerd, P., Bandettini, P.A., Kriegeskorte, N., 2012.

Categorical, yet graded - Single-image activation profiles of human category-selective cortical regions. J. Neurosci. 32 (25), 8649–8662.

Niemeier, M., Goltz, H.C., Kuchinad, A., Tweed, D.B., Vilis, T., 2005. A contralateral preference in the lateral occipital area: sensory and attentional mechanisms. Cereb. Cortex 15, 325–331.

Op de Beeck, H., Vogels, R, 2000. Spatial sensitivity of macaque inferior temporal neurons. J. Comp. Neurol. 426, 505–518.

Payne, B.R., 1994. Neuronal interactions in cat visual cortex mediated by the corpus callosum. Behav. Brain Res. 64, 55–64.

Rajimehr, R., Bilenko, N., Vanduffel, W., Tootell, R.B.H., 2014. Retinotopy Versus Face Selectivity in Macaque Visual Cortex. J. Cogn. Neurosci. 26, 2691–2700.

Rauschecker, A.M., Bowen, R.F., Parvizi, J., Wandell, B.A., 2012. Position sensitivity in the visual word form area. Proc. Natl. Acad. Sci. USA 109, E1568–E1577.

Rocha-Miranda, C.E., Bender, D.B., Gross, C.G., Mishkin, M., 1975. Visual activation of neurons in inferotemporal cortex depends on striate cortex and forebrain commissures. J. Neurophysiol. 38, 475–491.

Rolls, E.T., Tovee, M.J., 1995. The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive field. Exp. Brain Res. 103, 409–420.

Rolls, E.T., Aggelopoulos, N.C., Zheng, F., 2003. The receptive field of inferior temporal cortex neurons in natural scenes. J. Neurosci. 23, 339–348.

Sary, G., Vogels, R., Orban, G.A., 1993. Cue-invariant shape selectivity of macaque inferior temporal neurons. Science 260, 995–997.

Sato, T., 1989. Interactions of visual stimuli in the receptive fields of inferior temporal neurons in awake macaques. Exp. Brain Res. 77, 23–30.

Schwartz, E.L., Desimone, R., Albright, T.D., Gross, C.G., 1983. Shape recognition and inferior temporal neurons. Proc. Natl. Acad. Sci. USA 80, 5776–5778.

Schwarzlose, R.F., Swisher, J.D., Dang, S., Kanwisher, N., 2008. The distribution of category and location information across object-selective regions in human visual cortex. Proc. Natl. Acad. Sci. USA 105, 4447–4452.

Serre, T., Oliva, A., Poggio, T., 2007. A feedforward architecture accounts for rapid categorization. Proc. Natl. Acad. Sci. USA 104, 6424–6429.

Störmer, V.S., Alvarez, G.A., Cavanagh, P., 2014. Within-hemifield competition in early visual areas limits the ability to track multiple objects with attention. J. Neurosci. 34, 11526–11533.

Tanaka, K., Saito, H., Fukada, Y., Moriya, M., 1991. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. J. Neurophysiol. 66, 170–189.

Tsao, D.Y., Freiwald, W.A., Tootell, R.B., Livingstone, M.S., 2006. A cortical region consisting entirely of face-selective cells. Science 311, 670–674.

Tovee, M.J., Rolls, E.T., Azzopardi, P., 1994. Translation invariance in the responses to faces of single neurons in the temporal visual cortical areas of the alert macaque. J. Neurophysiol. 72, 1049–1060.

Troje, N., Bülthoff, N.N., 1996. Face recognition under varying poses: the role of texture and shape. Vision Res. 36, 1761–1771.

Wandell, B.A., Winawer, J., 2011. Imaging retinotopic maps in the human brain. Vision Res. 51, 718–737.

Wandell, B.A., Winawer, J., 2015. Computational neuroimaging and population receptive fields. Trends Cogn. Sci. 19, 349–357.

Weiner, K.S., Grill-Spector, K., 2010. Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. NeuroImage 52, 1559–1573.

Weiner, K.S., Grill-Spector, K., 2012. The improbable simplicity of the fusiform face area. Trends Cogn. Sci. 16, 251–254.

Zoccolan, D., Cox, D.D., DiCarlo, J.J., 2005. Multiple object response normalization in monkey inferotemporal cortex. J. Neurosci. 25, 8150–8164.

Zoccolan, D., Kouh, M., Poggio, T., DiCarlo, J.J., 2007. Trade-off between object selectivity and tolerance in monkey inferotemporal cortex. J. Neurosci. 27, 12292–12307.