Imagery and Perception Share Cortical Representations of Content and Location

Radoslaw M. Cichy[1,2], Jakob Heinzle[1] and John-Dylan Haynes[1,2,3]

[1]Bernstein Center for Computational Neuroscience Berlin, Charité—Universitätsmedizin Berlin, 10115 Berlin, Germany, [2]Berlin School of Mind and Brain, Humboldt Universität zu Berlin, 10099 Berlin, Germany and [3]Max Planck Institute for Human Cognitive and Brain Sciences, 04303 Leipzig, Germany

Address correspondence to Radoslaw Martin Cichy, Charité—Universitätsmedizin Berlin, Bernstein Center for Computational Neuroscience, Haus 6, Philippstrasse 13, 10115 Berlin, Germany. Email: rmcichy@gmail.com.

Visual imagery allows us to vividly imagine scenes in the absence of visual stimulation. The likeness of visual imagery to visual perception suggests that they might share neural mechanisms in the brain. Here, we directly investigated whether perception and visual imagery share cortical representations. Specifically, we used a combination of functional magnetic resonance imaging (fMRI) and multivariate pattern classification to assess whether imagery and perception encode the "category" of objects and their "location" in a similar fashion. Our results indicate that the fMRI response patterns for different categories of imagined objects can be used to predict the fMRI response patters for seen objects. Similarly, we found a shared representation of location in low-level and high-level ventral visual cortex. Thus, our results support the view that imagery and perception are based on similar neural representations.

Keywords: high-level vision, imagery, multivoxel pattern classification

Introduction

Visual mental imagery refers to our ability to conjure up a visual experience in the absence of retinal stimulation. This experience is phenomenologically similar to the experience of seeing and thus often related to as "seeing with the mind's eye" (Tye 2000; Kosslyn et al. 2001; Pylyshyn 2002). If perception and imagery can yield similar experiences, it could be assumed that this is because they share similar representations in the brain (Kosslyn et al. 1997). We investigated this question directly using functional magnetic resonance imaging (fMRI) and multivariate pattern classification (Haxby et al. 2001; Spiridon and Kanwisher 2002; Cox and Savoy 2003; Haynes and Rees 2005a, 2005b; Kamitani and Tong 2005; Kriegeskorte and Bandettini 2007; Williams et al. 2008). Specifically, we asked 2 questions. First, do imagery and perception share representations of the "content," that is, of the category of object a person was seeing? Second, do imagery and perception share representations of the "location," that is, where an object is seen to be?

Prior fMRI studies based on blood oxygen level–dependent (BOLD) activation levels already suggested that imagery and perception share representations of content in high-level ventral visual cortex (O'Craven and Kanwisher 2000; Ishai et al. 2000). Category-selective regions, that is, specific specialized regions in high-level ventral visual cortex that during perception activate more for objects of a particular category, than to any other category (Grill-Spector and Malach 2004; Reddy and Kanwisher 2007; Op de Beeck et al. 2008), showed similar activations during imagery as during perception. However, similar BOLD activation in a category-selective region cannot conclusively establish whether imagery and perception evoke the same representations. In contrast, multivariate pattern classification can establish the encoding of specific contents (Mika et al. 2001; Haxby et al. 2001; Spiridon and Kanwisher 2002; Cox and Savoy 2003; Carlson et al. 2003; Kamitani and Tong 2005; Haynes and Rees 2005a, 2006; Norman et al. 2006; Haynes et al. 2007). Recent studies investigated directly whether imagery and perception share representations of simple shapes (Stokes et al. 2009) and object categories (Reddy et al. 2010) in lateral occipital complex (LOC) (Malach et al. 1995; Grill-Spector and Malach 2004). Here, we aimed to extend this previous work and investigate the link between the category of imagined objects and category-selective regions, such as the fusiform face area (FFA) and occipital face area (OFA), (Puce et al. 1995; Clark et al. 1996; Kanwisher et al. 1997; Halgren et al. 1999; Gauthier et al. 2000; Haxby et al. 2000), the fusiform body area (FBA) and extrastriate body area (EBA) (Downing et al. 2001; Peelen and Downing 2005; Schwarzlose et al. 2005), the parahippocampal place area (PPA) and transverse occipital sulcus (TOS) (Aguirre et al. 1998; Epstein and Kanwisher 1998; Hasson et al. 2003).

Furthermore, we also wanted to assess which regions share encoding of the location of an object between perception and imagery. Prior psychophysical (Farah 1985, 1989; Craver-Lemley and Reeves 1992; Ishai and Sagi 1995) and neuro-imaging studies (Kosslyn et al. 1995; Tootell 1998; Klein et al. 2004; Slotnick et al. 2005; Thirion et al. 2006) indicated that imagery engages low-level visual cortex in a topographical manner comparable to perception. In contrast, the role of high-level ventral visual cortex in the representation of location during imagery remains unknown. Recent studies of object perception indicate that object location is represented in ventral visual cortex beyond low-level visual cortex (MacEvoy and Epstein 2007; Schwarzlose et al. 2008; Sayres and Grill-Spector 2008; Carlson et al. 2009). Thus, here we asked whether perception and imagery share representations of location in low- and high-level ventral visual cortex.

Materials and Methods

Participants and Experimental Design

Sixteen volunteers (age 22–33 years) gave written informed consent to participate in the experiment. The experiment was approved by the ethics committee of the Max Planck Institute of Human Cognitive and Brain Sciences (Leipzig) and conducted according to the Declaration of Helsinki. All participants were right handed and had normal or corrected-to-normal vision.

Each subject completed 5 runs (duration 642 s per run) of the main experiment. During each run, participants either viewed (perception condition) or were instructed to imagine pictures (imagery condition)

of 3 different object exemplars in 4 different categories (Fig. 1A). In the perception condition, the pictures (size 4.8°) were presented for 4 s at a position either 6° left or right of fixation (Fig. 1B,C) in pseudorandom order. In the imagery condition, participants received auditory cues that indicated the location at which to imagine an object (left or right of fixation) and which of the objects to imagine (Fig. 1B,C). Participants had 4 s to conjure up the indicated image as similar as possible in position and appearance to the visually presented pictures. Stimulus presentations and imagery were interleaved with randomly jittered interstimulus intervals of 2 to 6-s duration during which a gray background screen was shown. During stimulus presentation and imagery, participants were instructed to fixate a central white square. Between-stimulus perception or imagery participants were engaged in a Landolt-C task on fixation to assure that participants' attention was directed to fixation and to prevent them from being involved in further unwanted imagery (Bieg et al. 2010). In detail, the fixation square turned red and opened either left or right every 1000 ms (open 800 ms, closed 200 ms). Participants pressed a button indicating the direction of the opening of the Landolt-C.

In a practice session before the scan, participants completed 1 to 3 modified runs of the main experiment to learn the fixation task and to familiarize themselves with the experiment. The modified runs were identical in setup to the main experiment with the exception of the substitution of the gray background with a random-dot display of black and white dots. The display inverted pixel luminance in every frame. When participants held fixation as required, the alternating displays gave the impression of a uniform gray background. In contrast, eye movements led to a striking experience of a flash (Guzman-Martinez et al. 2009). Participants were asked to maintain fixation to avoid the experience of the flash. All participants reported noticing the effect and using it as feedback to improve fixation.

After the main fMRI experiment, each participant performed 5 localizer runs (duration 180 s each) to identify category-selective regions for bodies, places, and houses, as well as high-level cortex that responds more to pictures of objects than to scrambled counterparts without a clear category preference. Participants viewed blocks of images from 5 different stimulus classes (bodies, faces, scenes, houses, everyday objects, and grid-scrambled objects). In each run, 2 blocks of images from each of the 5 different stimulus classes were shown, that is, in total 10 image blocks. Blocks of images were interleaved with 16-s periods of a uniform black background. Each image block had a duration of 16 s and consisted of 20 images (presentation time 600 ms, 200 ms gap). To preclude a foveal bias in the cortical activation, the same picture was presented simultaneously at 3 adjacent positions along the horizontal meridian. Participants were asked to maintain fixation on a central fixation dot. Participants performed a one-back task on repetitions of an image in order to sustain attention to the images. In each block, at random 4 of the 20 images were repeated. Participants indicated their answer via a button press. The serial order of conditions was counterbalanced within participants.

Finally, 10 out of the 16 participants participated in a retinotopic mapping session to identify low-level visual regions V1, V2, and V3 using the standard travelling wave method with a double wedge and expanding ring stimuli (Sereno et al. 1995; DeYoe et al. 1996; Wandell et al. 2007). Participants completed 3–4 runs of angular mapping to map the boarders between visual regions and 2–3 runs of eccentricity mapping. Gray matter segmentation of anatomical images was conducted using FreeSurfer (Dale et al. 1999), and mrGray was used for cortical flattening (Wandell et al. 2000). We defined the borders between ventral and dorsal areas V1, V2, and V3 by visual inspection of flattened polar angle maps. Eccentricity maps were used to check whether anterior limits of the regions as defined by polar angle maps corresponded with eccentricity mapping. Borders of areas V1, V2, and V3 could be reliably defined in all 10 participants bilaterally. The quality of eccentricity and polar angle mapping did not allow reliable identification of the borders between areas hV4 or V3A/B in all subjects, so that we did not define retinotopic areas beyond area V3. The regions of interest (ROIs) for visual areas V1 to V3 were based on the borders on the flattened surface and transformed back into the functional space of the main experiment. Ventral and dorsal maps of areas V1 to V3 were combined into common ROIs, respectively.
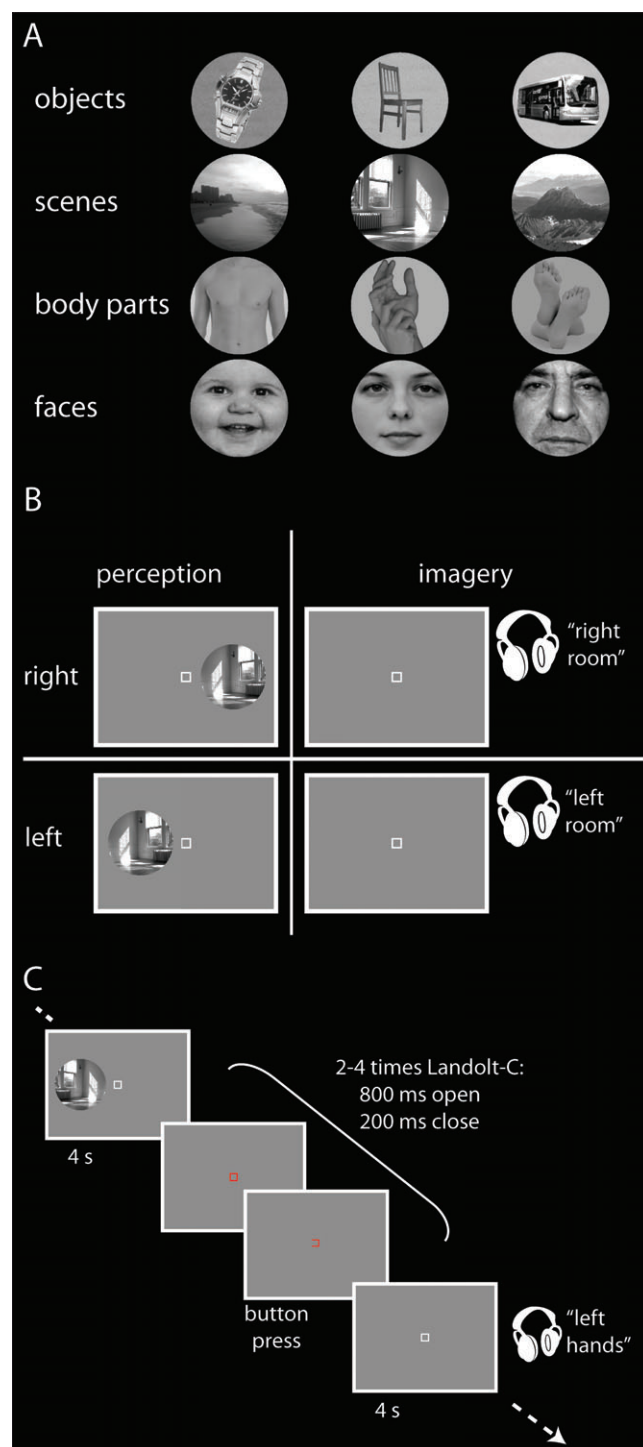
Figure 1. Stimulus set and experimental design. (A) The stimulus set comprised pictures of 3 object exemplars in 4 categories: objects, scenes, body parts, and faces. (B) Schematic drawing of the experimental conditions. Stimuli were either presented visually, or participants were verbally instructed to imagine stimuli either left or right of a central fixation square. Participants were asked to imagine objects as similar as possible in position and appearance to the visually presented pictures. (C) Periods of stimulus presentation and stimulus imagery were interleaved with periods of a gray background screen with a variable duration of 2–6 s. During the whole experiment, participants were instructed to fixate a central white square. At the end of a stimulus presentation or imagery, the central fixation square turned red. Between stimulus presentations, participants completed a Landolt-C task.

fMRI Acquisition

A 3T Trio scanner (Siemens, Erlangen, Germany) with a 12-channel head coil was used to acquire MRI data. Structural images were acquired with a T_1-weighted sequence (192 sagittal slices, field of view [FOV] = 256 mm^2, time repetition [TR] = 1900 ms, time echo [TE] = 2.52 ms, flip angle = 9°). For the main experiment, 5 runs of 321 volumes were acquired for each participant (gradient-echo echo-planar imaging [EPI] sequence: TR = 2000 ms, TE = 30 ms, flip angle = 70°, FOV = 256 × 192 mm^2, FOV phase = 75%, matrix = 128 × 96, ascending acquisition, gap = 10%, resolution = 2 mm isotropic, slices = 24). Slices were positioned parallel to the temporal lobe, such that the fMRI volume covered the ventral visual regions from low-level visual to anterior temporal cortex. For the 5 localizer scans, consisting of 90 volumes each, the parameters were identical. For the retinotopic mapping, 6 to 8 runs of 160 volumes were acquired for each participant (gradient-echo EPI sequence: TR = 1500 ms, TE = 30 ms, flip angle = 90°, FOV = 256 × 192 mm^2, matrix = 128 × 86, ascending acquisition, gap = 50%, resolution = 2 mm isotropic, slices = 25). The slices were positioned parallel to the calcarine sulcus.

fMRI Analysis

All functional data were initially processed using SPM2 (http://http://www.fil.ion.ucl.ac.uk/spm). Data were realigned and slice-time corrected. In the following, we will first describe the analysis of the functional localizers that served the definition of ROIs. We then explain the analysis of the main experiment.

Localizers and Definition of Regions of Interest

First, we modeled the fMRI response in the "independent" localizer runs to identify category-selective regions and object-selective regions. Functional data of the localizer runs were spatially smoothed with a 4-mm FWHM Gaussian kernel. The data were modeled with a general linear model (GLM) that included the 5 stimulus classes as conditions (faces, places, bodies, objects and scrambled objects). Next, we identified voxels that showed category preference by contrasting parameter estimates evoked by the specific category in question with parameter estimates evoked by objects. In this manner, face-selective (T-contrast faces > objects), body-selective (T-contrast bodies > objects and place-selective (T-contrast places > objects) voxels were defined. Similarly, we identified voxels activated more by pictures of objects than by their scrambled counterparts (T-contrast objects > scrambled objects). Next, we defined regions of interest (ROIs) in a multistep process. First, we identified the most activated voxel in each contrast (thresholded at $P < 0.0001$, uncorrected) in lateral–occipital and ventral–temporal positions on the left and right hemisphere of the cortical surface, respectively. Then, we defined a sphere with a 7-voxel radius around this peak voxel. This step limited further voxel selection by vicinity to the most activated voxel and by anatomical location. Finally, within this sphere, we selected only the 300 most activated voxels in each contrast. This yielded up to 12 category-selective ROIs in each subject: the FFA and OFA for faces (Puce et al. 1995; Clark et al. 1996; Kanwisher et al. 1997; Halgren et al. 1999; Gauthier et al. 2000; Haxby et al. 2000), the FBA and EBA for bodies (Downing et al. 2001; Peelen and Downing 2005; Schwarzlose et al. 2005), and the PPA and TOS for places and scenes (Epstein and Kanwisher 1998; Aguirre et al. 1998; Hasson et al. 2003). All areas were defined in the right and the left hemisphere, respectively. In addition, up to 4 object-selective ROIs were identified in the same fashion: the fusiform gyrus (FUS) and lateral–occipital activation (LO) (Malach et al. 1995; Grill-Spector and Malach 2004; Eger et al. 2008a,b) in the right and the left hemisphere, respectively. Note that not every ROI was present in each hemisphere in all participants (Supplementary Table 1). In total, we identified the following numbers of ROIs: FFA (28), OFA (24), PPA (32), TOS (24), FBA (20), EBA (32), FUS (32), and LO (32). Importantly, our ROI identification procedure takes into account individual differences in the location of category-selective regions and guarantees equality of ROI size across ROIs and subjects. Finally, we selected voxels in low-level visual regions V1, V2, and V3 as defined by retinotopic mapping. For this, we calculated a T-contrast all classes of visual stimulation > baseline and chose the 300 most activated voxels in

V1, V2, and V3 each in the left and right hemisphere, respectively. As each lower-level visual region could be defined in each participant, we identified 20 ROIs for V1, V2, and V3 each.

Analysis of Main Experiment

Participants successfully performed the Landolt-C fixation task (mean ± standard error of the mean = 88.43 ± 3.95% correct). Thus, no fMRI data were rejected from further analysis. We modeled the cortical response to the experimental conditions in the main experiment for each run separately. For this, we treated all exemplars belonging to the same category as the same condition. This resulted in a 4 (categories) × 2 (locations) × 2 (perception vs. imagery) design. The onsets and durations of the stimulus presentations were entered into a GLM as regressors and convolved with a hemodynamic response function. The estimation of this model yielded 16 parameter estimates per run, representing the responsiveness of each voxel to the 4 different object categories at either of the 2 different locations in either the perception or the imagery condition.

Pattern Classification

Data from the main experiment were subjected to 2 multivoxel pattern classification analyses (Muller et al. 2001; Haynes and Rees 2006; Kriegeskorte et al. 2006; Norman et al. 2006) using a linear support vector classifier (SVC) with a fixed regularization parameter $C = 1$ in the LibSVM implementation (http://http://www.csie.ntu.edu.tw/~cjlin/libsvm). The 2 analyses investigated whether imagery and perception share representations of 1) object category and 2) object location. Each analysis shared a basic framework that was adapted. Analyses were conducted independently for each ROI and for each subject. Please recall that ROIs were defined based on the independent localizer runs. For each run, we extracted parameter estimates for the experimental conditions under investigation (see below). These parameter estimates constituted the pattern vectors (length of 300, corresponding to 300 voxels) that entered the pattern classification. Pattern vectors from 4 out of 5 runs were assigned to a training data set, which was used to train the SVC. The trained SVC was used to classify pattern vectors from the independent test data set consisting of the fifth run.

Attribution of pattern vectors to training and test sets was based on the following reasoning. If imagery and perception share representations, that is, share the same neural code, they will evoke similar activation patterns in fMRI. Thus, training a SVC on activation patterns evoked during imagery and testing it on activation patterns evoked during perception amounts to testing whether imagery and perception share representations.

5-fold cross-validation was carried out by repeating the classification procedure, each time with pattern vectors from a different run assigned to the independent test data set. Decoding results (decoding accuracies) were averaged over these 5 iterations. We conducted second-level analyses across identified ROIs on decoding accuracies by means of repeated-measures analyses of variance (ANOVAs), one-sample t-tests against chance level and paired t-tests. For repeated-measures ANOVAs and paired t-tests comparing category-selective regions with each other, missing ROIs were excluded case by case. All t-tests were Bonferroni corrected.

Analysis 1: Representation of Category

We investigated whether perception and imagery share representations of object category in category-selective regions. Thus, a SVC was trained to discriminate between activation patterns evoked by imagery of object categories and tested on activation patterns evoked by perception of the same categories. In detail, activation patterns evoked by object imagery in both locations were assigned to the training set. Activation patterns evoked by object perception were assigned to the test sets (Fig. 2A), and the SVC was tested separately for each location, that is, twice. This analysis was conducted for all possible category pairs separately in each of the category-selective regions. For each category-selective region, we grouped decoding results as indicating either preferred or nonpreferred category
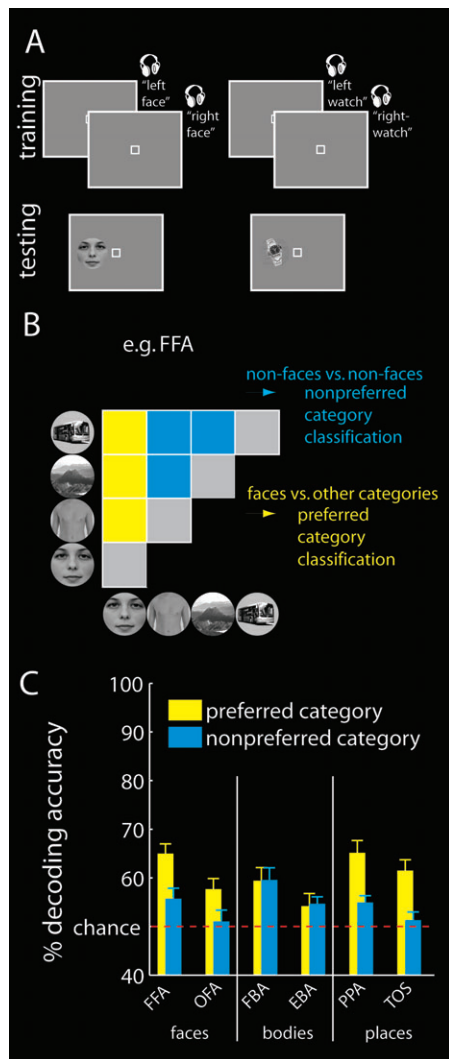
Figure 2. Imagery and perception share representations of category in category-selective regions. (*A*) We investigated whether imagery and perception share representations of category. Thus, we trained a support vector classifier (SVC) to discriminate between activation patterns evoked by imagery of categories and then tested on activation patterns evoked during perception of categories. In detail, patterns evoked during perception both left and right of fixation were used to train the SVC. Then, the SVC was tested separately for each location, left and right (only the left condition is shown here). (*B*) The matrix represents a schema of decoding accuracy for all possible category pair discriminations. For each category-selective region, we grouped decoding results as indicating either preferred or nonpreferred category information. In detail, for each category-selective region (here: FFA), decoding results of discriminations involving the preferred category (here: faces) were averaged and considered to indicate preferred category information. In contrast, decoding results of discriminations not involving the preferred category (e.g., scenes vs. body parts) were averaged and considered to indicate nonpreferred category information. (*C*) Category representations shared by imagery and location in category-selective regions. In all, 5 of 6 category-selective regions (except EBA) contained information about the preferred category. Also, 4 regions (except OFA and TOS) contained information about nonpreferred categories. This indicates that category-selective regions contain representations about both preferred and nonpreferred categories shared by perception and imagery. Further, face-and place-selective regions contained more information about their preferred category than about nonpreferred categories, indicating that category selectivity is retained to the same representations in perception and imagery.

information (Fig. 2*B*). In detail, for each category-selective region (e.g., FFA) decoding results of discriminations involving the preferred category (i.e., faces) were averaged and considered to indicate "preferred category information." In contrast, decoding results of

discriminations not involving the preferred category (e.g., in FFA scenes vs. body parts) were averaged and considered to indicate "nonpreferred category information."

Analysis 2: Representation of Object Location
To investigate whether perception and imagery share representations of location, we asked whether activation patterns evoked by imagery of categories predict the location of perceived categories. A SVC was trained to distinguish activation patterns evoked by a category imagined either left or right of fixation (Fig. 3*A*). Then the SVC was tested on activation patterns evoked by another category perceived at identical locations. This analysis was conducted for all possible category pairs and the results were averaged across categories. We carried out an identical analysis in each category-selective regions and subregions of LOC, as well as in low-level visual cortex (V1, V2, and V3).

Results

Representations of Category Shared by Imagery and Perception in Category-Selective Regions
We investigated whether different category-selective regions contain representations of preferred and nonpreferred categories shared by imagery and perception (Fig. 2*A,B*). In 5 of 6 category-selective regions (except EBA), the preferred category could be decoded with accuracies significantly above chance (all $P < 0.002$, Fig. 2*C*, Supplementary Table 2). Remarkably, 4 of 6 regions (except OFA and TOS) showed significant decoding accuracies for nonpreferred categories as well (all $P < 0.05$). This indicates that most category-selective regions contain representations shared by perception and imagery of preferred as well as nonpreferred categories.

Next, we asked whether category-selective regions retain their perceptually defined category preference (Supplementary Analysis 1) to the same representations also during imagery. Category-selective regions have been suggested and shown to differ systematically in their response profile during perception dependent on whether they are situated laterally or ventrally (Hasson et al. 2003; Schwarzlose et al. 2008). Thus, we also asked whether a systematic difference was present for category information shared by imagery and perception. We grouped category-selective regions into pairs, such that each region pair was defined to the same category preference with one region in each position on the cortical surface (ventral: FFA, FBA, PPA; lateral: OFA, EBA, TOS). For each of the resulting 3 ROI pairs, we conducted a repeated-measures ANOVA on decoding accuracies for category classification with factors "cortical position" (ventral vs. lateral) and "category preference" (preferred vs. nonpreferred category). We found a main effect of category preference in face- and place-selective regions (all $P < 0.001$, Supplementary Table 3). This indicates that face- and place-selective regions retain their preference during imagery to the same representations as during perception. We also found a significant main effect of cortical position for face- and body-selective regions (all $P < 0.05$) and a trend for place-selective regions ($P = 0.062$). This indicates that regions positioned ventrally contained representations shared by imagery and perception to a greater extent than regions positioned laterally. No significant interaction effect was observed for any of the ANOVAs (all $P > 0.1$).
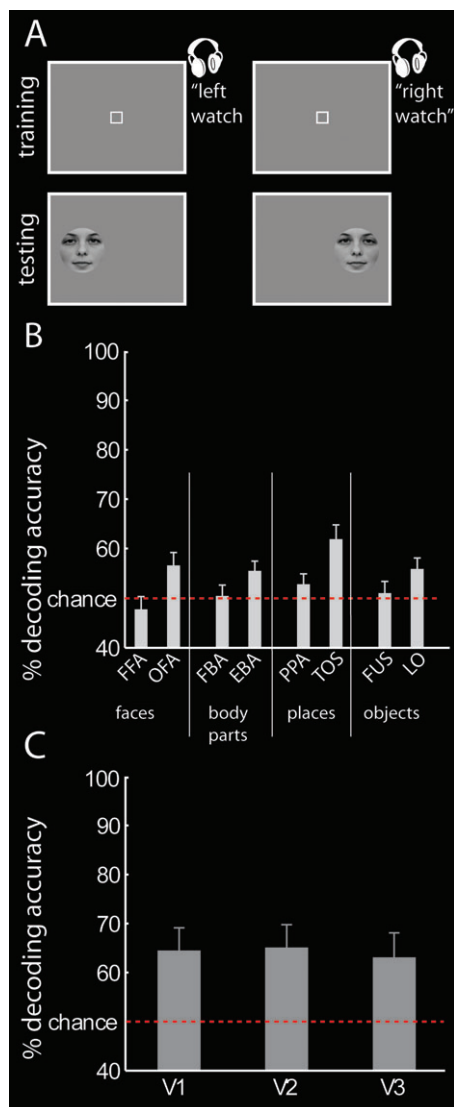
Figure 3. Imagery and perception share representations of location. (*A*) We investigated whether perception and imagery share representations of location. Thus, we asked whether activation patterns evoked by imagery of categories predict the location of perceived categories. A SVC was trained to distinguish activation patterns evoked by a category imagined either left or right of fixation. Then the SVC was tested on activation patterns evoked by another category perceived at identical locations. (*B*) Imagery and perception share location representations in high-level visual cortex. All lateral category-selective regions contained information about object location. In contrast, we found no evidence that ventral regions contained information about object location. The same was true for LO and FUS, the subregions of LOC. Further, for 3 out of 4 region pairs (except for EBA/FBA), there was more location information in lateral than in ventral regions. These results indicate that in high-level visual cortex representations of location shared by imagery and perception are distributed in a systematic fashion. (*C*) Imagery and perception share location representations in low-level visual areas V1, V2, and V3.

Representation of Object Location Shared by Imagery and Perception

We asked whether imagery and perception share representations of location in category-selective regions (Fig. 3*A*). As this analysis does not distinguish between preferred and non-preferred categories, we included the subregions of general object-selective LOC, that is, LO and FUS into the analysis. Thus, for each region, we tested decoding accuracies for location classification against chance by one-sample *t*-tests. The results

indicate that all regions positioned laterally contained significant information about object location (Fig. 3*B*, all $P < 0.005$, Supplementary Table 4). In contrast, we found no evidence for location information in regions positioned on the ventral surface of cortex (all $P > 0.2$, Supplementary Table 4). Please note that the lack of location information in ventral regions is not simply due to a lack of signal in these regions: as mentioned above, we found even more category information shared by imagery and perception in ventral than in lateral regions. A systematic difference in location information was ascertained by 4 repeated-measures *t*-tests on decoding accuracies for location classification for 4 region pairs as defined above. For 3 out of 4 region pairs (except for FBA/EBA) location representations proved to be present to a greater extent in lateral than in ventral high-level visual cortex (all $P < 0.05$, Supplementary Table 5). This indicates that imagery and perception share representations of object location to a greater extent in lateral than in ventral high-level visual cortex.

Next, we investigated whether perception and imagery share representations of location in low-level visual cortex. We conducted 3 one-sample *t*-tests (for V1, V2, and V3, respectively) of decoding accuracies for location classification against chance. Low-level visual regions V1, V2, and V3 (Fig. 3*C*, Supplementary Table 6) contained significant above-chance location information (all $P < 0.05$). This indicates that in low-level visual cortex imagery and perception share representations of location.

Discussion

In this study, we used fMRI and multivoxel pattern classification to investigate in 2 ways to which extent similarities between imagery and perception are reflected in their neural representations. First, we asked whether imagery and perception share representations of content, that is, the category of object a person was seeing. We found that imagery and perception share category representations in category-selective regions, and they do so to a greater extent in regions on the ventral than on the lateral cortical surface. Category-selective regions retained category selectivity during imagery to the same representations as during perception. Interestingly, also nonpreferred categories shared representations in category-selective regions. Second, we investigated whether perception and imagery share representations of location, that is, where an object is seen to be. We found that low-level and high-level ventral visual cortex shared representations of object location. In high-level ventral visual cortex, the extent of location representation was dependent on the position of the cortical region: Lateral regions contained more location information than ventral regions. Our results have interesting implications in 2 domains, as outlined below. First, they elucidate the nature of representations underlying visual imagery. Second, they inform our understanding of the way high-level visual cortex subtends object recognition.

Shared Representations in Imagery and Perception

Prior imaging studies suggested that imagery and perception share representations of content in high-level ventral visual cortex based on similar overall BOLD activation levels during imagery and perception (Ishai et al. 2000; O'Craven and Kanwisher 2000). In line with this research a supplementary ROI analysis showed that similar to perception, imagery of

preferred categories activated category-selective regions more than imagery of nonpreferred categories (Supplementary Analysis 2). Further, a supplementary analysis suggested that voxels activated more by the preferred than by nonpreferred category in category-selective regions tend to overlap (Supplementary Analysis 11). However, it is important to establish content selectivity, that is, that the same similarity also holds at the level of individual representations of specific categories. Using multivoxel pattern classification, we showed that imagery and perception share representations of content, that is, category in most category-selective regions. Further, a supplementary analysis revealed that object-selective LOC also contained representations shared by imagery and perception for all categories investigated in this study (Supplementary Analysis 3). Importantly, the decoding of shared representations of category in imagery and perception in this experiment cannot be explained by mean activation differences alone. Visualization of discriminant weight patterns of the classifiers across imagery and perception indicated that classifiers relied on activation patterns, not mean-activation differences alone (Supplementary Figure 1). Moreover, supplementary analyses separating effects due to mean activation differences and due to differences in patterns only indicated that activation patterns contain information about both preferred and non-preferred categories, whereas mean activation level contained information about preferred categories only (Supplementary Analysis 4).

Thus, together with recent findings that demonstrated shared representations in imagery and perception in high-level cortex (Stokes et al. 2009; Reddy et al. 2010), our results provide a systematic and comprehensive survey of a common neural substrate for the content of imagery and perception in ventral visual cortex (Daselaar et al. 2010).

In contrast to high-level ventral visual cortex, a supplementary analysis did not reveal shared representations of category in low-level ventral visual cortex (Supplementary Analysis 5). The role of low-level visual cortex in imagery remains debated (for a review, see Kosslyn and Thompson 2003). We can only speculate about the reasons for our failure to find shared involvement of low-level visual cortex in imagery and perception of content. The concurrent spatial task might have had an interfering effect, or the reconstruction of high-resolution visual information during imagery was not sufficient to necessitate engagement of low-level visual cortex (Kosslyn et al. 2001; Kosslyn and Thompson 2003). However, we exclude the possibility that this result was trivially due to lack of information about object category in low-level visual cortex during perception (Supplementary Analysis 6). Thus, the extent to which imagery and perception engage the same representations in low-level visual cortex requires further investigation.

Importantly, for full compatibility between imagery and perception not only representations of content but also representations of the location of the content must be shared by imagery and perception. Here, we comprehensively investigated whether the ventral visual cortex encodes the location of the content of perception and imagery similarly. We found that imagery and perception share encoding of object location and that they did so to a greater extent in lateral than in ventral regions in high-level visual cortex. In addition, perception and imagery shared representations of stimulus location in low-level visual cortex, corroborating prior studies (Kosslyn

et al. 1995; Tootell 1998; Klein et al. 2004; Slotnick et al. 2005; Thirion et al. 2006). A supplementary analysis revealed that the decoding of shared representations of location in imagery and perception did not rely on mean activation differences (Supplementary Analysis 2). Furthermore, visualization of discriminant weight patterns learned by a classifier of object location indicated that the classifier learned activation patterns, not mean-activation differences for location classification only (Supplementary Figure 2). Thus, imagery and perception share representations of location by activation patterns, not by mean activation differences.

A complementary analysis classifying across imagery and perception in the reverse direction, that is, predicting imagined objects from activation patterns evoked during the perception of objects, yielded similar results for both category (Supplementary Analysis 7) and location classification (Supplementary Analysis 8). However, effect sizes were partly reduced and thus did not reach statistical significance in some cases. At first sight, this result may appear to be counterintuitive. However, imagery is less detailed than veridical perception. Correspondingly, the neural representation of an imagined object is presumably less detailed that the neural representation of a perceived object. If the discriminative features learned under imagery work better on perception than vice versa, this would suggest that the imagery representation comprises only a subset of features of the veridical representation. Or to put it differently: All the features learned under imagery are discriminative for representations in perception but not the other way round.

The above interpretation also yields the prediction that classification within imagery only should be less robust than classification across imagery and perception. A supplementary analysis conducted to classify location and category within imagery supports this interpretation. Although qualitatively similar to classification between imagery and perception, this analysis yielded relatively weak and unreliable results (Supplementary Analysis 9 and 10).

Prior imagery studies using multivoxel pattern classification yielded mixed results on decoding accuracy for within imagery and across imagery and perception decoding of object category in higher-level visual cortex. Reddy et al. (2010) reported very similar decoding accuracies for object classification within and across both imagery and perception. In contrast, Stokes et al. (2009) reported stronger decoding accuracy within imagery than across imagery and perception. However, the present study differs strongly in design (block and event related) and stimulus material (simple shapes, colour images gray-scale images) from the aforementioned studies. This suggests that strength of activation patterns in higher-level visual cortex during imagery and their similarity to patterns evoked during perception may depend strongly on the features of the imagined objects and on the time available for imagery. It is an interesting question for further research how both these factors, that is, object features and time, influence imagery and the neural of imagined contents.

It is unlikely that our results can be explained by eye movements. First, it is unlikely that participants made large eye movements, for example, fixated the pictures of the objects because doing so would have strongly deteriorated participants' task performance on the Landolt-C fixation task which requires central fixation. Also, if subjects had systematically made large eye movements toward the object and fixated it,

object information would be encoded in both hemispheres in early visual cortex, whereas we found encoding of perceived objects in early visual areas only in the hemisphere contralateral to fixation and no encoding in the ipsilateral hemisphere (Supplementary Analysis 6). We cannot finally exclude that participants fixated slightly differently or differentially suppressed eye movements for different objects or locations. However, small differences in fixation are unlikely to account for the results of shared representation between imagery and perception. For this, eye movements would have to be not only specific for each experimental condition but also identical for imagery and perception.

Taken together, our results demonstrate that imagery and perception share representations of both content and location in a systematic fashion in ventral visual cortex. Our results suggest that the similarity in experience between imagery and perception is mirrored in similar neural representations in the brain.

The Role of High-Level Visual Cortex in Object Recognition

Pattern classification indicates a statistical dependency between stimulus conditions and spatially distributed activation patterns (Kriegeskorte et al. 2006; Mur et al. 2009). Thus, to argue that the information read out by pattern classifiers actually is used by the brain, and does not merely reflect epiphenomenal engagement, always requires further argumentation (Kriegeskorte and Bandettini 2007). We suggest that our results provide a plausibility argument for the representations of category and location in high-level ventral visual cortex. When imagery, that is, internally and top-down–driven neural processing, and perception, that is, externally and bottom-up–driven processing, evoke the same representations, the functional role of these representations during perception gains plausibility. It is unlikely that 2 processes, which differ so dramatically in origin, would evoke the same cortical activations as a mere by-product. Based on this reasoning, our results have 3 implications for the role of visual representations in high-level visual cortex.

First, shared mechanisms of category preference in imagery and perception provide evidence for category preference as an organizing principle of conscious object representation in the brain (Op de Beeck et al. 2008; Mahon et al. 2009). Importantly, though prevalent, the category preference observed in our study was not absolute: Not only preferred categories but also nonpreferred categories shared representations during imagery and perception in most category-selective regions. The role of nonpreferred responses in representing objects of the nonpreferred category remains debated. One view holds that nonpreferred responses are part of a distributed and overlapping representation in ventral visual cortex (Haxby et al. 2001). In contrast, another view claims that nonpreferred responses do not play a role in representing objects of the nonpreferred category, but rather indicate mere epiphenomenal, automatic bottom-up processing of any visual stimulus (Spiridon and Kanwisher 2002). As in our experiment during imagery no visual stimulus was present, the presence of category representations shared by imagery and perception cannot be explained by automatic bottom-up visual processing. Specifically, this is the case for the presence of representations of nonpreferred categories in category-selective regions shared by imagery and perception.

Thus, by excluding the explanation by automatic bottom-up processing, our result suggests a role for nonpreferred responses in category-selective regions in the representation of the nonpreferred categories. Therefore, our results suggest an intermediate position (O'Toole et al. 2005) between models of object representation which are fully distributed (Haxby et al. 2001; Ewbank et al. 2005) and which are fully modular (Spiridon and Kanwisher 2002).

Second, imagery and location encoded object location in lateral regions of high-level visual cortex. This suggests that location information not only in the dorsal but also in parts of the ventral stream might be used by the brain during perception (Edelman and Intrator 2000). Thus, our results provide further support for recent studies (MacEvoy and Epstein 2007; Sayres and Grill-Spector 2008; Schwarzlose et al. 2008; Carlson et al. 2009) that question the description of the ventral visual stream as invariant to changes in object location (Ungerleider and Mishkin 1982).

Third, our results indicate a functional differentiation within the ventral visual stream into 2 parts, a lateral and a ventral part, based on the amount of information about content and location shared by imagery and perception. Ventral category-selective regions contained more information about category shared by imagery and perception than lateral regions. For location information, this pattern was reversed: Lateral category-selective regions contained more information about location than ventral regions. The same was true for lateral and ventral subdivisions of LOC. A possible explanation for greater presence of location information in lateral than ventral areas might be that ventral regions have larger receptive fields than lateral regions. This hypothesis could be tested by future studies that use stimuli optimized to elicit responses in high-level ventral visual cortex, rather than flickering checkerboard stimuli (Smith et al. 2001; Dumoulin and Wandell 2008). This difference in sensitivity to location and extent of feedback processing suggests different computational roles for high-level lateral and ventral cortex (Hasson et al. 2003; Schwarzlose et al. 2008). Whereas the lateral part might be more engaged in the processing of spatial aspects of objects, the ventral part might play a greater role in the processing of object identity tolerant to changes in viewing conditions (Grill-Spector et al. 1999; Beauchamp et al. 2002; Eger et al. 2008b). Thus, our results suggest a tentative explanation for the existence of 2 regions for each category in each hemisphere (Hasson et al. 2003). However, more experimental support is needed to conclusively establish these subdivisions.

Summary

Taken together, our results provide a comprehensive account of the extent to which imagery and perception share representations in high-level ventral visual cortex. Thereby, they delineate to which degree the experiential similarity between imagery and perception is mirrored in neural activity. Furthermore, they inform about the role of the ventral stream in object recognition as an interface between bottom-up sensory-driven and top-down memory-driven processing.

Supplementary Material

Supplementary Tables 1–6, Figures 1 and 2, and Analyses 1–11 can be found at: http://www.cercor.oxfordjournals.org/

Notes

Conflict of Interest: None declared.

References

Aguirre GK, Zarahn E, D'Esposito M. 1998. An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. Neuron. 21:373-383.

Bieg HJ, Chuang LL, Bülthoff HH. 2010. Does adding a visual task component affect fixation accuracy? Perception. 39:35.

Beauchamp MS, Lee KE, Haxby JV, Martin A. 2002. Parallel visual motion processing streams for manipulable objects and human movements. Neuron. 34:149-159.

Carlson T, Hogendoorn H, Fonteijn H, Verstraten FAJ. 2009. Spatial coding and invariance in object-selective cortex. Cortex. 47:14-22.

Carlson TA, Schrater P, He S. 2003. Patterns of activity in the categorical representations of objects. J Cogn Neurosci. 15:704-717.

Clark VP, Keil K, Maisog JM, Courtney S, Ungerleider LG, Haxby JV. 1996. Functional magnetic resonance imaging of human visual cortex during face matching: a comparison with positron emission tomography. Neuroimage. 4:1-15.

Cox DD, Savoy RL. 2003. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage. 19:261-270.

Craver-Lemley C, Reeves A. 1992. How visual imagery interferes with vision. Psychol Rev. 99:633-649.

Dale AM, Fischl B, Sereno MI. 1999. Cortical surface-based analysis: i. segmentation and surface reconstruction. Neuroimage. 9:179-194.

Daselaar SM, Porat Y, Huijbers W, Pennartz CM. 2010. Modality-specific and modality-independent components of the human imagery system. Neuroimage. 52:677-685.

DeYoe EA, Carman GJ, Bandettini P, Glickman S, Wieser J, Cox R, Miller D, Neitz J. 1996. Mapping striate and extrastriate visual areas in human cerebral cortex. Proc Natl Acad Sci U S A. 93:2382-2386.

Downing PE, Jiang Y, Shuman M, Kanwisher N. 2001. A cortical area selective for visual processing of the human body. Science. 293:2470-2473.

Dumoulin SO, Wandell BA. 2008. Population receptive field estimates in human visual cortex. Neuroimage. 39:647-660.

Edelman S, Intrator N. 2000. (Coarse coding of shape fragments) + (Retinotopy) representation of structure. Spatial Vis. 13:255-264.

Eger E, Ashburner J, Haynes J, Dolan RJ, Rees G. 2008a. fMRI activity patterns in human LOC carry information about object exemplars within category. J Cogn Neurosci. 20:356-370.

Eger E, Kell CA, Kleinschmidt A. 2008b. Graded size sensitivity of object-exemplar-evoked activity patterns within human LOC subregions. J Neurophysiol. 100:2038-2047.

Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. Nature. 392:598-601.

Ewbank MP, Schluppeck D, Andrews TJ. 2005. fMR-adaptation reveals a distributed representation of inanimate objects and places in human visual cortex. Neuroimage. 28:268-279.

Farah MJ. 1985. Psychophysical evidence for a shared representational medium for mental images and percepts. J Exp Psychol Gen. 114:91-103.

Farah MJ. 1989. Mechanisms of imagery-perception interaction. J Exp Psychol Hum Percept Perform. 15:203-211.

Gauthier I, Tarr MJ, Moylan J, Skudlarski P, Gore JC, Anderson AW. 2000. The fusiform "face area" is part of a network that processes faces at the individual level. J Cogn Neurosci. 12:495-504.

Grill-Spector K, Kushnir 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, Malach R. 2004. The human visual cortex. Annu Rev Neurosci. 27:649-677.

Guzman-Martinez E, Leung P, Franconeri S, Grabowecky M, Suzuki S. 2009. Rapid eye-fixation training without eyetracking. Psychon Bull Rev. 16:491-496.

Halgren E, Dale AM, Sereno MI, Tootell RB, Marinkovic K, Rosen BR. 1999. Location of human face-selective cortex with respect to retinotopic areas. Human Brain Map. 7:29-37.

Hasson U, Harel M, Levy I, Malach R. 2003. Large-scale mirror-symmetry organization of human occipito-temporal object areas. Neuron. 37:1027-1041.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science. 293:2425-2430.

Haxby JV, Hoffman EA, Gobbini MI. 2000. The distributed human neural system for face perception. Trends Cogn Sci. 4:223-233.

Haynes J, Rees G. 2005a. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat Neurosci. 8:686-691.

Haynes J, Rees G. 2005b. Predicting the stream of consciousness from activity in human visual cortex. Curr Biol. 15:1301-1307.

Haynes J, Rees G. 2006. Decoding mental states from brain activity in humans. Nat Rev Neurosci. 7:523-534.

Haynes J, Sakai K, Rees G, Gilbert S, Frith C, Passingham RE. 2007. Reading hidden intentions in the human brain. Curr Biol. 17:323-328.

Ishai A, Sagi D. 1995. Common mechanisms of visual imagery and perception. Science. 268:1772-1774.

Ishai A, Ungerleider LG, Haxby JV. 2000. Distributed neural systems for the generation of visual images. Neuron. 28:979-990.

Kamitani Y, Tong F. 2005. Decoding the visual and subjective contents of the human brain. Nat Neurosci. 8:679-685.

Kanwisher N, McDermott J, Chun MM. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci. 17:4302-4311.

Klein I, Dubois J, Mangin J, Kherif F, Flandin G, Poline J, Denis M, Kosslyn SM, Le Bihan D. 2004. Retinotopic organization of visual mental images as revealed by functional magnetic resonance imaging. Cogn Brain Res. 22:26-31.

Kosslyn SM, Thompson WL. 2003. When is early visual cortex activated during visual mental imagery? Psychol Bull. 129:723-746.

Kosslyn SM, Ganis G, Thompson WL. 2001. Neural foundations of imagery. Nat Rev Neurosci. 2:635-642.

Kosslyn SM, Thompson WL, Alpert NM. 1997. Neural systems shared by visual imagery and visual perception: a Positron Emission Tomography Study. Neuroimage. 6:320-334.

Kosslyn SM, Thompson WL, Klm IJ, Alpert NM. 1995. Topographical representations of mental images in primary visual cortex. Nature. 378:496-498.

Kriegeskorte N, Bandettini P. 2007. Analyzing for information, not activation, to exploit high-resolution fMRI. Neuroimage. 38:649-662.

Kriegeskorte N, Goebel R, Bandettini P. 2006. Information-based functional brain mapping. Proc Natl Acad Sci U S A. 103:3863-3868.

MacEvoy SP, Epstein RA. 2007. Position selectivity in scene- and object-responsive occipitotemporal regions. J Neurophysiol. 98:2089-2098.

Mahon BZ, Anzellotti S, Schwarzbach J, Zampini M, Caramazza A. 2009. Category-specific organization in the human brain does not require visual experience. Neuron. 63:397-405.

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB. 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci U S A. 92:8135-8139.

Mika S, Ratsch G, Tsuda K, Scholkopf B. 2001. An introduction ot kernel-based learning algorithms. IEEE Trans Neural Netw. 12:181-201.

Muller K, Mika S, Ratsch G, Tsuda K, Scholkopf B. 2001. An introduction to kernel-based learning algorithms. IEEE Trans Neural Netw. 12:181-201.

Mur M, Bandettini PA, Kriegeskorte N. 2009. Revealing representational content with pattern-information fMRI—an introductory guide. Soc Cogn Affect Neurosci. 4:101-109.

Norman KA, Polyn SM, Detre GJ, Haxby JV. 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn Sci. 10:424-430.

O'Craven KM, Kanwisher N. 2000. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. J Cogn Neurosci. 12:1013-1023.

Op de Beeck HP, Haushofer J, Kanwisher NG. 2008. Interpreting fMRI data: maps, modules and dimensions. Nat Rev Neurosci. 9:123-135.

O'Toole AJ, Jiang F, Abdi H, Haxby JV. 2005. Partially distributed representations of objects and faces in ventral temporal cortex. J Cogn Neurosci. 17:580-590.

Peelen MV, Downing PE. 2005. Selectivity for the human body in the fusiform gyrus. J Neurophysiol. 93:603-608.

Puce A, Allison T, Gore JC, McCarthy G. 1995. Face-sensitive regions in human extrastriate cortex studied by functional MRI. J Neurophysiol. 74:1192-1199.

Pylyshyn Z. 2002. Mental Imagery: in search of a theory. Behav Brain Sci. 25:157-182.

Reddy L, Kanwisher N. 2007. Category selectivity in the ventral visual pathway confers robustness to clutter and diverted attention. Curr Biol. 17:2067-2072.

Reddy L, Tsuchiya N, Serre T. 2010. Reading the mind's eye: decoding category information during mental imagery. Neuroimage. 50:818-825.

Sayres R, Grill-Spector K. 2008. Relating retinotopic and object-selective responses in human lateral occipital cortex. J Neurophysiol. 100:249-267.

Schwarzlose RF, Baker CI, Kanwisher N. 2005. Separate face and body selectivity on the fusiform gyrus. J Neurosci. 25:11055-11059.

Schwarzlose RF, Swisher JD, Dang S, Kanwisher N. 2008. The distribution of category and location information across object-selective regions in human visual cortex. Proc Natl Acad Sci U S A. 105:4447-4453.

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB. 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science. 268:889-893.

Slotnick SD, Thompson WL, Kosslyn SM. 2005. Visual mental imagery induces retinotopically organized activation of early visual areas. Cereb Cortex. 15:1570-1583.

Smith AT, Singh KD, Williams AL, Greenlee MW. 2001. Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. Cereb Cortex. 11:1182-1190.

Spiridon M, Kanwisher N. 2002. How distributed is visual category information in human occipito-temporal cortex? An fMRI study. Neuron. 35:1157-1165.

Stokes M, Thompson R, Cusack R, Duncan J. 2009. Top-down activation of shape-specific population codes in visual cortex during mental imagery. J Neurosci. 29:1565-1572.

Thirion B, Duchesnay E, Hubbard E, Dubois J, Poline J, Lebihan D, Dehaene S. 2006. Inverse retinotopy: inferring the visual content of images from brain activation patterns. Neuroimage. 33:1104-1116.

Tootell R. 1998. From retinotopy to recognition: fMRI in human visual cortex. Trends Cogn Sci. 2:174-183.

Tye M. 2000. The Imagery Debate. Cambridge (MA): MIT Press.

Ungerleider LG, Mishkin M. 1982. Two cortical visual systems. In: Ingle DJ, Goodale MA, Mansfield, editors. Analysis of visual behavior. Cambridge (MA): MIT Press.

Wandell BA, Chial S, Backus BT. 2000. Visualization and measurement of the cortical surface. J Cogn Neurosci. 12:739-752.

Wandell BA, Dumoulin SO, Brewer AA. 2007. Visual field maps in human cortex. Neuron. 56:366-383.

Williams MA, Baker CI, Op de Beeck HP, Mok Shim W, Dang S, Triantafyllou C, Kanwisher N. 2008. Feedback of visual object information to foveal retinotopic cortex. Nat Neurosci. 11:1439-1445.