Principles governing the large-scale organization of object selectivity in ventral visual cortex

by
Rebecca Frye Schwarzlose

B.A. Psychology
Northwestern University, 2002

Submitted to the Department of Brain and Cognitive Sciences in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Neuroscience
at the
Massachusetts Institute of Technology
June, 2008

Signature of Author: _____

Department of Brain and Cognitive Sciences
February 7th, 2008

Certified by: _____

Nancy Kanwisher
Ellen Swallow Richards Professor of Cognitive Neuroscience
Thesis Supervisor

Accepted by: _____

Matthew Wilson
Professor of Neurobiology
Chairman, Committee for Graduate Studies

Principles governing the large-scale organization of object selectivity in ventral visual cortex

by

Rebecca Frye Schwarzlose

Submitted to the Department of Brain and Cognitive Sciences on February 7th, 2007 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Neuroscience

Abstract

As our understanding of the brain grows, neuroscientists find themselves increasingly in the role of cartographer. Thus far, cortical maps have been found primarily in early input and late output areas, however they may also occur in higher-level regions of the brain that perform more complex functions. An example of such a region is the object-selective ventral visual cortex (VVC) in humans. This region, which is involved in the high-level task of object recognition, is comprised of several functionally defined, category-selective subregions that are laid out with remarkable systematicity and consistency across individuals.

In this thesis, I use fMRI to test several hypotheses about the nature of object representations and the dimensions along which object-selective cortex might be organized. In the first study, I find evidence supporting the existence of domain-specific regions. Results from the second set of studies suggest that temporal associations do not guide the overall organization of VVC, and also provide contradictory evidence against a long-standing hypothesis that the VVC is organized based on conceptual knowledge about objects and, specifically, the distinction between animate and inanimate objects. Instead, my results suggest that associations between objects and motor actions may play a role in the location of category selectivities for a subset of object classes. Results from a third set of studies demonstrate that computational demands for acuity or spatial integration cannot account for location biases in category-selective regions, and instead suggest that experience with objects at specific retinal locations may serve as an organizing dimension. Moreover, these studies reveal systematic differences in the amount of location information contained in category-selective regions on the ventral temporal versus lateral occipital surfaces.

In sum, the studies described in this thesis address several hypotheses about the large-scale organization of VVC, and, in doing so, advance our understanding of the principles that govern the layout of maps in higher-level, object-selective cortex.

Acknowledgments

Thank you . . .

To my advisor, Nancy Kanwisher, who has been a patient teacher, fierce advocate, and a dear friend. She's been wonderful beyond what any reasonable student could hope for, and I can't thank her (and my lucky stars) enough.

To my committee members Marvin Chun, Bob Desimone, and Pawan Sinha, and my advisory committee members Molly Potter and Jim DiCarlo, for their invaluable feedback and guidance.

To Ann Graybiel, for bringing me to MIT and sharing her energy and wisdom with me, and to the entire Graybiel lab, who welcomed me here. I especially want to thank Esen Saka, who was an amazing teacher and a wonderful listener.

To all of my teachers at MIT, and to the other faculty and support staff, especially the perennially-supportive Denise Heintze and Bettiann McKay. I would also like to thank Mrs. Singleton and Mr. Razin, for their generosity and belief in the power of neuroscience to heal.

To the essential technical guidance of many wonderful people whose time, effort and patience made my work possible, especially Nick Knouf, Sabin Dang, Jason Webster, Christina Triantafyllou, Pat Harlan, and Henry Hall.

To everyone at the Kanwisher lab, past and present, whose scholarly feedback and hilarious antics made every day one of growth and laughter (particularly Danny Dilks and Mark Williams, who gave me both in abundance). Most of all, I want to thank Chris Baker, who was with me from the first slice prescription and has been there in one way or another at every step since.

To the fabulous grad students and friends I've met and come to love during my time at MIT, especially Alex Rivest, Amy Pooler, Arvind Govindarajan, Kristin MacCully and Theresa Desrochers, whose ears I've bent and shoulders I've leaned on often enough to deserve a restraining order.

To my family: John, Darlene, and Terri, who cheered me on, and my brother Dan, who always supported his nerdy little sister and wore his MIT shirt with pride.

To my partner Sabin, whose love and encouragement made the path feel shorter and the road less rocky. His unwavering belief that I could do anything has been a precious gift.

And to my mom, who is my best friend and hero, for her never-ending support in all things. There is truly nothing I could say or do that would be deep and wide and endless enough to contain my gratitude.

This thesis is dedicated to my other best friend, my father, who would have loved to see this day.

Table of Contents

1. Introduction

Over the past few decades, neuroscientists have hit upon a recurrent theme in the organization of the brain: maps. Neural processing at the early stages of nearly all sensory input takes place within well-defined cortical maps, as do the final stages of motor output. While neuroscience has come a long way in characterizing the properties of these primary input and output maps, very little is know about what topographic axes govern the layout of higher-level areas that process more complex representations, or even if such maps exist. Human ventral visual cortex (VVC), a large region of cortex dedicated to the complex task of face and object recognition, is one such higher-level area that contains numerous subregions with distinct selectivities for specific object categories. The fundamental question addressed in this thesis is whether the functionally defined subregions in VVC are components of a larger cortical map and, specifically, what dimensions might govern the layout of such a map.

The VVC in humans and the homologous inferotemporal cortex (IT) in monkeys

are thought to be devoted to recognition of complex shapes and objects. Extracellular recordings in monkey IT have revealed considerable diversity in the types of shapes and objects that cells prefer (Gross et al., 1972; Desimone and Gross, 1979; Sato et al., 1980; Fuster and Jervey, 1982), and have found that neurons with similar response properties may cluster together to form cortical columns (Tanaka, 1997). Clusters of category selectivity have even been found at a scale large enough to be seen with fMRI in monkey extrastriate cortex (Tsao et al., 2003; Pinsk et al., 2005; Op de Beeck et al., 2007). Functional MRI studies in humans have also revealed large-scale clustering of category selectivity in extrastriate cortex (specifically for faces, bodies, and scenes). These functionally defined regions of selectivity include two face-selective areas (the fusiform face area, or FFA, and occipital face area, OFA (Allison et al., 1994; Kanwisher et al., 1997; McCarthy et al., 1997)), two body-selective areas (the fusiform body area, FBA (Peelen and Downing, 2005), and extrastriate body area, EBA (Downing et al., 2001)), two scene-selective areas (the parahippocampal place area (Aguirre et al., 1996; Epstein and Kanwisher, 1998), PPA, and a transverse occipital sulcus area (Nakamura et al., 2000; Grill-Spector, 2003), TOS), as well as two general shape-selective areas that comprise the lateral occipital complex, or LOC (Kourtzi and Kanwisher, 2000; Grill-Spector et al., 2001): a posterior fusiform area, pFs, and a lateral occipital area, LO. Of the two regions selective for each category, one is located on the ventral temporal surface of the brain while the other is found on the lateral occipital surface. Crucially, the locations of these category-selective regions in cortex and with respect to one another are largely consistent across testing sessions (Peelen and Downing, 2005) and across individuals (Spiridon et al., 2006), suggesting that the layout of these regions does not arise in a stochastic fashion and does not change appreciably over time. Indeed, fMRI studies of children as young as five to seven years of age have identified some of these regions in the same general cortical locations where

they are found in adulthood, although the size of some regions may continue to expand throughout childhood (Golarai et al., 2007; Scherf et al., 2007). Moreover, the fMRI pattern of response in VVC to faces and scenes (but not chairs or letter strings) were more similar in monozygotic than dizygotic twins, suggesting that hard-wired, genetic factors may guide the locations of selectivity for these specific object categories (Polk et al., 2007).

An ongoing debate focuses on whether the functionally defined regions in human extrastriate cortex are dedicated to processing a single object category (Kanwisher, 2000) or whether they form part of an overlapping map of graded object representations (Haxby et al., 2001). The first step in addressing how the VVC is organized will be to resolve this fundamental question. While Haxby and colleagues (Haxby et al., 2001) found that patterns of activation in the face-selective FFA could discriminate between non-face stimuli and the patterns of activation in the scene-selective PPA could discriminate between non-house stimuli, subsequent studies using independent stimulus sets (Spiridon and Kanwisher, 2002) and principal components analysis (O'Toole et al., 2005) found poor discrimination of non-preferred object categories in these regions. Further, such discrimination between nonpreferred categories drops to chance when more than a single object is present at a time (Reddy and Kanwisher, 2007). These results suggest that the functionally defined category-selective regions may be exclusively dedicated to processing their preferred categories. Extracellular recordings from a face-selective patch in monkey extrastriate cortex support the same conclusion, revealing strong selectivity for faces in 97% of visually-responsive cells (Tsao et al., 2006). However fMRI studies of the face-selective area FFA in humans have challenged this specificity, finding apparent selectivity for a specific non-face category: bodies and body parts (Kanwisher et al., 1997; Peelen and Downing, 2005; Spiridon et al., 2006). Chapter 2 of this thesis tests whether the FFA is selective for both face and body stimuli and finds a striking dissociation of face- and body

selectivity on the fusiform gyrus (Schwarzlose et al., 2005). These results further support the theory that functionally defined extrastriate regions may be uniquely dedicated to the processing of only one object category.

Based on the evidence that the category-selective regions of VVC are functionally distinct from each other and occupy consistent anatomical locations across individuals, one can ask whether their layout constitutes a larger map in object-selective cortex and, if so, what rules might govern the organization of this map. On a basic level, maps equate to the spatial layout of cells with similar response properties such that adjacent cells respond to similar stimuli or outputs, yet differ along a critical dimension. Multiple dimensions can even be mapped onto the same cortical region, as seen with the overlapping maps of retinotopic location, orientation, and spatial frequency in primary visual cortex. Similarly, the spatial pattern of object selectivities in VVC might be organized along any number of dimensions, theoretically as many dimensions as there are properties along which objects can differ. These organizing principles could include the basic visual features of the object, the pattern of associations of objects in the statistics of everyday visual experience, the computations required to recognize particular object categories, or the conceptual properties of the object itself. Chapters 3 through 5 of this thesis are dedicated to testing several specific hypotheses for dimensions that may determine the layout of object-selective extrastriate cortex.

Perhaps the most straightforward map that could be proposed for the VVC would be based on shape. Consistent with this idea, a recent fMRI study in monkeys shows a distribution of selectivity for novel objects spanning much of IT cortex, suggesting a pure "shape map" independent of that familiarity and meaning (Op de Beeck et al., 2007). However, only two monkeys were included in this study, therefore it would be impossible to

ascertain based on this study whether the layout of such a map would be consistent across individuals. Moreover, object shape per se does not appear to be the only dimension in the VVC, at least not among category-selective regions. For example, the body-selective region EBA responds to a variety of disparate shapes, including whole bodies in various positions, body parts such as hands and legs, and stick figures of bodies (Downing et al., 2001). This robust selectivity for category irrespective of shape indicates that any cortical organization based on shape would not be universal across the entire VVC.

Another dimension that may play a role in determining the layout of VVC is based on the statistics of our visual experience. At any given time, we see many different objects in our visual field, and some objects tend to co-occur with one another (such as tables and chairs). Perhaps objects that are temporally associated have similar patterns of response in the cortex. Electrophysiological studies in monkeys support this hypothesis; the response patterns in IT neurons to two temporally associated unrelated shapes tend to be more similar as the association is learned (Sakai and Miyashita, 1991; Erickson and Desimone, 1999; Messinger et al., 2001). The hypothesis considered in Chapter 3 of this thesis is that temporal association may influence the large-scale organization of object selectivity in human VVC.

Beyond visual and statistical properties, objects also differ in meaning. Evidence from studies of patients with brain damage has led to the proposal that the neural representations of objects are segregated by whether the objects are animate or inanimate (Warrington and McCarthy, 1983). Both human fMRI and monkey electrophysiology studies using standard and multivariate techniques have found evidence that the animate/inanimate distinction is a fundamental distinction reflected in neural representations of objects (Downing et al., 2006; Kiani et al., 2007; Kriegeskorte, Mur, Ruf, Kiani, Bodurka,

and Bandettini, HBM 2007). In Chapter 4 of this thesis we test the hypothesis that object animacy serves as an organizing dimension of VVC.

Another major hypothesis that has been put forth to explain the organization of VVC is that object-selective regions are arranged based on the computational requirements for processing different object categories (Levy et al., 2001; Malach et al., 2002). According to this idea, face selectivities land in parts of the cortex best suited for fine-grained acuity (i.e., fovea-biased cortex), whereas scene-selective regions of cortex land in parts of cortex best suited for large-scale integration (i.e., periphery-biased cortex) (Levy et al., 2001; Malach et al., 2002). Therefore, computational demands (and secondarily, eccentricity biases) may serve as organizing dimensions in VVC. Chapter 5 of this thesis is dedicated to testing and refining this hypothesis.

An intriguing and salient feature of VVC is the fact that category-selective regions come in pairs. Specifically, two functionally defined, extrastriate regions have been discovered for each of the major category selectivities (faces, bodies, scenes, and general shape), and in each case one is located on the ventral temporal cortical surface and the other is located on the lateral occipital cortical surface (Levy et al., 2001; Beauchamp et al., 2002; Hasson et al., 2003). As yet, little is known about the functional differences between the regions in each pair, although a few studies have found greater sensitivity to motion in the regions on the lateral surface (Beauchamp et al., 2002; Hasson et al., 2003). However, these results may be due to the overlap of these regions with the motion-selective area MT (Spiridon et al., 2006) and thus may not reflect the properties of the category-selective regions themselves (Downing et al., 2007). Another possibility is that the category-selective regions of VVC differ in their sensitivity to location. The long-standing Dual Pathway Model (Underleider and Mishkin, 1982) divides extrastriate cortex into two pathways:

a dorsal pathway dedicated to processing object shape and a ventral one dedicated to processing object location. The set of category-selective regions on the lateral occipital surface falls anatomically between the cortex associated with the two pathways. Along the same lines, this set of regions may fall functionally between the focus on shape in the ventral regions and the focus on location in the dorsal regions, and therefore may demonstrate more sensitivity to location than their ventral counterparts. In Chapter 5 of this thesis, we test the hypothesis that regions on the lateral surface of the temporal lobe contain more location information than regions on the ventral surface.

The hypotheses listed here all propose different dimensions that may characterize the consistent and systematic layout of functional subregions in VVC. By testing these hypotheses, this thesis will address the fundamental question of what principles govern the large-scale organization of object-selective cortex and, in doing so, whether neural maps might occur in high-level cortical areas.

References

Aguirre GK, Detre JA, Alsop DC, D'Esposito M (1996) The parahippocampus subserves topographical learning in man. Cereb Cortex 6:823-829.

Allison T, Ginter H, McCarthy G, Nobre AC, Puce A, Luby M, Spencer DD (1994) Face recognition in human extrastriate cortex. J Neurophysiol 71:821-825.

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

Desimone R, Gross CG (1979) Visual areas in the temporal cortex of the macaque. Brain Research 178:363-380.

Downing PE, Wiggett AJ, Peelen MV (2007) Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. J Neurosci 27:226-233.

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

Downing PE, Chan AW, Peelen MV, Dodds CM, Kanwisher N (2006) Domain specificity in visual cortex. Cereb Cortex 16:1453-1461.

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

Erickson CA, Desimone R (1999) Responses of macaque perirhinal neurons during and after visual stimulus association learning. J Neurosci 19:10404-10416.

Fuster JM, Jervey JP (1982) Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. J Neurosci 2:361-375.

Golarai G, Ghahremani DG, Whitfield-Gabrieli S, Reiss A, Eberhardt JL, Gabrieli JD, Grill-Spector K (2007) Differential development of high-level visual cortex correlates with category-specific recognition memory. Nat Neurosci 10:512-522.

Grill-Spector K (2003) The neural basis of object perception. Curr Opin Neurobiol 13:159-166.

Grill-Spector K, Kourtzi Z, Kanwisher N (2001) The lateral occipital complex and its role in object recognition. Vision Res 41:1409-1422.

Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of the macaque. J Neurophysiol 35:96-111.

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.

Kanwisher N (2000) Domain specificity in face perception. Nat Neurosci 3:759-763.

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

Kiani R, Esteky H, Mirpour K, Tanaka K (2007) Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. J Neurophysiol 97:4296-4309.

Kourtzi Z, Kanwisher N (2000) Cortical regions involved in perceiving object shape. J Neurosci 20:3310-3318.

Levy I, Hasson U, Avidan G, Hendler T, Malach R (2001) Center-periphery organization of human object areas. Nat Neurosci 4:533-539.

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

McCarthy G, Puce A, Gore JC, Allison T (1997) Face-specific processing in the human fusiform gyrus. J Cogn Neurosci 9:605-610.

Messinger A, Squire LR, Zola SM, Albright TD (2001) Neuronal representations of stimulus associations develop in the temporal lobe during learning. Proc Natl Acad Sci U S A 98:12239-12244.

Nakamura K, Kawashima R, Sato N, Nakamura A, Sugiura M, Kato T, Hatano K, Ito K, Fukuda H, Schormann T, Zilles K (2000) Functional delineation of the human occipito-temporal areas related to face and scene processing. A PET study. Brain 123 (Pt 9):1903-1912.

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.

Op de Beeck HP, Deutsch JA, Vanduffel W, Kanwisher NG, Dicarlo JJ (2007) A Stable Topography of Selectivity for Unfamiliar Shape Classes in Monkey Inferior Temporal Cortex. Cereb Cortex.

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

Pinsk MA, DeSimone K, Moore T, Gross CG, Kastner S (2005) Representations of faces and body parts in macaque temporal cortex: a functional MRI study. Proc Natl Acad Sci U S A 102:6996-7001.

Polk TA, Park J, Smith MR, Park DC (2007) Nature versus nurture in ventral visual cortex: a functional magnetic resonance imaging study of twins. J Neurosci 27:13921-13925.

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

Sakai K, Miyashita Y (1991) Neuronal organization for the long-term memory of paired associates. Nature 354:152-155.

Sato T, Kawamura T, Iwai E (1980) Responsiveness of inferotemporal single units to visual pattern stimuli in monkeys performing discrimination. Exp Brain Res 38:313-319.

Scherf KS, Behrmann M, Humphreys K, Luna B (2007) Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. Dev Sci 10:F15-30.

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

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

Spiridon M, Fischl B, Kanwisher N (2006) Location and spatial profile of category-specific regions in human extrastriate cortex. Hum Brain Mapp 27:77-89.

Tanaka K (1997) Columnar organization in the inferotemporal cortex. In: Cerebral Cortex: Extrastriate Cortex in Primates (Rockland KS, Kaas JH, Peters A, eds), pp 469-498. New York: Plenum.

Tsao DY, Freiwald WA, Tootell RB, Livingstone MS (2006) A cortical region consisting entirely of face-selective cells. Science 311:670-674.

Tsao DY, Freiwald WA, Knutsen TA, Mandeville JB, Tootell RB (2003) Faces and objects in macaque cerebral cortex. Nat Neurosci 6:989-995.

Underleider LG, Mishkin M (1982) Two Cortical Visual Systems. In: Analysis of Visual Behavior (Ingle MA, Goodale MI, Masfield RJW, eds), pp 549-586. Cambridge, MA: MIT Press.

Warrington EK, McCarthy R (1983) Category specific access dysphasia. Brain 106 (Pt 4):859-878.

2. Separate Face and Body Selectivity on the Fusiform Gyrus

The following chapter appeared as:

Schwarzlose RF, Baker CI, Kanwisher N (2005). Separate Face and Body Selectivity on the Fusiform Gyrus. Journal of Neuroscience 25:11055-11059.

Introduction

Does the ventral visual pathway contain cortical regions that are selectively involved in processing just a single class of visual stimuli (Kanwisher et al., 1997;Allison et al., 1994), or are all regions of the ventral pathway instead involved in graded and overlapping representations of multiple stimulus classes (Haxby et al., 2001)? Faces have served as a key test case for this debate, based in part on the fact that the fusiform face area (FFA) is activated considerably more strongly by images of faces than by other object classes. In this chapter we address an important challenge to the claimed face selectivity of the FFA that arises from recent reports that the FFA may also respond strongly to images of bodies.

Specifically, two studies have found responses in the FFA that were higher to headless bodies than to control objects, though lower to bodies than faces (Peelen and Downing, 2005;Kanwisher et al., 1999; see also Cox et al., 2004;Hadjikhani and de Gelder, 2003 for responses to body stimuli in the FFA). One study even found that the FFA response was not significantly lower to body parts than to faces (Spiridon et al., 2006). Peelen and Downing also report a fusiform region that we will call the "fusiform body area," or FBA, that is adjacent to and overlapping with the FFA and that responds more strongly to headless bodies than to objects, but equally to headless bodies and faces. (Note that the FBA is located on the ventral surface of the brain, far from the extrastriate body area, or EBA, (Downing et al., 2001) which is on the lateral surface of the temporal lobe.) Collectively, these findings suggest a graded and overlapping pattern of responses in the fusiform gyrus (Haxby et al., 2001) rather than a strict spatial segregation of responses to faces and bodies. Here we used a scanning resolution higher than that of previous studies to test the hypothesis that the apparent dual selectivity of both the FFA and the FBA for both faces and bodies may result from blurring of the responses from two adjacent but distinct cortical regions, one selectively responsive only for faces, and the other only for bodies.

To do this, we identified the FFA with a blocked localizer scan and then tested its response magnitude to a variety of face, body, and assorted everyday object stimuli with an event-related design administered in the same participants and scan sessions. We conducted this study at standard resolution (3.125 x 3.125 x 4.0mm) in Experiment 1, and at a higher resolution (1.4 x 1.4 x 2.0mm) in Experiment 2. Using these methods, we found that the response to faces and bodies on the fusiform gyrus could be clearly dissociated with higher resolution imaging.

Materials and Methods

Stimuli and Design. Participants performed both localizer scans (to identify regions of interest) and event-related scans to test the selectivity of the ROIs. They completed five runs of the localizer scan, each of which included three 16-second fixation periods and two 16-second blocks of each of five stimulus classes (faces, headless bodies, scenes, assorted everyday objects, and scrambled versions of the everyday objects). The order of conditions was palindromic within a scan and the serial position of each condition was counterbalanced within participants across runs. Within each block, participants viewed twenty images of a single stimulus class (300 ms per image, with a 500 ms ISI). Scrambled object stimuli were constructed by superimposing a grid over the objects and relocating the component squares randomly. As participants watched these stimuli, they performed a 1-back task in which they were asked to make a key-press whenever an image was repeated consecutively. The images were jittered slightly in their location on the screen to preclude use of low-level transients in performing the 1-back task. Participants completed 6-8 runs of the event-related experiment in the same scan session. Each of the runs was composed of a quasi-random order of stimuli from the following four stimulus conditions: faces, headless bodies, body parts, and cars. A fifth condition of assorted everyday objects was added for the high resolution scans in Experiment 2. Other stimulus conditions that were included in the event-related design to test different hypotheses will not be discussed here. Each image moved either downward or to the left and the participants' task was to identify the direction of motion of each stimulus by pressing one key to indicate movement to the left and another to indicate movement down. There were fifteen images per stimulus category and thirteen image presentations per stimulus category per run. Different images were used in the localizer and the event-related experiment.

Functional Imaging. Participants were scanned on a 3.0 T Siemens Trio scanner at the Martinos Center for Biomedical Imaging in Charlestown, MA. Images were acquired with a Siemens 8-channel phased-array head coil and gradient echo single-shot echo planar imaging sequence. For Experiment 1 conducted at standard resolution: 28 slices covered the whole brain (dimensions 3.125 x 3.125 x 4.0mm, interslice gap 0.8mm; repetition time 2 seconds, echo time 30 ms). For Experiment 2 conducted at higher resolution, 15-18 slices were oriented roughly perpendicular to the calcarine sulcus (1.4 x 1.4 x 2.0mm; interslice gap 0.4mm; repetition time 2 seconds; echo time 33 ms). For both experiments, high-resolution MPRAGE anatomical images were also acquired for each participant. Seven participants were scanned for Experiment 1 and ten for Experiment 2. The data from one of the participants in Experiment 2 were excluded from the analysis because of excessive head motion.

Data Analysis was performed using Freesurfer and FS-FAST software (http://surfer.nmr.mgh.harvard.edu/). Before statistical analysis, images were motion corrected (Cox and Jesmanowicz, 1999) and smoothed, for the localizer runs only (5 mm full width at half maximum Gaussian kernel for Experiment 1, 3 mm for Experiment 2).

Our fMRI analyses focused on the right hemisphere because prior work has shown that the FFA is larger and more consistent in the right hemisphere (Kanwisher et al., 1997) and because it is only the right FFA that has been claimed to be strongly category selective (Grill-Spector et al., 2004). Regions of interest were visualized on slices and defined individually for each participant using the blocked localizer scans (as described below); we then used fROI (http://froi.sourceforge.net/) to extract the time courses of response for the event-related experiments in each ROI (see Figure 2). Critically, the data used to define the ROIs were independent of the data used to calculate the response magnitudes for each stimulus

category in each ROI.

Cortical surfaces were reconstructed using Freesurfer for three of the participants in Experiment 2 based on prior anatomical scans of those participants.

Results

Experiment 1. The results of our standard resolution scans replicated prior studies (Spiridon et al., 2006;Peelen and Downing, 2005). The localizer data were used to identify both a face-selective right FFA (using a contrast of faces > objects), and a body-selective right FBA (using headless bodies > objects) with a threshold for both contrasts of $p < 0.0001$ uncorrected in individual participants. A right FFA was identified in every participant and a replicable right FBA was found in five of the seven participants. One participant had no FBA and another had a very small FBA (eight voxels) which failed to replicate the body selectivity observed in the localizer scans in subsequent event-related scans; these two participants were excluded from further FBA analyses. The FFA and FBA overlapped in all participants who showed both ROIs. The average size of the right FFA ROIs was 1.53 cm^3. Of the five participants with a right FBA, the average size of the FBA ROI was 0.86 cm^3 and the overlap averaged 0.45 cm^3 per participant.

The time course of the response from the event-related runs for each stimulus category in the right FFA and right FBA ROIs (defined from the localizer scans) are shown in Figure 2a and 2b. In an ANOVA on the peak response magnitudes among the five participants that demonstrated both FFA and FBA ROIs, the interaction of stimulus condition by ROI did not reach significance ($F(2,3) = 5.8$, $p = 0.09$). Planned comparisons revealed that the response to headless bodies in the FFA was significantly greater than to cars ($p < 0.005$),

while still significantly lower than to faces ($p < 0.01$). In contrast, in the FBA the responses to images of both faces and headless bodies were significantly higher than to cars (both $p < 0.05$), while there was no difference in the degree of activation for face and headless body stimuli in this region ($p > 0.8$). These data replicate the results reported by Peelen and Downing (2005) at a similar scanning resolution (3.75 x 3.75 x 5.0mm), indicating elevated responses to both faces and bodies in both the FFA and FBA.

To test whether the high FFA response we observed to bodies might be due to the inference of a face from the headless body stimuli (Cox et al., 2004) we compared the FFA response to headless bodies (where a face might be inferred) with its response to assorted body parts (where the inference of a face is unlikely). In the FFA, the response to body parts was greater than to cars ($p < 0.005$) and lower than to faces ($p < 0.007$), while there was no significant difference between the responses to body parts and headless bodies ($p > 0.9$). These results argue against the possibility that the high responses to body stimuli in the right FFA are due to the inference of a face.

Experiment 2. The second experiment was the same as Experiment 1 except that it was conducted at higher resolution (voxel size 1.4 x 1.4 x 2.0mm) and a second baseline condition of assorted everyday objects was added. We isolated the right FFA and FBA ROIs using the same methods described above for Experiment 1. See Figure 1 for examples of slices and surface plots showing these ROIs at high resolution. Figure 2c and 2d show the time courses of the responses to each stimulus condition in the event-related runs for the high resolution FFA and FBA ROIs.

An ANOVA on the magnitude of the peak response in the event-related experiments revealed a significant interaction of ROI by stimulus condition, ($F(2,7) = 23.3$, $p < 0.002$). Planned comparisons confirmed, as expected, that the selectivity of the FFA and the FBA

for their preferred categories (on the basis of which of these regions were identified in the localizer scan) was replicated in the event-related scans: the FFA responded significantly more strongly to faces than to cars, objects, headless bodies, and body parts (all four $p < 0.005$), and the FBA responded significantly more strongly to headless bodies and body parts than to mixed objects and cars (all four $p < 0.005$).

Following up on the interaction of ROI by stimulus category, planned comparisons tested whether any selectivity for bodies could be found in the FFA and whether any selectivity for faces could be found in the FBA. In the FFA, the response to headless bodies was not significantly greater than to cars or mixed objects, nor was it significantly different from the response to body parts (all $p > 0.20$). Although the response to body parts was not significantly different from that to cars ($p > 0.10$), it trended toward a higher response than to mixed objects ($p < 0.05$ uncorrected, a significance level that would not survive a correction for the four comparisons of body stimuli to control object stimuli).

In the FBA, the response to the headless body stimuli trended toward a higher response than to faces ($p < 0.05$ uncorrected), while the response to body parts did not significantly differ from that to faces ($p > 0.30$). Nonetheless, the FBA response to faces was still higher than to mixed objects ($p < 0.01$), while not significantly different from that to cars ($p = 0.09$).

Thus, the selectivity of the FFA and FBA ROIs are stronger at high resolution than standard resolution, a result supported by a significant triple interaction (among the participants who had both FFAs and FBAs) of ROI, stimulus type, and experiment ($F(2,11) = 5.07$, $p < 0.05$). This difference in selectivity might, at least in part, be attributable to the decrease in partial voluming between face-selective and body-selective regions at high resolution. This potential explanation is supported by the reduced volume of the overlap

between the FFA and FBA ROIs from standard to high resolution. Specifically, the average overlap at standard resolution constituted 0.45 cm^3, or 27% of the total sum of the FFA and FBA ROIs, across the five standard resolution participants that demonstrated both ROIs, while the corresponding overlap was 0.17 cm^3, or 18% of the summed FFA and FBA ROI volumes across high resolution participants. However despite this reduction of overlap at higher resolution, we still observed trends of higher responses to body parts than to control stimuli in the FFA, and higher responses to faces than to control stimuli in the FBA.

We next attempted a stronger test of our hypothesis that the dual selectivity of the FFA and FBA may result from the pooling of responses from two adjacent but distinct cortical regions, one selective for only faces and the other selective for only bodies. To do this, we used a new ROI selection method in which we omitted from the right FFA ROI (which had an average of 162 voxels, or 0.64 cm^3) all voxels that were also included in the FBA ROI, to generate a new FFA* ROI (mean 121 voxels, or 0.47 cm^3), and we omitted from the right FBA ROI (mean 96 voxels, or 0.38 cm^3) all voxels that were also included in the FFA ROI, to generate a new FBA* ROI (mean 55 voxels, or 0.22 cm^3). Importantly, as with our earlier analyses, the blocked localizer data used to identify the FFA* and FBA* were independent from the event-related data we used to assess selectivity profiles in these ROIs (see Figure 2).

An ANOVA on the magnitude of the peak responses in the event-related data in these new ROIs revealed a strong interaction of ROI (FFA* vs FBA*) by stimulus condition (faces, headless bodies, and cars), ($F(2,7) = 31.4$, $p < 0.001$). As expected from previous analyses, the FFA* response was significantly higher to faces than to headless bodies, body parts, cars, and mixed objects (all four $p < 0.001$). More importantly, neither headless

bodies nor body parts produced a higher response in the FFA* than did either cars or mixed objects (all four $p > 0.3$). Conversely, the FBA* responses to headless bodies and body parts were significantly greater than to faces, cars, and mixed objects (all six $p < 0.02$), while faces no longer produced a higher response than to cars or mixed objects (both $p > 0.4$). These results demonstrate selective responses (above control objects) only for faces in the FFA* and only for bodies in the FBA*.

To assess the relative locations of the FFA* and FBA* ROIs, we calculated the center of mass (COM) locations for these two ROIs in each individual high resolution participant by taking the average of the in-slice row and column numbers in the matrix, as well as the slice number, across all voxels in each given ROI. As can be seen in the surface plots of Figure 1, we found a significant difference between the location of the right FFA* and FBA* COMs along the medial-lateral axis, with FFA* medial to FBA* ($p > 0.005$). The average distance between these COMs was 2.2 voxels, or 3.1 mm. Because our slices were oriented perpendicular to the calcarine sulcus and were therefore not aligned precisely from one participant to the next, it was difficult to accurately compare the ROI locations in the anterior-posterior and dorsal-ventral dimensions between participants.

Finally, we addressed the question of whether exclusive selectivity on the fusiform gyrus could be demonstrated at standard resolution if the effects of partial voluming were minimized. We did this by selecting a single voxel in the FFA of each of our standard resolution participants that most reliably demonstrated a greater response to faces than to objects as measured by the p-value of this contrast in data from our blocked localizer runs. Crucially, the selection of these voxels was independent of their selectivity for bodies and the pattern of response in these peak voxels was evaluated using our independent event-related data set. We found that the peak FFA voxels averaged across the standard resolution

participants demonstrated a high response to faces (1.08 percent signal change) with a response to headless bodies (0.43) and body parts (0.52) that was no greater than to cars (0.41), $p > 0.7$ and $p > 0.3$, respectively. This finding demonstrates that exclusive selectivity on the fusiform gyrus can also be observed at standard resolution in circumstances for which the effects of partial voluming are minimized.

Discussion

In this study we used high resolution scanning techniques that uncovered a clear and striking dissociation between face and body selectivities on the fusiform gyrus. At a standard fMRI scanning resolution, face and body selectivity overlapped considerably, with substantial responses to body stimuli in regions identified as face selective and vice versa (see Figure 2a,b), as reported by Peelen and Downing (2005). However, at higher resolution the observed selectivities become stronger, with responses to body stimuli in the FFA only slightly higher than to control objects (see Figure 2c). Finally, when new ROIs that omit regions of overlapping selectivity for faces and bodies were created, we found one region (the FFA*) that is selectively responsive only to faces, not bodies, and another region (the FBA*) that is selectively responsive only to bodies, not faces (see Figure 2e,f). These findings support our hypothesis that the dual selectivity of the FFA for both faces and bodies observed at standard resolution results from the pooling of responses from two adjacent but distinct cortical regions, one selective for only faces, and another selective for only bodies.

In keeping with prior evidence from intracranial recordings (Allison et al., 1994), stimulation studies (Mundel et al., 2003;Puce et al., 1999) and neuropsychological studies (see Wada and Yamamoto, 2001), our finding that some regions in the ventral visual pathway are apparently strongly selective for a single class of visual stimuli would

seem to argue against the idea that all regions in the ventral visual pathway participate in the representation of each object (Haxby et al., 2001). However, two caveats must be mentioned here. First, the fact that strong and separate cortical selectivities exist for faces and bodies in the ventral visual pathway does not mean that the same will be found for all stimulus categories. Indeed, current evidence suggests that the cortical selectivities for faces and bodies may be unusual cases, contrasting with the more distributed and overlapping responses to multiple object categories in other cortical regions such as the lateral occipital complex (Malach et al., 1995). Second, although the FFA* and FBA* are uniquely selective for faces and bodies, respectively, compared to control stimuli (mixed objects and cars), both of these regions produce positive responses to non-preferred stimuli compared to a fixation baseline. The role of these non-preferred responses in the coding of objects is an important open question that is now being tested using a variety of neuroimaging methods (Haxby et al., 2001;Spiridon and Kanwisher, 2002;Grill-Spector et al., 2004). Currently, the strongest evidence that face processing regions do not play an important role in the recognition of non-face objects comes from studies of neurological patients with very selective deficits in face recognition but not in general object recognition (Wada and Yamamoto, 2001).

The results of our study also have methodological relevance in highlighting the importance of scanning resolution when investigating functional segregation in the cortex (see also Beauchamp et al., 2004). Regions selectively responsive to faces and bodies that were clearly dissociable at high resolution were not dissociable at standard resolution. How can we determine how much resolution is enough to make such a distinction for any given study? The answer will depend on the grain of the cortical organization under investigation, with the response profiles of relatively large cortical regions such as the PPA less dependent on scanning resolution than smaller regions that may only be detected at high resolution.

These considerations lead to an important asymmetry in the conclusions that can be drawn from fMRI studies: when clear functional dissociations are demonstrated between adjacent cortical regions, such results can not be overturned by future studies at higher resolution, whereas any failure to find a functional dissociation (e.g. Shuman and Kanwisher, 2004) will always be contingent on the outcome of future studies at higher resolution. For example, the current results suggest that it will be necessary to revisit prior claims that the FFA may be responsive not only to faces but also to biological motion (Grossman and Blake, 2002), animations implying intentional agency (Schultz et al., 2003), visual expertise (Gauthier et al., 2000;Xu, 2005;Gauthier et al., 1999), and animals (Chao et al., 1999). Indeed, it seems possible that many of these activations previously attributed to the FFA arise not from the FFA* but from the FBA* or another adjacent but distinct cortical region.

Beyond their methodological implications, the present results also raise a host of questions for future research. What is the function of the FBA, and how does it differ from that of the EBA? Given that lesions affecting the FFA are likely to also affect the FBA, do acquired prosopagnosic patients show deficits in body perception, and if so in what aspects of body perception? More generally, why do face and body selectivities land nearby in the cortex, not only in the fusiform gyrus, but also in lateral temporal cortex in both humans and monkeys (Tsao et al., 2003;Pinsk et al., 2005)?

Another question raised by our findings concerns the nature of the overlap region between the face and body selectivities on the fusiform gyrus. One prior study has suggested that the area of overlap between two functional regions might play a role in the integration of information processed by the neighboring regions (Beauchamp et al., 2004). While this explanation is possible both for their case and ours, another possibility is that the observed dual selectivity reflects distinct but interleaved neural populations that perform no

integrative function. Distinguishing between these hypotheses will require other methods such as fMRI adaptation.

In summary, our findings demonstrate two adjacent regions in the fusiform gyrus, one selectively responsive to bodies but not faces, and an adjacent region selectively responsive to faces but not bodies. The striking dissociation in the category selectivity of these regions was not clear when standard scanning methods were used (Peelen and Downing, 2005), underlining the importance of resolution for investigations of functional specificity of the cortex.

References

Allison T, McCarthy G, Nobre A, Puce A, Belger A (1994) Human extrastriate cortex and the perception of faces, words, numbers and colors. Cereb Cortex 4:544-54.

Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004) Unraveling multisensory integration: patchy organization within human STS multisensory cortex. Nat Neurosci 7:1190-2.

Chao LL, Martin A, Haxby JV (1999) Are face-responsive regions selective only for faces? Neuroreport 10:2945-50.

Cox D, Meyers E, Sinha P (2004) Contextually evoked object-specific responses in human visual cortex. Science 304:115-7.

Cox R, Jesmanowicz A (1999) Real-time 3D image registration for functional MRI. Magn Reson Med 42:1014-1018.

Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A cortical area selective for visual processing of the human body. Science 293:2470-3.

Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC (1999) Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. Nat Neurosci 2:568-73.

Gauthier I, Skudlarski P, Gore JC, Anderson AW (2000) Expertise for cars and birds recruits brain areas involved in face recognition. Nat Neurosci 3:191-7.

Grill-Spector K, Knouf N, Kanwisher N (2004) Fusiform face area subserves face perception, not generic within-category identification. Nat Neurosci 7:555-62.

Grossman ED, Blake R (2002) Brain areas active during visual perception of biological motion. Neuron 35:1167-75.

Hadjikhani N, de Gelder B (2003) Seeing fearful body expressions activated the fusiform cortex and amygdala. Curr Biol 13:2201-5.

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 293:2425-30.

Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 17:4302-11.

Kanwisher N, Stanley D, Harris A (1999) The fusiform face area is selective for faces not animals. Neuroreport 10:183-7.

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 92:8135-9.

Mundel T, Milton JG, Dimitrov A, Wilson HW, Pelizzari C, Uftring S, Torres I, Erickson RK, Spire JP, Towel VL (2003) Transient inability to distinguish between faces: electrophysiologic studies. J Clin Neurophysiol 20:102-10.

Peelen MV, Downing PE (2005) Selectivity for the human body in the fusiform gyrus. J Neurophysiol 93:603-8.

Pinsk MA, DeSimone K, Moor T, Gross CG, Kastner S (2005) Representations of faces and body parts in macaque temporal cortex: A functional MRI study. Proc Natl Acad Sci 102:6996-7001.

Puce A, Allison T, McCarthy G (1999) Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. Cereb Cortex 9:445-58.

Schultz RT, Grelotti DJ, Klin A, Kleinman J, Van der Gaag C, Marois R, Skudlarski P (2003) The role of the fusiform face area in social cognition: implications for the pathobiology of autism. Philos Trans R Soc Lond B Biol Sci 358:415-27.

Shuman M, Kanwisher N (2004) Numerical magnitude in the human parietal lobe: tests of representational generality and domain specificity. Neuron 44:557-69.

Spiridon M, Kanwisher N (2002). How distributed is visual category information in human occipitotemporal cortex? An fMRI study. Neuron 35:1157-65.

Spiridon M, Fischl B, Kanwisher N (2006) Location and spatial profile of category-specific regions in human extrastriate cortex. Hum Brain Mapp 27:77-89.

Tsao DY, Freiwald WA, Knutsen TA, Mandeville JB, Tootell RB (2003) Faces and objects in macaque cerebral cortex. Nat Neurosci 6:989-95.

Wada Y, Yamamoto T (2001) Selective impairment of facial recognition due to a haematoma restricted to the right fusiform and lateral occipital region. J Neurol Neurosurg Psychiatry 71:254-7.

Xu Y (2005) Revisiting the role of the fusiform face area in visual expertise. Cereb Cortex 15:1234-42.

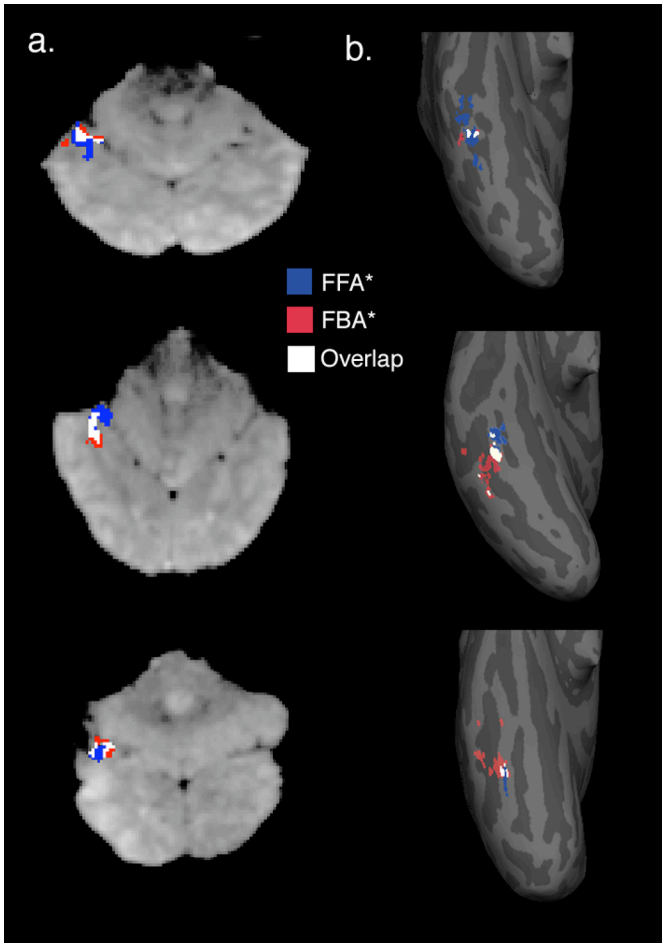

Figure 1: Examples of Face and Body ROIs at High Resolution.

a: Examples from three participants of FFA* (blue) and FBA*(red) ROIs, as well as the overlap (white) between FFA (defined by faces<objects) and FBA (defined by bodies > objects). ROIs are shown on functional image slices from three participants. The slices are left-right reversed, with posterior regions shown at the bottom of each image and the cerebellum at the top. b: The same regions in the same three participants mapped to each participant's inflated cortical surface. The view shown here is of the ventral temporal surface of the posterior portion of the right hemisphere, with the lower tip of each inflated hemisphere representing the occipital pole. The FFA* (shown in blue), FBA* (shown in red), and the overlap (white) show considerable variation in their sizes and relative locations on the cortex.

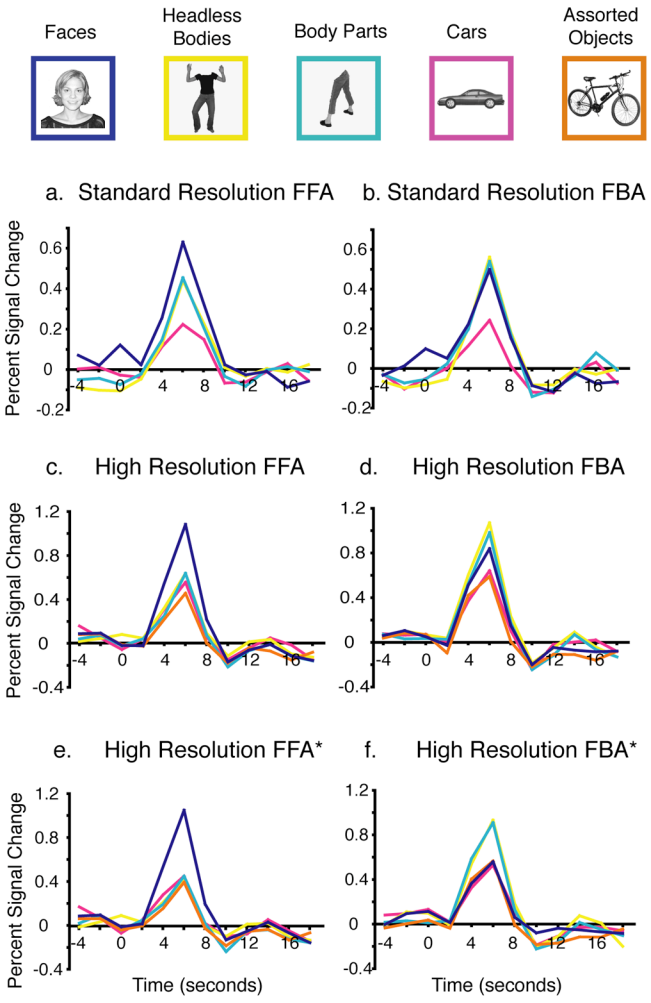

Figure 2: Time Courses of the Hemodynamic Response in Regions of Interest.

Examples of each of the stimulus conditions and their color code are shown across the top. Time courses of the hemodynamic response for each stimulus condition for the event-related runs averaged across participants are shown below for the FFA (2a) and FBA (2b) from Experiment 1 at standard resolution, as well as for the FFA (2c), FBA (2d), FFA* (2e), and FBA* (2f) in Experiment 2 at high resolution.

3. Is Spatiotemporal Association an Organizing Dimension of Ventral Visual Cortex?

Introduction

What determines the systematic spatial organization of stimulus selectivities across the ventral visual pathway? In this chapter we consider the hypothesis that pairs of stimuli that co-occur in daily experience come to activate the same or nearby regions of cortex. This hypothesis is based on several considerations. First, our visual experience is statistically structured such that some pairs of objects are seen together or in rapid succession much more frequently than are other pairs. Second, humans and other primates are sensitive to this statistical structure of experience, and it affects their behavioral performance (Bar and Ullman, 1996; Chun and Jiang, 1999; Bar, 2004; Oliva and Torralba, 2007) and the responses of single units in IT cortex (Sakai & Miyashita, 1991). Third, the case of cortical responses to faces and bodies is suggestive: virtually every time we see a face it is attached to a body and vice versa, and, consistent with our hypothesis, selectivities for faces and

bodies occupy adjacent cortical areas (see Chapter 2). Here we test the hypothesis that objects that co-occur in visual experience develop cortical selectivities that are in the same or nearby locations.

The idea that object representations can become more similar through spatiotemporal association is not new. Behavioral evidence shows that under some circumstances visual stimuli that are temporally associated over a short training period come to appear more similar to each other (Wallis and Bulthoff, 2001; Cox et al., 2005). Further, electrophysiology studies of single neurons in inferotemporal and perirhinal cortex have found a disproportionately high number of shape-selective neurons that respond to both shapes in an associated pair (Sakai & Miyashita, 1991; Erickson & Desimone, 1999), and one study found in a subset of neurons that the responses to the two stimuli in a pair became more similar within just one testing session (Messinger et al, 2001).

Although these physiological studies demonstrate that neuronal responses are affected by temporal associations on a time scale of hours or days, these effects were found in only a subset of the neurons tested and were small enough in magnitude that they might not be detectable with fMRI. Moreover, the associations learned in these studies were under reward-based circumstances and were the sole focus of the animal's attention throughout the testing session. It is not clear that these changes in neuronal response would take place in humans under normal conditions of incidental exposure to the objects and would be widespread enough to be seen with fMRI. We therefore decided to test our hypothesis of similar cortical responses for co-occurring stimuli using object categories that people have had extensive experience with over their lifetime.

For this study we chose a category of object that is strongly spatiotemporally associated with faces: eye glasses. Glasses occur frequently in our visual world and, crucially, they are

usually located on a face. Moreover, glasses and faces can both be seen simultaneously at the same retinal location, yet they are semantically unrelated (unlike bodies and faces, which are both semantically linked to people). For these reasons, glasses and faces make a perfect test case for our hypothesis that spatiotemporal associations guide the large-scale organization of the ventral visual cortex (VVC). Thus, we scanned subjects while they viewed a variety of stimulus classes in order to test the prediction of our hypothesis that faces and glasses will produce similar or nearby cortical responses. To do so, we used one half of the data to functionally identify face-selective and body-selective cortical regions, specifically the fusiform face area or FFA (Kanwisher et al., 1997; McCarthy et al., 1997) and the extrastriate body area or EBA (Downing et al., 2001). The other half of the data was used to quantify the magnitude of response in these regions. Our hypothesis predicts a higher response to glasses than to other control objects in or around the FFA, but not the EBA. This study is comprised of two experiments; the first provides an initial test of our hypothesis and the second provides a further test with additional stimulus conditions.

Materials and Methods

Stimuli and design

Experiment 1. Each subject was scanned using fMRI on ten experimental runs, as well as another four runs for an unrelated experiment that will not be described here. Each experimental run consisted of three 16 sec fixation blocks and ten 16 sec stimulus blocks (two blocks for each of five stimulus conditions.) Four of these conditions were: human faces, assorted human body parts, glasses, and assorted everyday objects. See Figure 1 for examples of stimuli. The fifth condition differed across participants and will not be discussed here. Conditions were presented in a palindromic order within each run, and the serial position of each condition was counterbalanced within participants across runs.

For each block, twenty images from a single stimulus class were foveally presented (300 ms per image, with a 500 ms ISI). Participants performed a 1-back task for which they were instructed to make a key-press whenever images were consecutively repeated, which happened twice per block.

The stimuli used in this experiment were drawn from two non-overlapping stimulus sets, each containing thirty images per condition. Each run drew from only one of these stimulus sets and the runs in a single scan session were evenly split such that five runs presented images from one stimulus set and five runs showed images from the other stimulus set. Runs from the two stimulus sets were alternated throughout the course of the scan session. The set of runs from one stimulus set ("defining runs") were subsequently used to define the regions-of-interest (ROIs), while the other set of runs ("evaluating runs") were used to measure the magnitude of response to the experimental conditions in those ROIs.

Experiment 2. The design of Experiment 2 was similar to that of Experiment 1. Participants performed a 1-back task while viewing 12-16 runs comprised of 16 sec blocks of foveally presented stimuli. Unlike Experiment 1, this experiment contained eight conditions: faces, body parts, glasses, bicycles, shoes, hats, vases, and grid-scrambled versions of the vases. Scrambled vase stimuli were made by superimposing a grid over the vases and randomly rearranging the component squares (Kourtzi and Kanwisher, 2000). See Figure 1 for examples of stimuli. Each run contained one block per condition, and the serial position of each condition was counterbalanced across runs for each participant. As in Experiment 1, each run contained images from one of two non-overlapping sets of stimuli. For each participant, half of the runs drew from one stimulus set and were subsequently used as defining runs, while the other half drew from the other stimulus set and were used as evaluating runs.

Functional imaging

Experiment 1. Eight participants were scanned on a 3.0 T Siemens (Erlangen, Germany) TimTrio scanner at the Martinos Center for Biomedical Imaging (Charlestown, MA). Images were acquired with a gradient echo single-shot echo planar imaging sequence with a repetition time of 2 sec, flip angle 90°, and echo time 33 ms. Twenty-two slices of thickness 2.0 mm were manually oriented roughly perpendicular to the calcarine sulcus. Voxel dimensions were 1.4 x 1.4 x 2.0 mm with a 0.4 mm interslice gap. High-resolution MPRAGE anatomical images were also acquired for each participant in the same scan session. The data from two participants were excluded from further analysis due to excessive head motion.

Experiment 2. Five participants were scanned on a 3.0 T Siemens (Erlangen, Germany) TimTrio scanner at the Martinos Imaging Center at the McGovern Institute (Cambridge, MA). Images were acquired with a gradient echo single-shot echo planar imaging sequence with a repetition time of 2 sec, flip angle 90°, and echo time 33 ms. Twenty to twenty-two slices of thickness 2 mm and a 0.4 mm interslice gap were manually oriented roughly perpendicular to the calcarine sulcus. In-plane voxel dimensions were 1.4 x 1.4 mm. High-resolution MPRAGE anatomical images were also acquired for each participant in the same scan session. The data from one participant were excluded from further analysis due to excessive head motion.

Experiments 1 and 2: Data Analysis. Data analysis was performed using Freesurfer and FS-FAST software (http://surfer.nmr.mgh.harvard.edu/). The acquired images were motion corrected (Cox and Jesmanowicz, 1999) prior to statistical analysis, and smoothed with a full width half maximum Gaussian kernel of 3 mm for data from the defining runs. Data from the evaluating runs were not smoothed.

For both Experiment 1 and Experiment 2, ROIs were individually defined for each participant using the smoothed data from the defining runs as a set of contiguous voxels with contrast difference of $p < 0.0001$ uncorrected. For Experiment 1, the face-selective FFA (Kanwisher et al., 1997) was defined using a faces > assorted objects contrast and the body-selective areas EBA (Downing et al., 2001) and FBA (Peelen and Downing, 2005; Schwarzlose et al., 2005) were defined with a contrast of body parts > assorted objects. For Experiment 2, the FFA was defined with a faces > vases contrast and the FBA was defined with a body parts > vases contrast. Since the FFA and FBA are typically adjacent and overlap, we excluded any overlapping voxels to create the functionally dissociated regions FFA* and FBA*, as described by Schwarzlose et al (Schwarzlose et al., 2005). We then used fROI (http://froi.sourceforge.net/) to extract the mean response magnitude across voxels in each ROI to the various conditions in the evaluating runs. The data and stimuli used to define the ROIs were entirely separate from those used to calculate the response magnitudes to each condition in each ROI.

In a subsequent analysis, new ROIs were created that encircle the FFA*. To make these FFA* ring ROIs, voxels adjacent to the previously defined FFA* ROIs were manually selected. One set of ring ROIs ("narrow FFA* rings") were made by selecting the set of single voxels adjacent to the most peripheral voxels in the FFA* ROIs. The other set of ring ROIs ("thick FFA* rings") were composed of all voxels in the narrow rings, as well as all individual voxels directly adjacent to and outside of that ring. As a result, the width of the thick ring was 2 voxels (2.8 – 4.0 mm), while the width of the narrow ring was one voxel (1.4 – 2.0 mm). Both sets of rings were selected in the native space (slices). Any voxels located near a susceptibility artifact were excluded from the ring ROIs. Response magnitudes in the rings were computed based on data from the evaluating runs only.

The error bars shown in Figures 2-5 of this chapter depict one standard error of the mean in each direction and reflect between-subject variability for each stimulus condition. However, all statistical analyses in this study were 2-tailed, paired t-tests conducted within subjects, therefore the error bars do not reflect the statistical significance of contrasts between stimulus conditions.

Experiment 1 Results

ROI localization

ROIs were identified using data from one half of the runs (the "defining runs") and were defined as described in the Methods section. FFA ROIs were found in the right hemispheres of all participants and in the left hemisphere of four participants, while EBA ROIs were found bilaterally in all participants. However among the six participants, only six FBA ROIs were identified (three right-hemisphere ROIs and three left-hemisphere ROIs). This may be due to the fact that individual body parts (hands, arms, legs, and feet) were used as body stimuli in this study, while the FBA is typically identified using a contrast of headless bodies > objects. The FBA has been shown to produce a greater response to whole bodies and large sections of bodies, such as the ones used in Schwarzlose et al. (Schwarzlose et al., 2005), than to individual body parts (Taylor et al., 2007). The FFA* ROI was made from the standard-definition FFA by excluding any voxels that overlapped with the FBA if one was found (Schwarzlose et al., 2005).

Mean response magnitude in FFA

To test the hypothesis that ventral visual cortex is organized based on principles of spatiotemporal association, we first analyzed the response of FFA* to the experimental conditions. The prediction according to this hypothesis is that the FFA* would produce a

greater response to glasses than to other object categories, since glasses and faces consistently co-occur in visual experience.

Using the data from the evaluating runs, we measured the magnitude of response in the FFA* to the four conditions that were presented to all participants (faces, body parts, glasses, and assorted objects). Results are shown in Figure 2. Bilaterally in the FFA*, the response to faces was significantly greater than to body parts, assorted objects, and glasses (all three, $p < 5x10^{-6}$), replicating the face selectivity used to define this region. The response to body parts in this region was not significantly greater than to objects ($p = 0.10$). The critical question for this analysis is whether glasses would elicit greater activation of the FFA* than assorted objects. We found that this was not the case; the response to glasses did not differ significantly from the response to body parts and objects ($p = 0.40$ and 0.45, respectively).

While these results appear to argue against a hypothesis of spatiotemporal association as a dimension of cortical organization, they are based upon a bilateral analysis of the FFA*. However, evidence exists that the right FFA may be more category-selective than the left FFA (Grill-Spector et al., 2004). It is possible that the inclusion of the left FFA* in this analysis washed out any elevated response to glasses that exists in the right FFA*. Therefore, we repeated these analyses on data from the right FFA* ROIs only. Despite the smaller sample size, this analysis also replicated the face selectivity of the FFA* compared to all non-face conditions (all three, $p < 0.0005$). The critical question is whether glasses would elicit a higher response than assorted objects in the right FFA*. Here again we found that they did not. The right FFA* responses for objects and glasses were nearly identical (see Figure 2), with no significant difference between them ($p = 0.83$). Although these results are suggestive of a lack of strong selectivity for glasses in the FFA*, the failure

to find a statistical difference between responses to glasses and objects does not prove the absence of an effect, and studies with larger sample sizes would be necessary to determine more definitively whether a modest selectivity for glasses might exist in FFA*.

Mean response magnitude in EBA

An alternate explanation for the failure to find strong selectivity for glasses in the FFA* is that the glasses stimuli used in this experiment are not adequately life-like or salient to capture attention and activate areas of object-selective cortex. To test whether the glasses stimuli do effectively activate areas of the VVC, we examined the response to these stimuli in the body-selective area EBA. Mean response magnitudes in EBA are shown in Figure 3. In this area, the response to glasses was significantly greater than the response to both assorted object and face conditions (both, $p < 0.001$). This result suggests that the low response to glasses in FFA* is not due to a failure of the stimuli to draw attention and drive neural responses. The heightened response to glasses over faces and objects in the body area is unexpected and does not directly fit with a hypothesis of temporal association (as glasses are paired more directly with faces than with bodies). See Chapter 4 for a study of this effect and a discussion of possible explanations.

Mean response magnitude in FFA rings*

Our failure to find selectivity for glasses in the FFA* does not support the hypothesis that the organization of VVC is guided by spatiotemporal association through visual experience. However, such association between objects could cause the cortical selectivities for these objects to develop adjacent to one another, rather than in the same cortical locus. Indeed, we found in Chapter 2 that bodies do not selectively activate the face area per se, but rather cortex adjacent to FFA*. Therefore, it is possible that glasses activate cortex adjacent to the FFA in the same way bodies do. To test this, we performed an analysis

on rings that encircle the FFA*. The magnitude of response to glasses and other object categories could then be tested in these 'ring' ROIs. If selectivity for glasses is located adjacent to the FFA*, then we should find a higher response to glasses than assorted objects in these rings. All individual voxels located adjacent to the most peripheral voxels in the FFA* ROIs were selected to form one set of ring ROIs ("narrow FFA* rings"). The second set of ring ROIs ("thick FFA* rings") were composed of the voxels in the narrow rings, as well as all individual voxels adjacent to and outside of that ring. Response magnitudes in the FFA* rings were computed based on the data from the evaluating runs only.

Results from the ring analyses are shown in Figure 4. A bilateral analysis of the narrow FFA* rings revealed a greater response to faces than to assorted objects ($p < 0.0005$), demonstrating that some degree of face selectivity persists beyond the boundaries of the standard-definition FFA*. This result is generally consistent with those of Spiridon et al. (Spiridon et al., 2006), which found that face selectivity persists beyond the FFA but drops off steeply with distance across the cortical surface from the FFA boundary, so that it is largely gone at 2 mm from the border of ROI. The narrow ring ROIs also produced a significantly greater response to body parts than assorted objects ($p < 0.05$), which is consistent with the fact that body selectivity is found adjacent to the FFA* (Schwarzlose et al., 2005). Critically, we found a greater response in the narrow rings to glasses than assorted objects, consistent with the spatiotemporal association hypothesis, however this effect was only marginally significant ($p = 0.07$).

The narrow rings may encompass too few voxels to pick up a localized selectivity for glasses. Since face selectivity persists at least 2 mm from the FFA boundary and our narrow rings were 1.4 – 2.0 mm in width, these rings would not include all of surrounding face-selective cortex and might miss adjacent or overlapping activation for glasses. Therefore,

we repeated the analysis on the thick FFA* rings (with a width of 2.8 –4.0 mm), to maximize the chance of finding any adjacent selectivity for glasses. Analysis of the thick FFA* rings showed the same profile of responses as the narrow rings, specifically greater response to both faces and body parts than to assorted objects (both, $p < 0.01$). Again, the response to glasses was greater than to assorted objects in the thick rings, however in this case the difference was significant ($p < 0.05$). These results suggest that, as with bodies, there may be a selectivity for glasses that lies adjacent to the FFA and possibly overlaps with the moderate face selectivity beyond its standard-definition borders. However, our measures for selectivity in this study are all judged against a single object baseline: the assorted object condition. Our findings would be consistent with a heightened response in the rings to glasses than other object categories; however, it would also be consistent with a lower response to the assorted object condition than to other object categories. Additional object conditions are necessary to test whether glasses actually do elicit a greater response around the FFA* than other object categories that do not co-occur with faces. Experiment 2 was designed to address this question.

Experiment 2 Results

In Experiment 1 we found a response to glasses that was higher than to assorted objects in the FFA* rings. Given the limited stimulus set used in Experiment 1, it is impossible to tell if this small effect is specific to glasses (possibly due to their temporal association with faces), or if it may be due to some other aspect of these specific stimuli. For example, the assorted object condition contains stimuli from many different object categories with a large variety of shapes, while the glasses stimuli all belong to the same category and assume a more limited range of shapes. In order to test whether glasses activate cortex around the FFA* more than other object categories, we scanned a new set of participants with several

object conditions that were not used in Experiment 1. Each of the new conditions contained stimuli from only one object category. They were: bicycles, hats, shoes, vases, and grid-scrambled versions of vases. See Figure 1 for examples of these stimuli. As in Experiment 1, runs were composed of blocks of stimuli from one of two non-overlapping stimulus sets. The runs drawing from one set of these stimuli ("defining runs") were subsequently used to define the ROIs, while independent data from the second set of runs ("evaluating runs") were used as assess the response magnitude to the stimulus conditions in those ROIs.

ROI localization

ROIs were identified based on the data from the defining runs and localized using the same criteria described for Experiment 1, except that in this case the face- and body-selective regions were defined by the contrasts faces > vases and body parts > vases, respectively. Vases were chosen as a baseline because they are not associated with people and are therefore least likely to co-occur with faces in visual experience. FFA ROIs were found bilaterally in all but one participant, who failed to show an FFA in the left hemisphere. FBA ROIs were found bilaterally in all four participants. As in Experiment 1, The FFA* ROI was made from the FFA by excluding any voxels overlapping with the FBA (Schwarzlose et al., 2005).

*Mean response magnitude in FFA**

The results of Experiment 1 failed to show a significantly greater response to glasses than assorted objects within the FFA* ROI. Here we test whether these findings replicate when the FFA* response to glasses is compared to several other object conditions.

Due to the limited number of participants in this experiment, ROIs were analyzed bilaterally to maximize the possibility of detecting significant differences between conditions.

See Figure 5 for the response magnitude in the FFA* ROIs. The FFA* response to faces was significantly higher than to all non-face conditions (all seven, $p < 0.005$), replicating the face selectivity of this region. However, the response to glasses was not significantly different from any other non-face conditions (five, $p > 0.20$) except scrambled vases ($p < 0.01$). In fact, all conditions elicited a greater response than scrambled vases (six, $p > 0.05$) except bicycles ($p = 0.09$). The failure to find a significant difference between glasses and other objects in the FFA* replicates the findings from Experiment 1, however it does not prove that this region contains no selectivity for glasses, particularly given the limited number of participants. Still, it does argue against a strong selectivity for glasses, since the FFA* response to glasses, body parts, hats, shoes, and vases were nearly identical to one another.

Mean response magnitude in FFA rings*

In Experiment 1, we found a higher response to glasses than assorted objects in the FFA* rings, rather than the FFA* itself. Here we tested whether the FFA* rings would produce a greater response to glasses than to the additional object categories. The results are shown in Figure 5. The response to faces was significantly greater than to all other object conditions in the narrow FFA* ring (all seven, $p < 0.0005$), and to all conditions except body parts in the thick FFA* ring (body parts, $p = 0.20$; all others, $p < 0.005$). In the narrow ring, body parts elicited a significantly greater response than all non-face conditions (five, $p < 0.05$) except hats ($p = 0.09$). In the thick FFA* ring, the response to body parts was significantly greater than to shoes and scrambled vases, and marginally greater than to glasses, bicycles, hats, and vases ($p = 0.06$, 0.05, 0.08, and 0.05, respectively). These results demonstrate that the rings are sensitive both to the face selectivity around the FFA* and the body selectivity adjacent to and overlapping with the standard FFA. However,

the key question is whether we find a greater response to glasses than to other non-face, non-body conditions. Indeed, glasses do not elicit a greater response than any of the other intact object conditions in the narrow ring (all four, $p \geq 0.37$) or in the thick ring (all four, $p \geq 0.14$). As discussed in Experiment 1, the failure to find a significant bias for glasses does not preclude the possibility that a selectivity exists in these regions, but that these analyses lack sufficient power to find the difference significant. This presents a particular problem in Experiment 2, which has a relatively small sample size ($n = 7$ ROIs). Nonetheless, the failure to reach significance demonstrates that any selectivity for glasses in these rings would be small in magnitude in comparison with those for faces and bodies. Moreover, an examination of the means in Figure 5 demonstrates that the mean responses to glasses, bicycles, shoes, hats, and vases are nearly identical in both the narrow and thick FFA* rings. If these conditions were placed in rank serial order according to the responses they elicit, glasses would rank second lowest out of the five conditions in the narrow FFA* ring and third lowest out of five conditions in the thick ring. These results further argue against a strong selectivity for glasses in or around the FFA*.

Discussion

In this study, we found that images of glasses do not activate either the FFA* or the cortex around FFA* more than images of other objects. These results suggest that spatiotemporal association between objects is not a large-scale organizing principle of the VVC, and therefore argue that the cortical adjacency of body selectivity and face selectivity is not due to the co-occurrence of these object categories.

Importantly, the results of this study cannot address the more basic question of what happens at the level of individual neurons. It may be that some neurons dually code both faces and glasses, but that this effect is small enough or takes place in few enough neurons

to be washed out at the voxel level. It is also possible that the object pairing tested here is not consistent enough to elicit an effect, since one can see faces without glasses and glasses that are not on faces. However, virtually no object pairings are perfectly consistent, therefore if spatiotemporal association were to serve as a major organizing dimension, it would have to do so with objects that are not exclusively seen together.

Another consideration when interpreting these results is that configuration and orientation may be important for the representation of spatiotemporally associated objects (Green and Hummel, 2006). Specifically, it may be that the FFA* is selective for glasses that are upright and have their arms unfolded, as one sees them when they are sitting on a face. The images included in the glasses stimuli were quite mixed in their orientation and configuration (see Figure 1 for examples). This was done to prevent the implication of a face, as prior work has shown that contextual cuing of a face may elicit FFA activation (Cox et al., 2004). Further studies using upright and unfolded glasses would be necessary to determine whether those stimuli will replicate the findings from these experiments.

Finally, analysis of the body-selective EBA in Experiment 1 revealed a significantly greater response to glasses than objects in that area. This finding cannot easily be explained based on the principle of spatiotemporal association; while glasses tend to co-occur with faces and faces tend to co-occur with bodies, the link between bodies and glasses is weaker than between bodies and faces. Therefore these findings are unexpected and warrant further study. The next chapter is dedicated to this investigation.

References

Bar M (2004) Visual objects in context. Nat Rev Neurosci 5:617-629.

Bar M, Ullman S (1996) Spatial context in recognition. Perception 25:343-352.

Chun MM, Jiang Y (1999) Top-down attentional guidance based on implicit learning of visual covariation. Psychol Sci 10:360-365.

Cox D, Meyers E, Sinha P (2004) Contextually evoked object-specific responses in human visual cortex. Science 304:115-117.

Cox DD, Meier P, Oertelt N, DiCarlo JJ (2005) 'Breaking' position-invariant object recognition. Nat Neurosci 8:1145-1147.

Cox R, Jesmanowicz A (1999) Real-time 3D image registration for functional MRI. Magn Reson Med 42:1014-1018.

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

Green C, Hummel JE (2006) Familiar interacting object pairs are perceptually grouped. J Exp Psychol Hum Percept Perform 32:1107-1119.

Grill-Spector K, Knouf N, Kanwisher N (2004) The fusiform face area subserves face perception, not generic within-category identification. Nat Neurosci 7:555-562.

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

Kourtzi Z, Kanwisher N (2000) Cortical regions involved in perceiving object shape. J Neurosci 20:3310-3318.

McCarthy G, Puce A, Gore JC, Allison T (1997) Face-specific processing in the human fusiform gyrus. J Cogn Neurosci 9:605-610.

Oliva A, Torralba A (2007) The role of context in object recognition. Trends Cogn Sci 11:520-527.

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

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

Spiridon M, Fischl B, Kanwisher N (2006) Location and spatial profile of category-specific regions in human extrastriate cortex. Hum Brain Mapp 27:77-89.

Taylor JC, Wiggett AJ, Downing PE (2007) Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. J Neurophysiol 98:1626-1633.

Wallis G, Bulthoff HH (2001) Effects of temporal association on recognition memory. Proc Natl Acad Sci U S A 98:4800-4804.

Figure 1: Examples of stimuli.

Five representative examples are shown for each of the stimulus conditions included in Experiment 1 (rows 1-4) and Experiment 2 (rows 2-9). By row, the conditions are: assorted objects, faces, assorted body parts, glasses, bicycles, shoes, hats, vases, and scrambled vases. In the actual experiment, stimuli were presented one at a time.

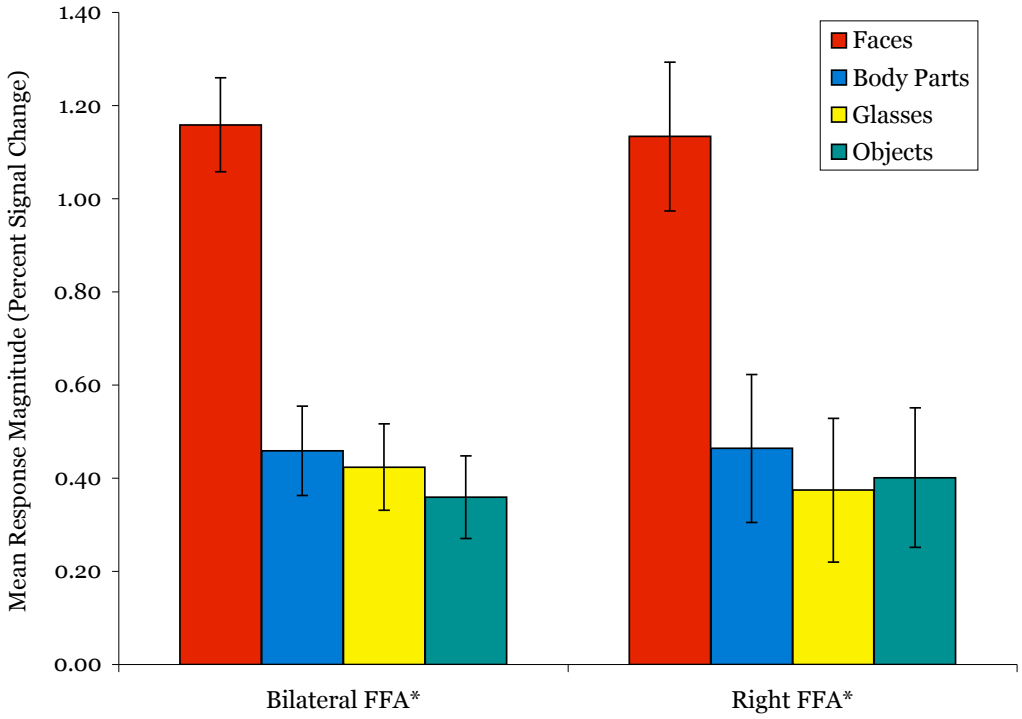

Figure 2: Category selectivity in FFA*.

The average magnitude of response to each of the four stimulus categories in Experiment 1 are shown here for the FFA* ROIs in both hemisphere (left) and for FFA* in the right hemisphere only (right). This figure demonstrates that the FFA* is not selective for glasses.

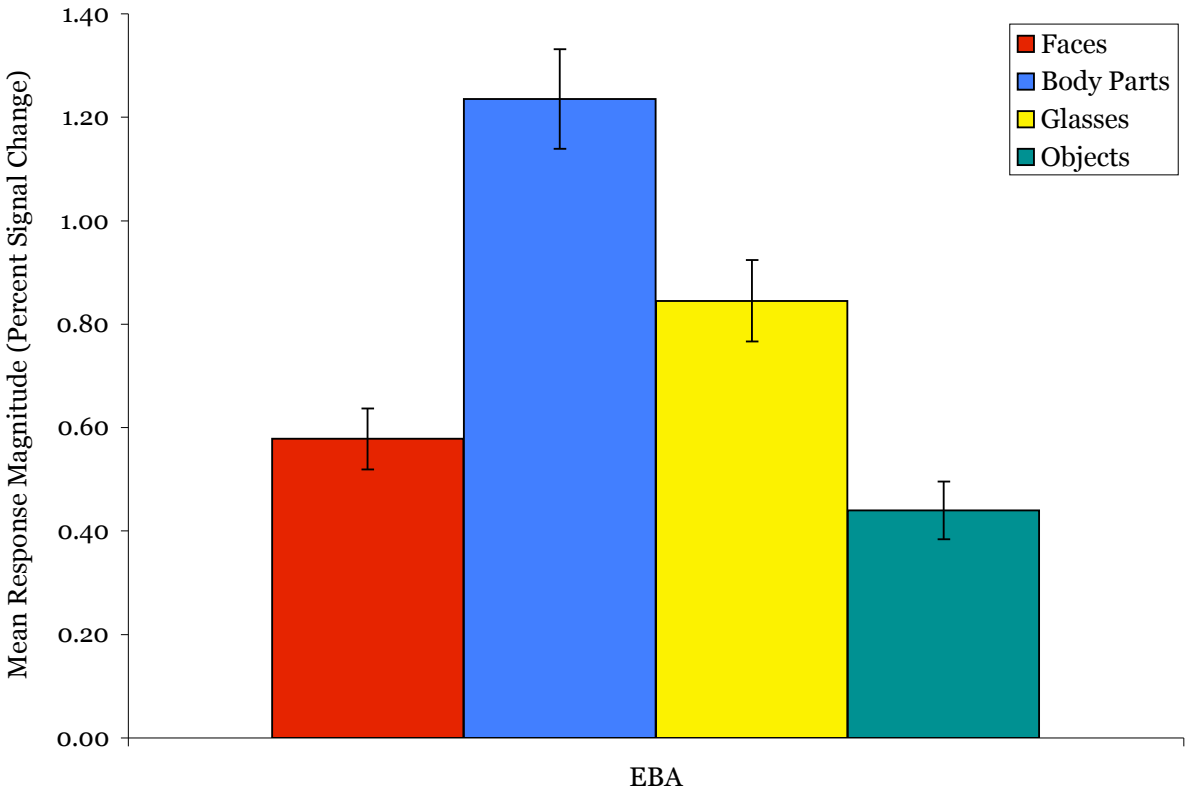

Figure 3: Category selectivity in the EBA.

The responses in the EBA to each of the stimulus categories in Experiment 1 are depicted here. Unlike the FFA*, the EBA shows a clear selectivity for glasses over assorted objects.

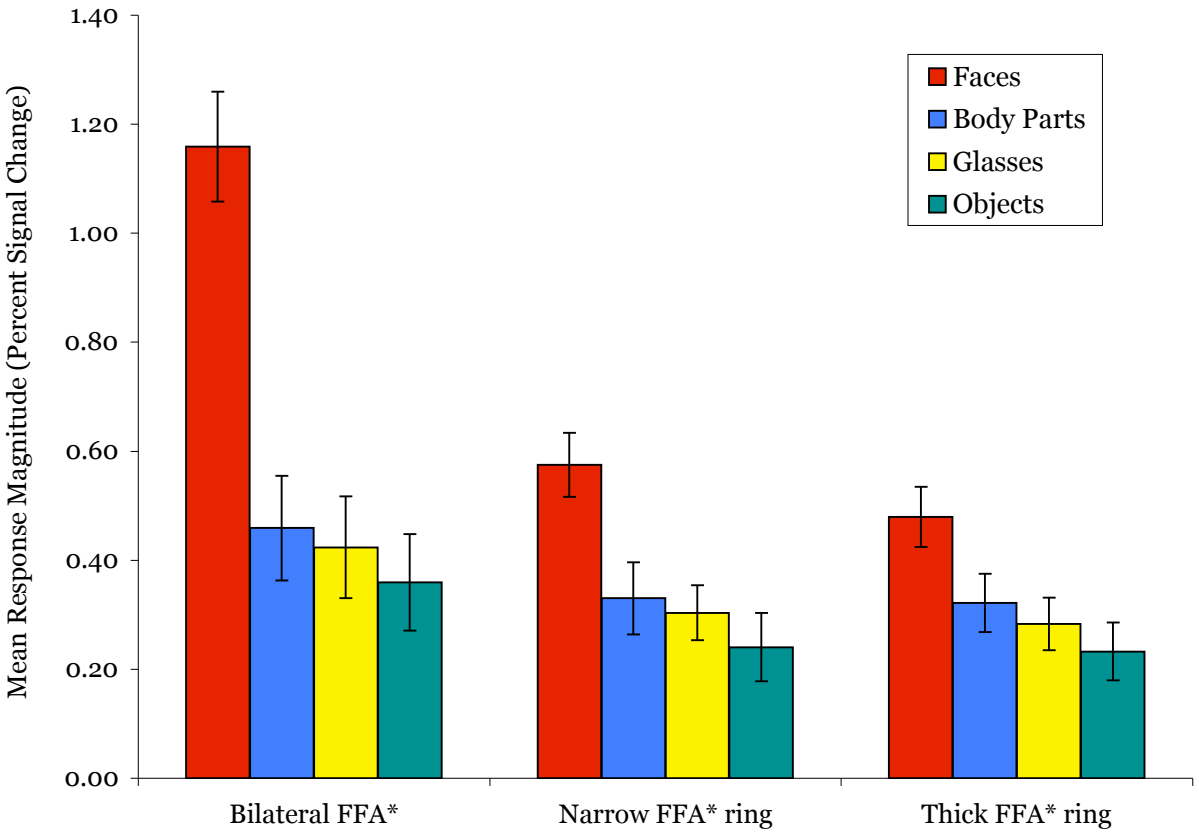

Figure 4: Category selectivity in FFA* rings.

Response magnitudes from Experiment 1 are plotted for the bilateral FFA* ROIs (left), the narrow FFA* rings (middle), and the thick FFA* rings (right).

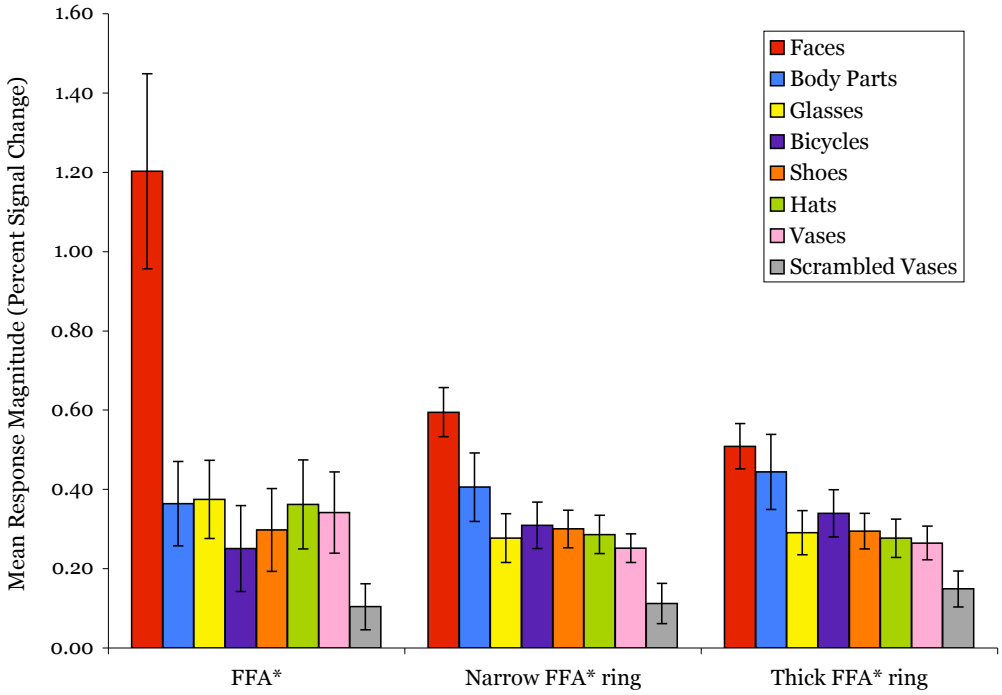

Figure 5: Category selectivity in FFA* rings.

Response magnitudes from Experiment 2 are plotted for the bilateral FFA* ROIs (left), the narrow FFA* rings (middle), and the thick FFA* rings (right). These results show no evidence for glasses selectivity in or around the FFA*.

4. Is Conceptual Knowledge an Organizing Dimension of Ventral Visual Cortex?

Introduction

According to one influential theory, semantic information in general, and the distinction between animate and inanimate objects in particular, is a major organizing principle that influences the spatial position of object selectivities within the ventral visual cortex (VVC). Although the VVC has traditionally been more strongly implicated in the processing of perceptual information about objects than conceptual information, several recent lines of evidence suggest that the functional organization of the VVC may nonetheless respect the conceptual distinction between animate and inanimate objects. We test this hypothesis in this chapter using fMRI with a region-of-interest (ROI) approach.

Early evidence for the idea that the brain respects the animate/inanimate distinction comes from neuropsychological observations of patients with brain damage who cannot

name or understand animate objects but do not show similar impairments for inanimate objects, and other patients who show the opposite pattern of deficit (Warrington and McCarthy, 1983; Warrington and McCarthy, 1987; McCarthy and Warrington, 1988). This double dissociation led scientists to hypothesize that conceptual knowledge about objects may be fundamentally divided along the lines of animacy. However, the patient data just described does not necessarily implicate the VVC per se as the locus of the animate/inanimate distinction.

More recently, functional imaging studies have found evidence for an animate/inanimate distinction in the ventral visual cortex. For example, Martin and colleagues found that regions of the lateral fusiform gyrus and superior temporal sulcus are selectively engaged in processing images of animate objects (animals or bodies), while the medial fusiform gyrus and the middle temporal gyrus are selectively engaged in processing images of tools, an inanimate object category (Beauchamp et al., 2002, 2003; Wheatley et al., 2005). The lateral fusiform activation described in these studies likely overlapped with or encompassed the face-selective FFA or body-selective FBA, and the STS activation probably included the body-selective EBA (Beauchamp et al., 2002), although these studies did not contain functional localizers designed to identify these particular ROIs. Further, Martin and colleagues have reported that category-selective regions within the VVC or their neighbors may also contain conceptual knowledge about objects and object properties (Martin and Chao, 2001; Martin, 2007): cortical areas adjacent to those activated by viewing animals and tools were activated by reading or silently generating the name of those same categories (Chao et al., 1999), even under conditions designed to minimize imagery (Wheatley et al., 2005). Finally, another neuroimaging study in humans (Downing et al., 2006) scanned participants on 20 object categories and used these to identify regions of VVC that responded more to animate than inanimate objects and vice versa. The

cortical areas with a strong preference for animate objects overlapped in part with the FFA and EBA, while cortical areas preferring inanimate objects overlapped in part with the scene-selective areas PPA and TOS. These studies collectively raise the question of whether the cortical locus of selectivities for faces, bodies, animals, tools, and other objects can be explained within a broad distinction between animate and inanimate objects.

In a rather different vein, a neurophysiological study of individual neurons in monkey inferotemporal cortex (Kiani et al., 2007) also found evidence for a neural correlate of the animate/inanimate distinction in the VVC: the population code of these neurons for more than 1,000 object images naturally formed two global clusters distinguishing animate items (e.g., faces, bodies, reptiles, and insects) from inanimate items (e.g., cars, plants, food, and artifacts). Similarly, a multivariate analysis of high-resolution fMRI in humans found that the voxelwise patterns of response in the VVC to various object images formed clusters based on this animate/inanimate distinction (Kriegeskorte, Mur, Ruff, Kiani, Bodurka, and Bandettini, HBM 2007).

An explanation of VVC organization based on conceptual knowledge and animacy would also be consistent with the findings in Chapter 2 of this thesis. This hypothesis would explain the adjacency of face and body selectivities as a result of their shared status as animate, living things. However, the hypothesis does not appear to be consistent with the finding in Chapter 3 of selectivity in the EBA for glasses. In this case, an area selective for animate objects (body parts) produced a higher response to an inanimate object category (glasses) than to another animate object category (faces). In this study, we use fMRI to follow up on this apparent evidence against the animacy hypothesis. In particular, we test whether it is the EBA itself that shows a preferential response to glasses compared to faces, rather than a distinct region overlapping with the EBA, and whether this effect generalizes

to other inanimate objects categories. This experiment will enable us to determine whether the high response to glasses in the EBA presents a fundamental challenge to the animacy hypothesis as an organizing dimension in the VVC.

Materials and Methods

The experimental stimuli, design, and functional imaging of both Experiment 1 and Experiment 2 in this chapter are identical to those described in Chapter 3 of this thesis, as the same data were used for both studies. These methods will be briefly described here, but see Chapter 3 for a more detailed description of the procedures.

Experiment 1: Stimuli and design. During the scan, participants completed 10 blocked runs for this experiment, each of which consisted of three 16-second fixation blocks and ten 16-second stimulus blocks (two blocks for each of the five stimulus conditions.) The four conditions presented to all participants were: faces, assorted body parts, glasses, and assorted everyday objects. Condition order was counterbalanced both within runs and across runs in each scan session.

Images presented in this experiment came from one of two separate sets of stimuli. These stimulus sets were non-overlapping, such that each image belonged to only one set. All of the stimuli presented within a given run came from the same stimulus set, and the runs were divided such that half of the runs came from one stimulus set and half came from the other. For each block, 20 images from a single stimulus class were presented foveally (300 ms per image, with a 500 ms ISI). To sustain the attention of the participant on the stimuli, they were instructed to perform a 1-back task, which required them to press a button whenever images were consecutively repeated.

Experiment 1: Functional imaging. Eight participants were scanned on a 3.0 T Siemens

(Erlangen, Germany) TimTrio scanner. Twenty-two slices (2.0 mm thick with a 0.4 interslice gap) were manually oriented roughly orthogonal to the calcarine sulcus. The in-plane resolution was 1.4 x 1.4 mm. The data from two of the participants were excluded from further analysis due to excessive head motion.

Experiment 2: Stimuli and Design. The design of Experiment 2 was identical to the design of Experiment 2 in Chapter 3. Participants performed a 1-back task while viewing 12-16 runs of centrally presented stimuli. This experiment contained eight conditions: faces, body parts, glasses, bicycles, shoes, hats, vases, and grid-scrambled versions of the vases. See Figure 1 in Chapter 3 for examples of stimuli. Each run contained one block per condition, and the serial position of each condition was counterbalanced both within and across runs for each participant. As in Experiment 1, runs drew from two non-overlapping sets of stimuli.

Experiment 2: Functional imaging. Five participants were scanned on a 3.0 T Siemens (Erlangen, Germany) TimTrio scanner. Twenty to twenty-two slices were manually aligned roughly perpendicular to the calcarine sulcus. Voxel dimensions were 1.4 x 1.4 x 2.0 mm with a 0.4 mm interslice gap. Data from 1 participant were excluded from further analysis due to excessive head motion.

Experiments 1 and 2: Data Analysis. Data analysis was performed using Freesurfer and FS-FAST software (http://surfer.nmr.mgh.harvard.edu/). The acquired images were motion corrected (Cox and Jesmanowicz, 1999) prior to functional analysis. For most of the analyses in Experiment 2, data from one set of runs were smoothed with a full width half maximum Gaussian kernel of 3 mm. Data from the other set of runs were not smoothed.

For both Experiment 1 and Experiment 2, regions-of-interest, or ROIs, were individually

defined for each participant using the smoothed data from one set of runs ("defining runs") as a set of contiguous voxels with a contrast difference of $p < 0.0001$ uncorrected. We then used fROI (http://froi.sourceforge.net/) to extract the mean of the signal magnitude across all voxels in each ROI to the various conditions in the unsmoothed data from the other set of runs ("evaluating runs"). The data and stimuli used to define the ROIs were entirely separate from those used to calculate the mean response magnitude to each stimulus in each ROI.

In a subsequent analysis, we identified the most selective voxel (or "peak voxel") from each EBA. The peak voxel is the single voxel from each participant's original EBA ROI that demonstrates the smallest p-value for the contrast of body parts > objects in the defining runs. Each peak voxel was then treated as a single-voxel ROI from which response magnitudes from the evaluating runs were extracted.

For a third type of analysis, we used the method of Downing and colleagues (Peelen et al., 2006; Downing et al., 2007b; Peelen and Downing, 2007), for which data are split into two sets and t-values are computed for key contrasts (e.g., faces > objects) in each voxel of the ROI for each data set. This analysis was performed on the same EBA ROIs that were identified with the data from the defining runs and were used in all other analyses in this experiment. In this case, however, both sets of runs (defining and evaluating) were also used to evaluate the patterns of response within the ROI. None of the data used to evaluate these patterns was spatially smoothed. For each of these two data sets, voxelwise maps were computed of the t-statistics based on the following three contrasts: body parts > objects, faces > objects, and glasses > objects. We then computed correlations across the two data sets of the three t-statistic maps.

It is important to note that the multi-voxel analysis described here breaks a

fundamental rule of ROI analyses, namely that data from defining and evaluating runs must be independent of one another. In this case, data from the same runs were used both to define the ROI and then to evaluate the pattern of response within it. This is problematic when analyzing mean response magnitude in a region, as the criteria used to identify the region will fundamentally bias the results you obtain. However, this confound does not extend to pattern analyses. Critically, in this study each voxel was selected for inclusion in the ROI based on its selectivity for body parts, however this method of selection is unrelated to the pattern of body selectivity across the individual voxels within the ROI. The method of voxel selection for ROIs (body parts > objects) could have potentially included voxels with particularly low responses to objects, rather than ones that have particularly high responses to body parts. If this were the case, we might expect to see high correlations amongst body-, face-, and glasses selectivities, as they all are measured with respect to the assorted object baseline. However, this explanation could not account for zero or negative correlations between selectivity patterns, nor could it account for any systematic differences in correlation magnitude between different category selectivity patterns. Therefore, any differences between correlation measures for the different pairs of selectivities would be due to the EBA response patterns to body parts, glasses, and faces, rather than to objects.

Experiment 1 Results

Experiment 1: ROI localization

The body-selective area EBA was defined with a contrast of body parts > assorted objects using a contrast threshold of $p < 0.0001$ uncorrected, and was found bilaterally in all participants.

Experiment 1: Mean response magnitude in EBA

All subsequent ROI analyses were performed using the independent data set from the evaluating runs. We measured the mean response magnitude to the four conditions (faces, body parts, glasses, and assorted objects) across all voxels in the EBA. Results are shown in Figure 1.

The response to body parts in the EBA was significantly greater than to faces, glasses, and objects (all three, $p < 5x10^{-5}$), replicating the body selectivity used to define this ROI. The response to assorted objects was significantly lower than to faces ($p < 0.005$) and glasses ($p < 5x10^{-7}$). Strikingly, the EBA response to glasses was significantly greater than to faces ($p < 0.001$). This finding appears to be incompatible with the hypothesis that objects processing is anatomically segregated based on object animacy.

Experiment 1: Response magnitude in single peak voxels

The functionally defined EBA occupies a large region of cortex and has been shown to overlap with object-selective area LO and motion-selective area MT (Spiridon et al., 2006; Downing et al., 2007a). An overlap between EBA and an unknown region preferring glasses over faces could result in an erroneous finding of glasses selectivity in the EBA, just as overlap with a neighboring body area resulted in an erroneous finding of body selectivity in the FFA in Chapter 2. This overlap could be due to neurons with dual selectivities, to overlapping boundaries of distinct sets of neurons, or to technical limitations such as inadequate spatial resolution, among other things. We first addressed this question by repeating the analyses on the peak voxel in the EBA that most reliably demonstrates body selectivity. The criteria used to select peak voxels are described in further detail in the Methods section.

Results from the analysis of EBA peak voxels (shown in Figure 2) demonstrate that these voxels produced a significantly greater response to body parts than to all other categories (all

three, $p < 0.0005$), replicating the body selectivity used to define the voxels. The response to faces was greater than to objects, although this difference failed to reach significance ($p = 0.08$). However, the critical contrasts involve the response of the peak voxel to glasses. The response to glasses was significantly greater than to both objects and faces (both, $p < 0.01$). Therefore, the analysis of the EBA peak voxel produced the same pattern of results found with the entire ROIs. These results provide further evidence that glasses selectivity is present within the EBA proper, rather than in an overlapping, adjacent cortical location.

Experiment 1: Multi-voxel pattern analysis

Although analysis of the single peak voxel provides a measure of the response to the experimental conditions at the most reliably selective point in the EBA, it does not tell us the relationship between object selectivities across the rest of the ROI. Moreover, this peak voxel may not be representative of the response in other voxels around the most selective part of the EBA, or it may be located on a blood vessel that is supplying separate but adjacent glasses and body regions. Another way to test the segregation of body and glasses selectivities in the EBA is to correlate the voxelwise selectivity for these conditions (Peelen and Downing, 2007). If the higher response to glasses than faces in EBA is due to overlap, then correlations between body parts and glasses should be positive and greater than correlations between body parts and faces.

To conduct this analysis, we used the method of Downing et al (Peelen et al., 2006; Downing et al., 2007b; Peelen and Downing, 2007). According to this method, we separately analyzed the two data sets (both unsmoothed) and computed t-values the contrast of interest (e.g., body parts > objects) in each voxel of the ROI for each data set. We conducted this analysis on the same ROIs used for the rest of Experiment 1, which were identified with the smoothed data from the defining runs. However in this case, both

sets of runs were also included in the analysis to identify the pattern of response in the ROIs. For both of these data sets, t-maps were separately computed for every individual voxel in the ROI based on the following three contrasts: bodies > objects, faces > objects, and glasses > objects. We then computed correlations of the resultant three t-maps across the two data sets. See Methods for a discussion of the potential dangers of using the same data to define and evaluate the ROI, and how those issues are addressed in this study.

The multivariate voxelwise approach allows us to ask whether the elevated mean selectivities to both bodies and glasses in the EBA arise from the same or different sets of voxels, which can be useful in asking whether the dual selectivities represent two overlapping or interdigitated selectivities (Peelen and Downing, 2007), or if they reflect shared neural machinery dedicated to processing both types of stimuli. See Table 1 for the results of this analysis. We find in the EBA, on a voxel-by-voxel level, strong correlations for the same selectivities across data sets (i.e., body parts with body parts, faces with faces, and glasses with glasses.) All three of these correlations were significantly above zero (all three, $p < 10^{-5}$). However, neither correlations between patterns of body- and face selectivity nor correlations between patterns of glasses- and face selectivity were significantly different from zero ($p = 0.56$ and 0.22, respectively). The fact that these correlations were close to zero indicates that correlations in this analysis are not artificially high due to low responses to assorted objects in these voxels. Crucially, the correlation between body selectivity and glasses selectivity is significantly above zero ($p < 5x10^{-4}$). Moreover, on a voxel by voxel basis, patterns of body selectivity were more correlated with glasses selectivity than with face selectivity ($p < 5x10^{-5}$).

Thus, all three analyses lead to the same conclusion, of striking overlap in selectivities for bodies and glasses. These findings provide further evidence against an animate/inanimate

distinction in the organization of the VVC. However, while the elevated response to glasses in the EBA argues against this hypothesis of cortical organization, it is not easily explained by other pre-existing hypotheses. Although several new hypotheses might be put forth to explain this pattern of results, it is first necessary to understand whether the preferential response to glasses is unique to that particular object category, or whether it generalizes to other nonliving types of objects. Experiment 2 was conducted to address this question.

Experiment 2 Results

To test whether inanimate objects other than glasses also activate the EBA more than faces do, we included additional object conditions to those used in Experiment 1. These conditions were: bicycles, hats, shoes, vases, and grid-scrambled versions of vases. All stimuli in each individual condition of this experiment came from the same object category; therefore the assorted object condition was not included in this experiment. As in Experiment 1, runs were composed of blocks of stimuli from one of two non-overlapping stimulus sets. Data from the runs with one set of stimuli were spatially smoothed and used to define the ROIs, while independent data from the second set of runs were used to assess the response magnitude to the stimulus conditions in those ROIs.

Stimulus conditions were selected to test various hypotheses that might explain the elevated response to glasses in the EBA. Hats were selected as another inanimate object that is associated with people, and specifically with faces. Shoes are also associated with people, but are specifically associated with bodies rather than faces. Bicycles were selected for their similarity of shape with glasses, although in both cases the specific configuration and viewpoint of the objects were varied across the individual images used for these stimuli. (See Figure 1 of Chapter 3 in this thesis for examples.) Vases were included to serve as a control category that is not associated with people and does not share shape features with

glasses, and grid-scrambled versions of vases were included to demonstrate the baseline response of the region to visual stimuli that do not depict any object category or coherent shape.

Experiment 2: ROI localization

The EBA was defined with a contrast of body parts > vases using a contrast threshold of $p < 0.0001$ uncorrected and was found bilaterally in all participants.

Experiment 2: Mean response magnitude in EBA

Due to the limited number of participants in Experiment 2, ROIs were analyzed bilaterally to maximize the possibility of detecting significant differences between conditions. Results from the analysis of bilateral EBA (shown in Figure 3) revealed greater activation to glasses than to faces in this region that nearly reached significance ($p = 0.05$). This result is similar to the results in Experiment 1. However, the key question in Experiment 2 is whether a preferential response to glasses in the EBA would generalize to any other object categories that share various properties with glasses. Indeed, we found that the EBA was activated more by bicycles than faces ($p < 0.005$), and was activated marginally more for shoes than faces, ($p = 0.06$). These results demonstrate that multiple nonliving object categories can elicit higher responses in the EBA than do faces. However, not all inanimate categories elicited responses above that to faces; The EBA produced greater responses to glasses, bicycles, and shoes than to hats (all three, $p < 0.05$) and vases (all three, $p < 0.005$). Moreover, the magnitude of response in the EBA to faces was similar to that for hats and vases, such that these means are not significantly different ($p = 0.56$ and 0.59, respectively). This pattern of results demonstrates that only a subset of inanimate object categories elicit a greater response than faces in the EBA. Further, the failure of hats to do so indicates that that distinction is not determined based on whether the object category is associated with

faces, or even with people. Moreover, the fact that shoes elicit a heightened response in addition to glasses and bicycles suggests that the EBA activation is not due to a preference for that particular object shape.

Discussion

In this study, we tested whether the response properties of the body-selective area EBA were consistent with the hypothesis that conceptual distinctions in general, and object animacy in particular, serve as an organizing dimension of VVC. Both the entire EBA ROI and the most body-selective voxels in this ROI showed a greater response to glasses than to faces. Moreover, a multivariate analysis of selectivity for bodies, faces, and glasses showed that, on a voxel-by-voxel basis, body selectivity co-segregated more with glasses than with faces. This finding serves as an example in which the similarity in the cortical layout of selectivities for one animate and one inanimate object category (body parts and glasses) is greater than between two animate objects (body parts and faces). Finally, we showed that the activation of EBA by glasses is not specific to that object category, but rather extends to other categories such as bicycles and shoes. These findings provide powerful evidence against the hypothesis that object animacy serves as a universal, large-scale organizing dimension of VVC.

While our results argue against the animacy hypothesis, they do not fit with any alternate established theory. If there is, in fact, some organizing principle to the distribution of selectivities in the VVC, then body parts, glasses, bicycles, and shoes (but not hats and vases) should have some property in common that would explain their shared cortical real estate and the computations performed therein. One possibility is that the shapes of glasses and bicycles and shoes are more similar to body parts than they are to the shapes of the less-preferred objects. However, the body selectivity in EBA includes whole bodies, any

recognizable body parts, and even stick figures of bodies, all in any number of positions and configurations.

A notable feature that is common to the stimuli that produce relatively strong responses in the EBA is that their configurations can be changed, while the configurations of the other stimuli are relatively fixed. In particular, glasses, bicycles, and bodies have articulated joints that can be bent to radically change their shapes. If this feature were relevant to the apparent preference of the EBA for these objects, at least two property dimensions could be proposed to explain it. In one case, objects with changeable configurations may share the property of requiring special computations for the purposes of visual recognition. Specifically, configural changes represent an additional transformation the visual system must overcome when recognizing an object. It is possible that the EBA performs these computations for objects that can undergo these changes, be they bodies and glasses or folding chairs and construction cranes. The intermediate response to shoes in the EBA does not fit as well with this hypothesis, although the laces and buckles on shoes do result in some degree of configural change.

Another property that ties together the preferred object categories in the EBA is the complex and somewhat stereotyped ways that humans interact with these objects. Glasses can be folded and unfolded, while bicycles can be pedaled and steered. Shoelaces can be tied and shoe buckles buckled. In contrast, hats and vases are generally associated with rigid motion and less specific motor actions. Therefore, it may be that the common property among the objects that elicit a strong response in the EBA is their association with complex motor actions. Prior studies have already shown that the EBA may be sensitive to motion and motor information. The EBA responds more strongly to moving than static stimuli (Spiridon et al., 2006), and is particularly sensitive to biological motion (Beauchamp

et al., 2002), even in the absence of the human form (Beauchamp et al., 2003; Peelen et al., 2006). Studies have also shown that the EBA or adjacent cortex is also sensitive to object-directed motion of one's own body (Astafiev et al., 2004; Peelen and Downing, 2005). Although our stimuli were static images of objects and participants lay passively during the scans, the images of these objects may have triggered motor representations or motor imagery, much as static images implying motion activate the motion-selective areas MT/MST (Kourtzi and Kanwisher, 2000), and learned shape-motion associations cause neurons in monkey MT to respond to those shapes when they are stationary (Schlack and Albright, 2007). Finally, when participants were trained to use novel objects as tools, later viewing of static images of those objects elicited activation of the left middle temporal gyrus (Weisberg et al., 2007), which is in the vicinity of a tool-selective area, the motion-selective MT, and the EBA.

In opposition to this hypothesis and our findings, the first report on the EBA found no difference in the response to articulated objects versus object parts (Downing et al., 2001). However, these stimuli were presented within each block as a mixture of different types of articulated objects, whereas Experiment 2 of this study used blocks comprised of only one object category. This difference may be relevant, since the activation of object-specific motor information may take additional time and could be prevented by a serial presentation of objects that are manipulated in different ways.

Future studies including more categories of objects will be required to assess whether the EBA consistently shows biases for articulated objects or objects associated with complex actions. The question also remains of whether responses in the EBA to non-body stimuli are relevant for behavior. Techniques such as TMS could be useful in addressing this question (Urgesi et al., 2004). While the results of this study demonstrate a striking contradictory

example to the animate-inanimate distinction, it does not negate the large body of work from neuropsychology, electrophysiology, and neuroimaging that have found evidence for neural dissociations between the representations of animate and inanimate objects. One explanation for this contradiction is that an animacy distinction does exist in parts of VVC, but it does not serve as a universal organizing dimension. Another possibility is that this distinction occurs because of an unrelated dimension that generally segregates object representations along the same lines. Further work is necessary to disentangle these hypotheses.

References

Astafiev SV, Stanley CM, Shulman GL, Corbetta M (2004) Extrastriate body area in human occipital cortex responds to the performance of motor actions. Nat Neurosci 7:542-548.

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

Beauchamp MS, Lee KE, Haxby JV, Martin A (2003) FMRI responses to video and point-light displays of moving humans and manipulable objects. J Cogn Neurosci 15:991-1001.

Chao LL, Haxby JV, Martin A (1999) Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. Nat Neurosci 2:913-919.

Cox R, Jesmanowicz A (1999) Real-time 3D image registration for functional MRI. Magn Reson Med 42:1014-1018.

Downing PE, Wiggett AJ, Peelen MV (2007a) Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. J Neurosci 27:226-233.

Downing PE, Wiggett AJ, Peelen MV (2007b) Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. J Neurosci 27:226-233.

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

Downing PE, Chan AW, Peelen MV, Dodds CM, Kanwisher N (2006) Domain specificity in visual cortex. Cereb Cortex 16:1453-1461.

Kiani R, Esteky H, Mirpour K, Tanaka K (2007) Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. J Neurophysiol 97:4296-4309.

Kourtzi Z, Kanwisher N (2000) Activation in human MT/MST by static images with implied motion. J Cogn Neurosci 12:48-55.

Martin A (2007) The representation of object concepts in the brain. Annu Rev Psychol 58:25-45.

Martin A, Chao LL (2001) Semantic memory and the brain: Structure and processes. Curr Opin Neurobiol 11:194-201.

McCarthy RA, Warrington EK (1988) Evidence for modality-specific meaning systems in the brain. Nature 334:428-430.

Peelen MV, Downing PE (2005) Is the extrastriate body area involved in motor actions? Nat Neurosci 8:125-126.

Peelen MV, Downing PE (2007) Using multi-voxel pattern analysis of fMRI data to interpret overlapping functional activations. Trends Cogn Sci 11:4-5.

Peelen MV, Wiggett AJ, Downing PE (2006) Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. Neuron 49:815-822.

Schlack A, Albright TD (2007) Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. Neuron 53:881-890.

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

Spiridon M, Fischl B, Kanwisher N (2006) Location and spatial profile of category-specific regions in human extrastriate cortex. Hum Brain Mapp 27:77-89.

Taylor JC, Wiggett AJ, Downing PE (2007) Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. J Neurophysiol 98:1626-1633.

Urgesi C, Berlucchi G, Aglioti SM (2004) Magnetic stimulation of extrastriate body area impairs visual processing of nonfacial body parts. Curr Biol 14:2130-2134.

Warrington EK, McCarthy R (1983) Category specific access dysphasia. Brain 106 (Pt 4):859-878.

Warrington EK, McCarthy RA (1987) Categories of knowledge. Further fractionations and an attempted integration. Brain 110 (Pt 5):1273-1296.

Weisberg J, van Turennout M, Martin A (2007) A neural system for learning about object function. Cereb Cortex 17:513-521.

Wheatley T, Weisberg J, Beauchamp MS, Martin A (2005) Automatic priming of semantically related words reduces activity in the fusiform gyrus. J Cogn Neurosci 17:1871-1885.

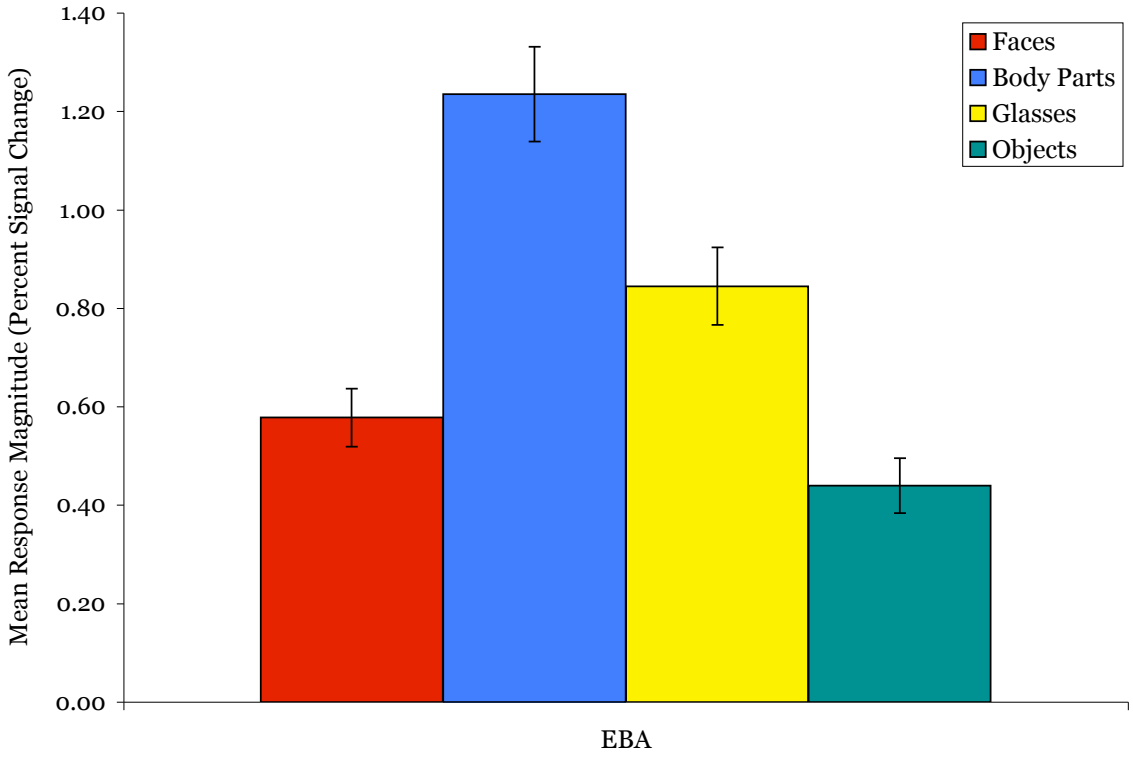

Figure 1: Mean response magnitude in the EBA

The average response in the EBA to each of the four stimulus conditions in Experiment 1 is shown here and demonstrates a strong selectivity for glasses in this region.

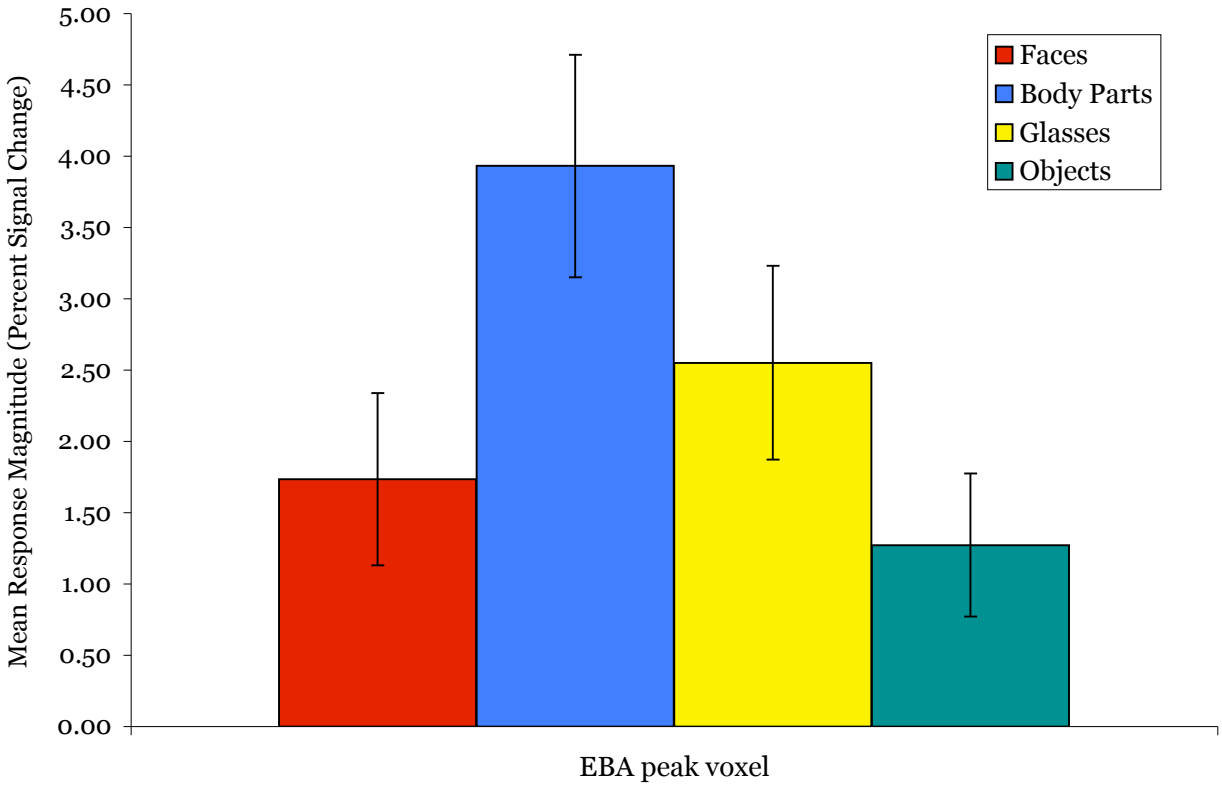

Figure 2: Response magnitude in the single peak EBA voxel

The responses of the EBA peak most selective voxel to each of the conditions in Experiment 1, shown here, replicate the glasses selectivity found with the entire ROI.

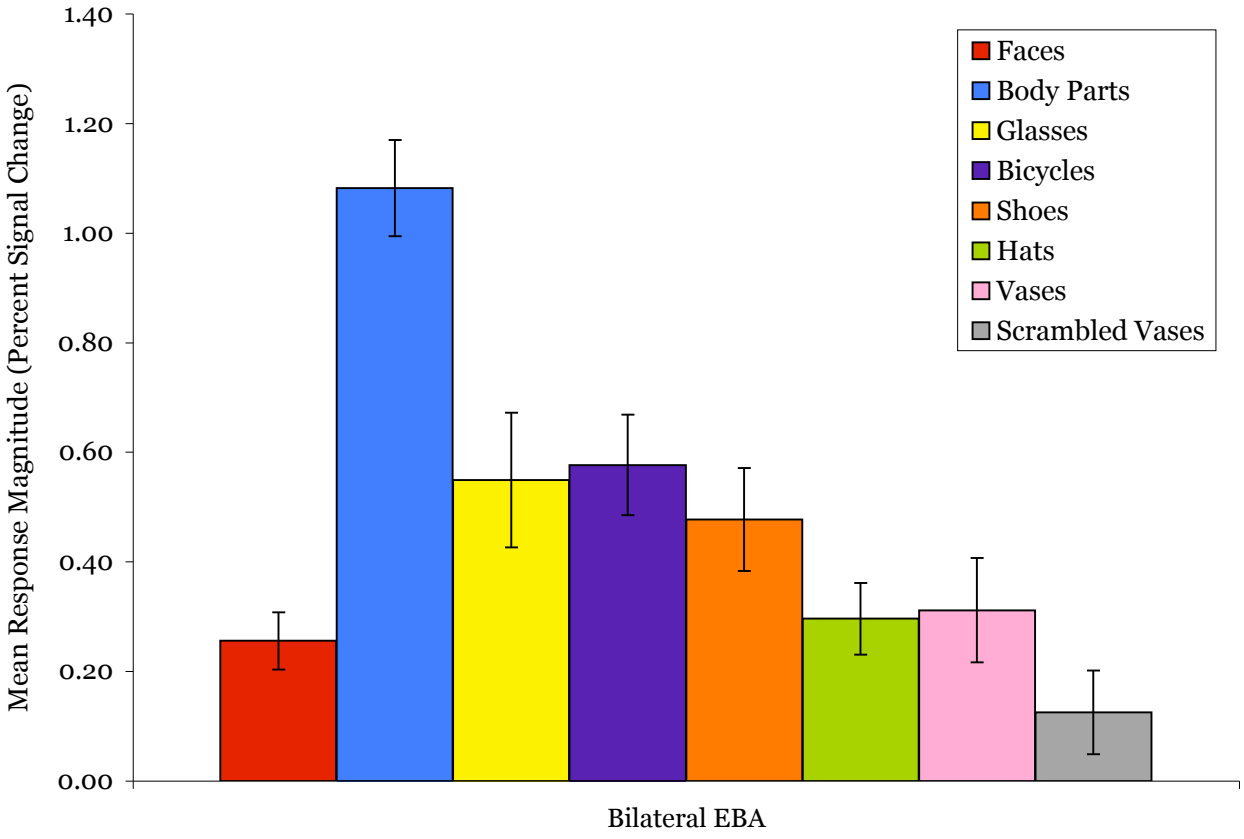

Figure 3: Mean response magnitude in the EBA

The average response in the EBA to each of the conditions in Experiment 2 demonstrates selectivity in this region for glasses, bicycles, and shoes in contrast to vases, hats, scrambled vases and, marginally, faces.

	Body Parts	Faces	Glasses
Body Parts	**0.47** (0.05)	0.02 (0.03)	**0.25** (0.03)
Faces		**0.40** (0.05)	-0.04 (0.03)
Glasses			**0.27** (0.03)

Table 1: Mean correlations between voxelwise patterns of selectivity in the EBA.

Correlations (with standard errors) between patterns of selectivity for body parts, faces, and glasses in the EBA are shown here, with those that differ significantly from zero (at $p < 0.05$ in a 2-tailed, single-sample t-test) in bold.

5. The Distribution of Category and Location Information in Ventral Visual Cortex

Introduction

Ungerleider and Mishkin (Underleider and Mishkin, 1982) argued in a seminal paper that information about form and location are segregated into separate processing streams in the primate visual system. Subsequent studies using lesions, neurophysiology, and fMRI have generally supported this hypothesis or its variants (Goodale and Milner, 1992). However, other evidence indicates that the two pathways are not completely distinct, but instead have multiple interconnections (Merigan and Maunsell, 1993), and that the occipitoparietal "where" pathway (Sereno and Maunsell, 1998) contains shape information and the occipitotemporal "what" pathway contains location information (Op De Beeck and Vogels, 2000; DiCarlo and Maunsell, 2003). Here we used mean population response and multivariate pattern methods (Haxby et al., 2001; Haynes and Rees, 2006; Norman et

al., 2006) with a region-of-interest (ROI) approach to ask how much location information is present in shape-selective cortex in humans, how that location information is distributed across specific functionally defined regions of occipitotemporal cortex, and how location information relates to category information in this pathway.

Extensive fMRI investigations over the last decade have characterized the functional organization of the occipitotemporal pathway in humans. Multiple cortical regions have been defined by their selectivity for general object shape, or by their selectivity for specific categories such as bodies, faces, and scenes. For each of these kinds of selectivity, two ROIs have been identified, one on the ventral surface of the brain and one on the lateral surface; For example, the body-selective fusiform body area, or FBA (Peelen and Downing, 2005; Schwarzlose et al., 2005), lies on the ventral surface and the extrastriate body area, or EBA (Downing et al., 2001), lies on the lateral surface.

Although much work has been done to characterize the shape or category selectivity of these regions, very little is known about whether they also process information about object location. At the most general level, some of these regions demonstrate contralateral field biases (Niemeier et al., 2005; Hemond et al., 2007; Macevoy and Epstein, 2007; McKyton and Zohary, 2007), and some of them (e.g., the parahippocampal place area, or PPA) respond more strongly to stimuli presented in the periphery, whereas others (e.g., the fusiform face area, or FFA) respond more strongly to foveal stimuli (Levy et al., 2001; Hasson et al., 2003). Studies conducted with retinotopic mapping (Brewer et al., 2005; Larsson and Heeger, 2006) have shown object-selective responses in certain retinotopically defined regions, although the degree to which these maps overlap object-selective cortex is not yet known. Other studies have found elevation biases (Niemeier et al., 2005), as well as repetition suppression sensitivity to translation around fixation (Grill-Spector et al.,

1999) in lateral occipital area LO. While each of the aforementioned studies has shown a specific kind of location information in a small number of regions, to date there has been no comprehensive examination of location information across the many category-selective functionally defined regions spanning both lateral and ventral surfaces. Here we set out to do just that, using a method sensitive to both retinotopic and spatiotopic location information.

Materials and Methods

Stimuli and design. Participants performed blocked Localizer scans to identify ROIs, as well as separate blocked Experimental scans to measure the response of these regions to stimuli of different categories in different locations. Scans of each type (Localizer and Experimental) alternated throughout the scan session.

Participants completed five or six runs of the Localizer scans, each of which consisted of three 16 sec blocks of fixation and two 16 sec blocks for each of five different stimulus classes (headless bodies, faces, outdoor scenes, assorted everyday objects, and grid-scrambled versions of those objects.) The conditions were presented in palindromic order within each run, and the serial position of each condition was counterbalanced within participants across the scan session. For each block of the Localizer scan, twenty images from a single stimulus class were foveally presented (300 ms per image, with a 500 ms interstimulus interval). Scrambled object stimuli were constructed by superimposing a grid over the objects and relocating the component squares randomly (Kourtzi and Kanwisher, 2000). To ensure that participants paid attention while they freely viewed the images, they performed a 1-back task in which they were asked to make a key-press whenever images were repeated consecutively, which happened twenty times per run.

In the same scan session as the Localizer runs, participants performed between eight and twelve runs of a blocked experiment designed to test category and location selectivity in the ROIs. For these scans, participants were instructed to fixate on a central cross while images were presented at one of three locations (at, above, or below fixation, with 5.25° of visual angle between the center of the image and the center of the fixation cross in the above and below conditions.) In order to roughly equate performance across conditions, peripheral images had to be scaled by more than the standard cortical magnification (Duncan and Boynton, 2003). Foveal stimuli were images 1.6° wide and high, whereas peripheral stimuli were 7.8° wide and high. Thus, foveal and peripheral images occupied non-overlapping locations in visual space, with an intervening gap of 0.55°, as shown in Figure 1. The stimuli used in these scans belonged to one of four categories: headless bodies, faces, cars, and scenes. Completely non-overlapping sets of stimulus images (40 images per condition in each) were used for the Localizer and Experimental scans, and all Experimental runs drew from the same set of stimuli. Each stimulus class was presented in each location for one block in every run (resulting in twelve conditions and twelve visual blocks per run). Each 16 sec block consisted of twenty image presentations (300 ms per image, with a 500 ms ISI). Location and stimulus class remained constant within a block. In addition, each run contained two 16 sec fixation blocks. During these Experimental scans, participants performed the same 1-back task described above. Conditions were counterbalanced across runs to control for block ordering effects.

Four participants took part in a separate retinotopic mapping scan session, during which they viewed chromatic, continuously rotating wedges or expanding/contracting rings while performing a contrast decrement detection task at fixation. These participants each viewed three or four runs of rotating wedge angular mapping and two runs of ring eccentricity mapping, as well as five or six Localizer runs. Full details of the retinotopic mapping stimuli

and analysis methods have been provided elsewhere (Swisher et al., 2007).

Functional imaging. Thirteen participants were scanned on a 3.0 T Siemens (Erlangen, Germany) TimTrio scanner at the Martinos Imaging Center at the McGovern Institute for Brain Research (Cambridge, MA). Images were acquired with a gradient echo single-shot echo planar imaging sequence with a repetition time of 2 sec, flip angle 90°, and echo time 33.7 – 34.0 ms. Twenty to twenty-six slices of thickness 2 mm were manually aligned roughly perpendicular to the calcarine sulcus to cover most of occipital, posterior parietal, and posterior temporal cortex. Voxel dimensions were 1.4 x 1.4 x 2.0 mm with a 0.4 mm interslice gap. In addition, 1-2 high-resolution MPRAGE anatomical scans were acquired for each participant in the same scan session. The same scan parameters and similar slice prescriptions were used in the retinotopic mapping sessions. For eight of the thirteen total participants, we monitored eye movements during the scans with an ISCAN (Burlington, MA) model RK-826PCI pupil/corneal reflection tracking system. The data from four participants (including three who were scanned with the eye tracker) were excluded from further analysis due to excessive head motion.

Data Analysis. Data analysis was performed using Freesurfer and FS-FAST software (http://surfer.nmr.mgh.harvard.edu/). The acquired images were motion corrected (Cox and Jesmanowicz, 1999) prior to statistical analysis, and smoothed with a full width half maximum Gaussian kernel of 3 mm for Localizer runs and 2 mm for retinotopic mapping scans. Data from Experimental runs were not smoothed.

ROIs were individually defined for each participant using the Localizer scans. We then used fROI (http://froi.sourceforge.net/) to extract the response magnitude for each voxel in each ROI to the various conditions in the separate Experimental runs. The data and stimuli used to define the ROIs were entirely separate from those used to calculate the

response magnitudes to each stimulus in each ROI. Response magnitudes were analyzed in two ways. First, we took the average of the response in all voxels within a given ROI to compute a mean percent signal change value for each condition. Second, for the pattern analysis we followed the method of Haxby et al. (Haxby et al., 2001). Specifically, we split data from the Experimental scans in half, such that the odd runs were assigned to one data set and the even runs were assigned to the other. Responses in each individual voxel were normalized separately for each data set by subtracting the voxel's mean response across all stimulus conditions from its response magnitude to each of the individual stimulus conditions. This resulted in normalized responses of each voxel for each condition in each of the two data sets (from even and odd runs), producing two voxelwise patterns of response for each condition in each data set and ROI. For each ROI, 144 correlations were computed between the patterns of response for the twelve stimulus conditions in each data set. Finally, these correlations were binned and averaged based on whether the correlated conditions were within category or location (e.g. face^{odd}-face^{even} or upper^{odd}-upper^{even}), or between category or location (e. g. face^{odd}-car^{even} or upper^{odd}-lower^{even}).

Although the mean population response magnitude and the voxelwise patterns of response can both demonstrate the presence of information about location and category in a region, they are orthogonal measurements that assess different neural phenomena. Pattern analyses assess the pattern of responses of subpopulations of neurons within an ROI and determine the degree to which this pattern is stable across data sets and conditions. If changes in mean occur uniformly across the voxels in the ROI, it will have no effect on the corresponding correlations. Conversely, the spatial pattern can differ greatly between two conditions, yet if the average of the responses across all voxels remains the same, then the mean will be unaffected by these changes. See Figure 2, which illustrates the independent effects of changes of means versus spatial patterns.

In a separate analysis, we identified the most selective voxel (or "peak voxel") individually for all FFA*, EBA, PPA, and TOS ROIs. The peak voxel of an ROI was defined as the single voxel that demonstrates the smallest p-value for the defining contrast of that ROI (such as bodies > objects for the EBA) in the Localizer runs. Each peak voxel was then treated like a single-voxel ROI from which response magnitudes from the Experimental runs were extracted.

Results

Eye movements were recorded inside the scanner in five of the participants included in this study to confirm that they maintained fixation throughout the Experimental runs. No significant differences were found in eye position across stimulus categories or locations (see Supporting Information).

ROI localization

The Localizer data were used to identify bilateral extrastriate regions selective for faces, bodies, scenes and objects. The face-selective FFA and the occipital face area (OFA) were defined using a faces > objects contrast. The body-selective areas EBA and FBA were defined with a contrast of bodies > objects, and the scene-selective PPA and an area in the transverse occipital sulcus (here referred to as TOS) were identified with a scenes > objects contrast. Finally, the broadly shape-selective areas that comprise the lateral occipital complex, namely LO and a posterior fusiform area (pFs), were identified with an objects > grid-scrambled objects comparison. All ROIs were defined using a contrast threshold of $p < 0.0001$ uncorrected. Since the FFA and FBA are adjacent and appear to overlap, we excluded any overlapping voxels to create the functionally dissociated regions FFA* and

FBA*, as described by Schwarzlose et al. (Schwarzlose et al., 2005). Not all ROIs were found in every participant, due in part to the limitations of high-resolution slice coverage, as well as the fact that only clusters comprised of twenty or more voxels were counted as ROIs and included in further analyses. The following ROIs were identified, with the number of participants in whom that ROI was identified in parenthesis: right FFA (8); left FFA (4); right FFA* (8); left FFA* (4); right OFA (7); left OFA (6); right EBA (9); left EBA (9); right FBA (8); left FBA (4); right FBA* (5); left FBA* (2); right PPA (5); left PPA (5); right TOS (6); left TOS (6); right LO (5); left LO (6); right pFs (5); left pFs (6). See Figure 3 for a mapping of the relative locations of these ROIs on a representative flattened occipital surface. We also analyzed a posterior, visually active region of cortex near the occipital pole for participants whose slice coverage extended that far back. These ROIs, here denoted as "earlyV" (7), were included in the analyses so that we could compare findings from high-level extrastriate cortex to those from early retinotopic regions (presumably V1 and/or V2).

Mean Response Magnitude in ROIs

All subsequent ROI analyses were performed on the independent data set from the Experimental runs. Using these data, we measured the mean magnitude of response to the twelve conditions (three locations x four categories) across all voxels in each ROI. Figure 4 shows the mean response in each ROI to each category (averaged across locations) and to each location (averaged across categories). These means are based on the data from those participants in whom the ROI could be identified. Each body-, face-, and scene-selective ROI produced a significantly greater response to its preferred category than to the second highest category ($p < 0.005$ for all six), replicating the category selectivity of these regions from prior studies and from the Localizer by which they were defined.

Eccentricity

Prior studies have shown that the magnitudes of response in the FFA, PPA, OFA, and TOS vary with the eccentricity of the stimuli (Levy et al., 2001; Hasson et al., 2003). In those studies, the peripheral stimuli consisted of multiple objects arranged into a ring or, alternately, a single object scaled so that its defining boundaries extended into the periphery (Levy et al., 2001; Levy et al., 2004). To determine whether our paradigm using individual objects replicated this finding, we pooled the mean responses across the categories and conducted paired t-tests comparing the response of images presented at the fovea with those presented in the lower or upper peripheral positions. These averaged response magnitudes are shown in Figure 4 and the corresponding p-values for each contrast are listed in Table 1. Our results support the findings of Hasson et al. (Hasson et al., 2003), demonstrating a peripheral bias in the scene-selective regions, and a foveal bias in face-selective regions.

Elevation

While eccentricity biases (Levy et al., 2001; Hasson et al., 2003) and contralateral biases (Niemeier et al., 2005; Hemond et al., 2007; Macevoy and Epstein, 2007; McKyton and Zohary, 2007) have been reported previously for some higher-level category-selective areas, no studies to date have systematically tested for elevation biases in regions other than LO, which shows a lower visual field bias (Niemeier et al., 2005). We therefore compared mean response magnitude in each ROI for upper and lower field stimuli at equal eccentricities. These response magnitudes are shown in Figure 4 and the corresponding p-values for this contrast are listed in Table 1. Remarkably, most of the ROIs demonstrated significant effects of elevation, such that the scene-selective areas PPA and TOS showed a significantly greater response to upper than lower field images, while the reverse was true of FFA*, EBA, and LO. These results provide new evidence that information about elevation is

widespread across higher-level category-selective regions and that these regions contain different patterns of location biases across measures of both eccentricity and elevation.

Peak Voxel Biases

It is important to note that the location biases found in the previous analysis may be sensitive to the criteria used to define the ROI. If the VVC contains underlying retinotopic maps, then the precise details of where the borders of an ROI land in cortex and what voxels are consequently included in the ROI could affect the location biases found in the mean responses of the ROI. For example, ROI boundaries could be affected by the statistical threshold used to determine the inclusion of voxels in the ROIs. In this case, we used a statistical cutoff of $p < 0.0001$, which is commonly used in the literature but is also arbitrary. Lower thresholds would expand the borders of the ROIs and possibly affect the biases we find in these regions. To address this issue, we measured the magnitude of response of the peak selective voxel in the FFA*, EBA, TOS, and PPA ROIs. The peak voxels were defined for each individual and each ROI as the single voxel within that ROI that demonstrates the highest significance level for its defining contrast (e.g., scenes > objects for the PPA) in the Localizer data.

Results from the peak voxel analysis are shown in Figure 5. Some of the elevation biases found in the whole-ROI analysis failed to reach significance in the peak voxel analysis, possibly due to the increased noise that results from sampling a single voxel rather than a population of voxels. However, critically, in all four cases (FFA*, EBA, TOS and PPA) the stimulus location that elicited the highest mean response from the whole-ROI also elicited the highest response from the ROI peak voxel. These results indicate that the location biases found in the whole-ROI analysis are not heavily skewed based on the threshold used to define the ROI and the resulting location of ROI borders in cortex.

Differences in Visual Feature Distribution

The stimuli used in this study were not symmetrical in the vertical dimension and were all presented upright. Consequently, features were unequally distributed in the upper and lower halves of the images, and these feature distributions varied considerably across the different stimulus categories. As a result, it is possible that location biases found in the ROIs were due to stimulus feature distributions rather than the inherent properties of the regions themselves. For example, the upper half of a face (where the eyes are located) contains more features and may be more diagnostic of faces than the lower half, and thus may drive the FFA to a greater extent. Therefore, the FFA may show an elevation bias because the eyes on a face are located closer to the fovea when the face is in the lower field than in the upper field.

Since the different stimulus categories have different distributions of features, we can use the consistency of location biases across stimulus categories as an indicator of whether the distribution of stimulus features could alternately account for our results. To demonstrate the consistency of location biases across preferred and nonpreferred categories, we separately plotted the mean response magnitudes in each category-selective ROI for the preferred object category and for the average of the three nonpreferred categories in each of the three locations. The resulting graph is shown in Figure 6. In all cases, the rank order of responses to stimuli in the three locations was the same for both preferred and nonpreferred categories, although the magnitudes of biases appear to differ in some of these ROIs.

To illustrate this point further, as well as to test the specific case of faces, we examined the four category-selective ROIs that demonstrated significant differences in their responses to upper versus lower stimuli when the stimuli were averaged across all four categories. These

regions were the FFA* and EBA (which showed lower visual field biases) and PPA and TOS (which showed upper visual field biases). In each of these regions, we used the difference in ROI mean response between upper and lower stimulus locations as a rudimentary indicator of the direction of these biases. This indicator was calculated independently for face, body, and scene stimuli in each ROI and is plotted in Figure 7. If the distribution of features for each stimulus category can account for elevation biases, we should find that the directions of these biases are consistent within stimulus category across all four ROIs. However, we find instead that the FFA* and EBA consistently prefer lower field stimuli for all three stimulus categories, just as the PPA and TOS consistently prefer upper field locations for all three categories. These results demonstrate that elevation biases do not arise from the distribution of features in the upper and lower halves of stimuli, but instead represent a stable property of the functionally defined regions themselves.

In sum, all ROIs showed category information in terms of significantly different responses to different stimulus classes, as expected from prior research. However, critically, nearly all ROIs also showed location information in the form of significantly different mean responses to different stimulus locations. Furthermore, these biases for specific locations are inherent to the functionally defined areas, rather than to their ROI boundaries or to the features of the stimulus categories.

Multivariate Pattern Analyses

Mean responses are only one way of characterizing the information contained in a given cortical region. Information may also be coded at a finer grain, in terms of the pattern of response across voxels in that ROI (Haxby et al., 2001). Prior studies have demonstrated that information not evident in mean response magnitude across an ROI may be discovered in the patterns of response across the voxels in that ROI (Haxby et al., 2001; Williams et

al., 2007). Here we used pattern analysis methods to ask what information is contained in each ROI about object category and object location. To conduct the pattern analyses, we followed the method of Haxby et al. (Haxby et al., 2001); see Methods. For each subject and ROI, 144 correlations were computed, one for each of the possible combinations of one pattern from the even runs (twelve conditions) with one pattern from the odd runs (twelve conditions). For each subject and ROI, we then averaged over these 144 correlation values as a function of whether the even and odd conditions were from the same (within) or different (between) category, and whether they were from the same (within) or different (between) location. The resulting means across subjects are shown in Figure 8.

Omnibus ANOVAs for each ROI

To determine whether location information can be found with pattern analyses we ran a 2x2 repeated-measures ANOVA with category (within or between) and location (within or between) as factors for each ROI. The results of this analysis can be seen in Table 2. In keeping with prior pattern classification findings (Ishai et al., 1999; Haxby et al., 2001; Spiridon and Kanwisher, 2002; O'Toole et al., 2005), all object-selective ROIs demonstrated category information, as indicated by a significant main effect of higher correlations within category than between categories (all eight, $p < 0.0005$). In contrast, there was no evidence of category information in the posterior, retinotopic area earlyV ($p = 0.14$). These findings were as predicted. However, strikingly, all regions except FBA* ($p = 0.40$) also demonstrated location information, as indicated by the significant main effect of higher correlations within location than between locations (eight, $p < 0.05$). Among the ROIs, only LO demonstrated a significant interaction between category and location, such that there was more location information within than between categories ($p < 0.005$). These results show that information about location exists in nearly all category-selective regions

in this study and that this location information is independent of category information for almost all of them.

Since foveal stimuli were considerably smaller than peripheral stimuli, the inclusion of data from all three retinal locations confounds location information with size information. To remove the size confound, we repeated the analysis on correlations from only the upper and lower stimulus positions. Results from this ANOVA are shown in Table 3 and Figure 9. This analysis yielded the same pattern of results, with all ROIs demonstrating a main effect of category information and all except FBA* ($p = 0.73$) demonstrating location information. Moreover, this analysis revealed no interaction between category and location information in any of the regions (all nine, $p \geq 0.10$).

Lateral vs Ventral Surfaces

The functional areas defined in this experiment lie either on the ventral temporal or lateral occipital cortical surfaces, and are laid out such that one region with each category selectivity (bodies, faces, scenes, and objects) is situated on each surface. We next asked whether ventral and lateral regions differ in the amount of information they contain about category and location. To address this question we conducted four repeated-measures ANOVAs of different ROI pairs. Each pair had the same category preference, with one ROI on each surface: bodies > objects for FBA* (ventral) and EBA (lateral); faces > objects for FFA* (ventral) and OFA (lateral); scenes > objects for PPA (ventral) and TOS (lateral); and objects > scrambled objects for pFs (ventral) and LO (lateral). These ANOVAs were run separately on each ROI pair to maximize the number of individual ROIs included in the analyses, since no single participant exhibited every functional region in both hemispheres. Each ANOVA had a 2x2x2 design, with surface (ventral or lateral), category (within or between), and location (within or between) as factors. The results are shown in

Table 4. All region pairs replicated the main effects of category (all four, $p < 0.001$) and location (all four, $p < 0.05$). None of the four ANOVAs showed a significant interaction between category and surface (all $p \geq 0.12$), suggesting that the lateral and ventral regions contain comparable amounts of category information. However, all four ANOVAs showed a significant interaction between cortical surface and location information, such that lateral regions contained more location information than did ventral regions (all four, $p < 0.05$). Therefore, the amount of location information is a distinguishing characteristic between lateral category-selective regions and their ventral counterparts.

The fact that ventral and lateral surfaces show differences in location information could alternately be accounted for by ROI size, since lateral ROIs are, on average, larger than their ventral counterparts. One of the best examples of this disparity is the difference in size between the scene-selective ROIs PPA (average size 68 voxels) and TOS (average size 160 voxels). To test whether ROI size may be mediating the effect, we created PPA and TOS control ROIs comprised of fifteen contiguous voxels each (see Supporting Information for details) and ran the same surface x location x category repeated-measures ANOVA described above. The complete statistical results from that test are shown in Table 5. Crucially, the analysis revealed a significant surface x location interaction ($p < 0.05$), replicating the original finding with whole ROIs of greater location information in lateral than ventral regions, and ruling out ROI size as a factor mediating this effect.

Position-Invariant Category Information

One of the central challenges of object recognition is the ability to identify an object independent of where it appears in the visual field (Ullman, 1996). The neural basis of this ability has been investigated at the level of individual neurons by asking whether the neuron's profile of response across different object categories is preserved despite changes

in the retinal location of the stimulus (Ito et al., 1995). The question can also be asked of population codes across multiple neurons (Hung et al., 2005; Cox, 2007) or voxels (Tong and Kim, VSS talk 2005), using pattern analysis methods. Here we asked whether the category information present in our ROIs is invariant to changes in stimulus position. Note that the presence of location information does not preclude position-invariant category information in an ROI; rather, the same neural pattern can contain both types of information (Hung et al., 2005; DiCarlo and Cox, 2007). Indeed, the fact that all ROIs except LO fail to show an interaction between category and location indicates that these ROIs do not contain significantly more category information when position is held constant (i.e. within locations) than when it is not (i.e. between locations). It is important to note, however, that the statistical independence of location and category information in the pattern analyses conducted here do not imply that the two kinds of information do not interact at the level of individual voxels or neurons.

To further test for position-invariant category information, we compared the amount of category information present when stimuli are displayed in different retinal locations (thus, within-category versus between-category correlations when both are between-location) using 2-tailed, paired t-tests; By this measure, all eight object-and shape-selective ROIs demonstrated position-invariant category information (all $p \leq 0.001$), while the retinotopic control area earlyV did not ($p = 0.98$). See Table 6 for individual p-values from these tests of position-invariant category information.

Just as the responses of a neuron or, on a larger scale, a cortical area can contain position-invariant category information, they can also contain category-invariant position information. As a measure of this, we used greater correlations within- than between-locations when the stimuli were from two different object categories. The results of this

analysis are shown in Table 6. We found significant category-invariant position information in all ROIs (seven, $p < 0.005$) except FBA* ($p = 0.23$) and marginally FFA* ($p = 0.07$).

The results of these analyses indicate that all of the object-selective ROIs contain category information that is independent of stimulus location. This finding is particularly striking given our other analyses showing that nearly all of these regions also contain substantial location information. Representations in these regions are not position-invariant in the strictest sense, because they change with stimulus position. Nonetheless, our data show that the category information represented by the profiles of response in these regions is preserved across changes in stimulus position.

Confirmation of Results with Independent Classification Method

To make sure that our results are not specific to the pattern analysis method we used, we also applied a linear support vector machine (SVM) to our Experimental data set. See Supporting Information for the details of this analysis. Classification performance revealed category information in all ROIs (all nine, $p < 0.005$) and location information in all ROIs (all $p < 0.05$) except PPA ($p = 0.09$). Classification performance for each ROI is shown in Figure 10 and individual p-values are listed in Table 7.

Discussion

Our study provides a broad-based survey of category and location information across functionally defined object-selective regions, as measured by both means and multivariate pattern analyses. A number of important new findings were revealed. First, a substantial amount of information about object location was found in all ROIs except FBA*, even though these ROIs were defined by their selectivity for object shape or category. Second, category and location information are independent of one another in all regions except

LO. Thus, despite the substantial amount of location information in nearly all ROIs, every object-selective ROI demonstrated significant position-invariant category information, in the sense that categories could be discriminated based on the pattern of response across voxels in that ROI even when this analysis was conducted across locations. Finally, we found more location information in the ROIs on the lateral surface (EBA, OFA, TOS, and LO) than in those on the ventral surface (FBA*, FFA*, PPA, and pFs), even though the two surfaces did not differ in the amount of information they contained about object category. These findings bear on a number of questions about the overall organization of the occipitotemporal pathway, which we discuss in turn.

Do Category and Location Information Coexist in Object-Selective Regions?

Although previous studies have shown the presence of location information in some cortical regions with strong selectivity for objects or categories (Op De Beeck and Vogels, 2000; Levy et al., 2001; DiCarlo and Maunsell, 2003; Hasson et al., 2003), our study is the first to show that location information is a systematic property found in nearly all of the known object-selective and category-selective regions in humans. This location information is manifested in most of the ROIs by each of our two independent measures (see Figure 2): i) significant differences in mean response to stimuli presented in different locations (Figure 4), and ii) higher correlations across voxels within than between locations (Figure 8). Category information is also present in all object-selective ROIs by both measures (Figures 4 and 8). Thus, contrary to the strict interpretation of the original "Dual Pathway Model" (Underleider and Mishkin, 1982), category information and location information coexist in object-selective extrastriate cortical regions in humans, including those in the occipitotemporal pathway.

Do Category and Location Information Interact within ROIs?

The finding that information about category and location coexist in the same cortical areas raises the question of whether these types of information interact. This question goes right to the core of our understanding of vision. It is frequently argued that the central problem of vision is object recognition (Ullman, 1996), and the crux of object recognition is solving the problem of invariance, that is, appreciating the sameness of an object despite the different images it casts as it moves across the retina. Segregating information into the "what" and "where" pathways is one way to achieve position-invariance. However, both kinds of information can be represented independently in the same neural population code, in the sense that either kind of information can be easily extracted (with a simple linear classifier) from the same population of neurons (Hung et al., 2005). Further, by keeping category information and location information together in the same neural substrate, it is possible not only to extract position-invariant category information, and category-invariant position information, but also to unite category and location information for perception, as needed to solve the "binding problem" (Treisman, 1996; Cox, 2007) and hence to "know what is where by looking" (Marr, 1982). Thus, the ideal representation would contain in the same neural substrate both position-invariant category information and category-invariant position information.

The ROIs investigated here appear to contain just such an ideal representation. Our pattern analyses showed that categories can be distinguished just as well across locations as within locations in nearly all ROIs, providing striking evidence for position invariance of category information in most of the object-selective regions in our analysis. Further, most ROIs also demonstrated a large amount of location information, which is just as strong within categories as between categories. Thus, analyses of fMRI patterns, like previous, more fine-grained analyses of local neural population codes (Hung et al., 2005), showed that location and category information coexist independently at the population level in

nearly all of the regions of occipitotemporal cortex.

Why do Category-selective Regions Come in Pairs?

A notable feature of extrastriate cortex is that functionally defined category-selective regions seem to come in pairs. This phenomenon has been described for several categories, including bodies, faces, scenes, tools, and shape-selective areas (Grill-Spector et al., 1999; Beauchamp et al., 2002; Hasson et al., 2003; Peelen and Downing, 2005; Schwarzlose et al., 2005). For each category pair, one of these regions is located on the ventral surface of the temporal lobe (FBA*, FFA*, PPA, pFs) while the other is situated on the lateral occipital surface (EBA, OFA, TOS, LO). The reason for this paired organization is not yet understood. Prior studies have shown that specific pairs of regions on the two surfaces differ in their sensitivity to features such as motion (Beauchamp et al., 2002; Hasson et al., 2003), eccentricity (Levy et al., 2001), size and location (Grill-Spector et al., 1999), and object completeness (Taylor et al., 2007). However, each of these studies tested only a small subset of object-selective regions. Our study, which systematically examined location information across a large set of object-selective ROIs using pattern analyses, found that lateral regions contain substantially more location information than do ventral regions, despite having equal amounts of category information. This systematic difference in the amount of location information between the two surfaces provides a preliminary clue of how the two surfaces differ in the representations they contain and computations they perform.

Why Are Combinations of Category and Location Selectivity Consistent across ROIs?

Prior studies have reported that scene-selective areas show a higher response to stimuli in the periphery, whereas face-selective regions show a higher response to foveal stimuli (Levy et al., 2001; Hasson et al., 2003). The studies have proposed that these particular

combinations of selectivities reflect the different computational requirements for processing each category: large-scale integration for scenes and fine-grained acuity for faces (Levy et al., 2001; Malach et al., 2002). Although the results of our study generally replicate the peripheral preference of scene-selective regions and the foveal preference of face-selective regions, we also found biases for elevation in some of these and other ROIs. For example, both scene-selective areas, PPA and TOS, responded more strongly to upper than lower visual field locations, even though stimulus eccentricities were matched. Similarly, the EBA preferred lower visual field stimuli to both foveal and upper visual field stimuli, which activated the region equally. These findings, like earlier reports of contralateral biases in object-selective regions (Niemeier et al., 2005; Hemond et al., 2007) do not fit within the fovea/periphery framework, and it is not clear how the computational demands hypothesis (Malach et al., 2002) could account for them.

An alternate explanation for consistent combinations of category and location selectivities appeals instead to the statistics of experience (Kanwisher, 2001): for example, to the extent that humans naturally tend to foveate faces (Malach et al., 2002), the foveal bias in face-selective areas and the lower visual field bias in the body-selective area EBA may reflect the locations where these stimuli are typically seen in daily life. Perhaps regions of cortex with a pre-existing category selectivity develop location preferences corresponding to the retinal location where that object is typically seen. Alternately, location biases might arise first in the cortex, with category selectivities arising in those regions of cortex already biased toward the location where that object typically occurs. Note, however, that the difference in location biases in the body-selective areas EBA and FBA* suggests that pairings of category and location selectivity are not perfectly consistent across ROIs. Thus, while experiential statistics would seem to provide a better account of the specific combinations of category and location selectivities than do computational requirements, neither can

account for all the data.

Further questions

While the results of this study yield several new insights about the relationship of location and category information in extrastriate cortex, they also raise many new questions. How precise is the location information contained in these category-selective areas? Here we sampled only three locations, many degrees apart; it is unclear how many different locations these ROIs can discriminate and whether such finer-grained location information can be detected with fMRI (at the present or perhaps higher spatial resolution). Second, does the location information reported here reflect retinotopic location or absolute location independent of eye position? Third, is the location information revealed in this study epiphenomenal, or does it contribute to perception and behavior (Williams et al., 2007)? Fourth, does the location information reflect retinotopic organization within these regions? The correlational analyses used here are blind to the adjacencies of voxels and so cannot answer this question, however the apparent overlap of some of our object-selective regions with retinotopic visual areas is suggestive. Finally, will the information in each ROI reported here be recoverable when participants view complex scenes containing multiple objects (Reddy and Kanwisher, 2007)? Regardless of how these questions are ultimately resolved, the present study provides the foundation for a better understanding of the information content, and hence the function, of each of the major object-selective regions in the occipitotemporal pathway.

Supporting Information

Analysis of Eye Position

Eye position was monitored in the scanner for five of the nine participants whose data

were subsequently included in the study. To test whether eye position (elevation) may have varied with stimulus location or category, we first ran a group 3 x 4 (location x category) repeated-measures ANOVA on the data from all five participants and found no significant effects of stimulus location (foveal, upper, or lower) or stimulus category (bodies, cars, faces, and scenes) on vertical eye position, nor an interaction between the two (all three, $p > 0.50$). Since the group ANOVA would not detect any condition-dependent eye position changes that are not systematic across participants, we also analyzed the eye position data from each individual separately. For each of the five participants, we ran a separate 3 x 4 (category x location) repeated-measures ANOVA on mean eye position for each of the individual stimulus blocks in the Experimental runs from that participant's scan session. In each of the five ANOVAs, there were no main effects of location (all five, $p > 0.65$) or category (all five, $p > 0.75$), as well as no interactions between them (all five, $p > 0.50$).

Pattern Analysis on Upper and Lower Stimulus Locations Only

Since the foveal stimuli in this study were smaller than the peripheral stimuli, the inclusion of data from all three retinal locations confounds location information with size information. To test whether object-selective ROIs demonstrate pattern information about location even when size is kept equal, we excluded correlations from all foveal conditions, thereby analyzing correlations from only upper and lower stimulus locations. The results of this analysis can be seen in Table 3 and Figure 9. They show the same pattern of results as the original analysis. Specifically, all ROIs demonstrate category information (eight, $p < 0.005$) except earlyV ($p = 0.62$). Furthermore, all ROIs except FBA* demonstrate location information (FBA* $p = 0.73$; all other $p < 0.05$). None of the ROIs show a significant interaction between category and location (all nine, $p \geq 0.10$). In this analysis, LO shows no interaction between category and location information, which is unlike the results of the

primary analysis. However, overall these results support the main findings from the primary analysis, namely that location information exists in nearly all object-selective ROIs and that this location information is independent of category information in these regions.

Test of Effect of ROI Size

Our pattern analyses revealed more location information in lateral than ventral ROIs, however lateral regions are also typically larger than their counterparts on the ventral surface. For example, the ventral scene-selective region PPA comprised 68 voxels on average, whereas the lateral scene-selective region TOS averaged 160 voxels. Given this size disparity, it is possible that the difference in voxel count, rather than an actual difference in location information, could have caused the apparent difference in location information between lateral and ventral surfaces. In order to test this hypothesis, we created and analyzed small ROIs at the center of the standard PPA and TOS ROIs used in the study. Specifically, in each ROI we selected fifteen voxels contiguous with the most selective voxel (the voxel with the smallest p-value in the scenes-objects Localizer contrast) in a single slice. Two of the PPA ROIs included in the prior 2 x 2 x 2 ANOVA of whole ROIs failed to contain fifteen voxels in the same slice as the peak voxel; Therefore, those PPA ROIs and the corresponding TOS ROIs were excluded from the present analysis, leaving seven 15-voxel PPAs and seven 15-voxel TOSs to be analyzed. To test whether the differences in surfaces persisted after voxel number was equated, we ran the same repeated-measures 2 x 2 x 2 ANOVA on cortical surface, category, and location for these 15-voxel control ROIs. The results are shown in Table 5. In keeping with our findings from the ANOVA on whole PPA and TOS ROIs, this ANOVA revealed significant main effects of category and location (both, $p < 0.05$), as well as a significant interaction of surface and location ($p < 0.05$), without a corresponding surface by category interaction ($p = 0.22$). These results indicate

that the greater amount of location information found in lateral (TOS) than ventral (PPA) surfaces with whole ROIs does not arise due to differences in ROI size.

Confirmation of Pattern Analysis Results with a Classification Method

To test whether both location and category information can be detected in the ROIs of this study by a different multivariate method, we reanalyzed our data with a linear support vector classifier using the OSU SVM toolbox based on the LIBSVM package. The twelve conditions were grouped either by location or category, depending on the feature to be classified. SVMs were then conducted using a leave-one-out design, such that they were trained on data from all but one run, then tested on the remaining run, a process that was iterated so that each run served as test data only once. The SVM results were then computed as the mean performance across each of these train-then-test iterations. All SVM training and testing was conducted on mean voxelwise responses across blocks, rather than individual stimulus presentations. Classification amongst multiple classes was based on a series of binary classifications between each pair of classes, and the 'winning' class was determined using a basic voting mechanism. Single-sample two-tailed t-tests run on classifier performance in each ROI showed that accuracy was significantly above chance for classification of category (all nine, $p < 0.000001$) and location (all nine, $p < 0.02$) in each ROI, confirming our findings using pattern analyses. Classification performance for each ROI is shown in Figure 10 and individual p-values are shown in Table 7.

References

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

Brewer AA, Liu J, Wade AR, Wandell BA (2005) Visual field maps and stimulus selectivity in human ventral occipital cortex. Nat Neurosci 8:1102-1109.

Cox DD (2007) Reverse Engineering Object Recognition. Dissertation: Massachusetts Institute of Technology.

Cox R, Jesmanowicz A (1999) Real-time 3D image registration for functional MRI. Magn Reson Med 42:1014-1018.

DiCarlo JJ, Maunsell JH (2003) Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. J Neurophysiol 89:3264-3278.

DiCarlo JJ, Cox DD (2007) Untangling invariant object recognition. Trends Cogn Sci 11:333-341.

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

Duncan RO, Boynton GM (2003) Cortical magnification within human primary visual cortex correlates with acuity thresholds. Neuron 38:659-671.

Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. Trends Neurosci 15:20-25.

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.

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.

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

Hemond CC, Kanwisher N, Op de Beeck HP (2007) A preference for contralateral stimuli in human object- and face-selective cortex. PLoS ONE 2.

Hung CP, Kreiman G, Poggio T, DiCarlo JJ (2005) Fast readout of object identity from macaque inferior temporal cortex. Science 310:863-866.

Ishai A, Ungerleider LG, Martin A, Schouten JL, Haxby JV (1999) Distributed representation of objects in the human ventral visual pathway. Proc Natl Acad Sci U S A 96:9379-9384.

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.

Kanwisher N (2001) Faces and places: of central (and peripheral) interest. Nat Neurosci 4:455-456.

Kourtzi Z, Kanwisher N (2000) Cortical regions involved in perceiving object shape. J Neurosci 20:3310-3318.

Larsson J, Heeger DJ (2006) Two retinotopic visual areas in human lateral occipital cortex. J Neurosci 2:13128-13142.

Levy I, Hasson U, Harel M, Malach R (2004) Functional analysis of the periphery effect in human building related areas. Hum Brain Mapp 22:15-26.

Levy I, Hasson U, Avidan G, Hendler T, Malach R (2001) Center-periphery organization of human object areas. Nat Neurosci 4:533-539.

Macevoy SP, Epstein RA (2007) Position selectivity in scene- and object-responsive occipitemporal regions. J Neurophysiol.

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

Marr D (1982) Vision: A Computational Investigation into the Human Representation and Processing of Visual Information: WH Freeman, New York, NY.

McKyton A, Zohary E (2007) Beyond retinotopic mapping: the spatial representation of objects in the human lateral occipital complex. Cereb Cortex 17:1164-1172.

Merigan WH, Maunsell JH (1993) How parallel are the primate visual pathways? Ann Rev Neurosci 16:369-402.

Niemeier M, Goltz HC, Kuchinad A, Tweed DB, Vilis T (2005) A contralateral preference in the lateral occipital area: sensory and attentional mechanisms. Cereb Cortex 15:325-331.

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'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.

Op De Beeck H, Vogels R (2000) Spatial sensitivity of macaque inferior temporal neurons. J Comp Neurol 426:505-518.

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

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

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

Sereno AB, Maunsell JH (1998) Shape selectivity in primate lateral intraparietal cortex. Nature 395:500-503.

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

Swisher JD, Halko MA, Merabet LB, McMains SA, Somers DC (2007) Visual topography of human intraparietal sulcus. J Neurosci 27:5326-5337.

Taylor JC, Wiggett AJ, Downing PE (2007) fMRI analysis of body and body part representations in the extrastriate and fusiform body areas. J Neurophysiol.

Treisman A (1996) The binding problem. Curr Opin Neurobiol 6:171-178.

Ullman S (1996) High-level vision: object recognition and visual cognition. Cambridge, MA: The MIT Press.

Underleider LG, Mishkin M (1982) Two Cortical Visual Systems. In: Analysis of Visual Behavior (Ingle MA, Goodale MI, Masfield RJW, eds), pp 549-586. Cambridge, MA: MIT Press.

Williams MA, Dang S, Kanwisher NG (2007) Only some spatial patterns of fMRI response are read out in task performance. Nat Neurosci 10:685-686.

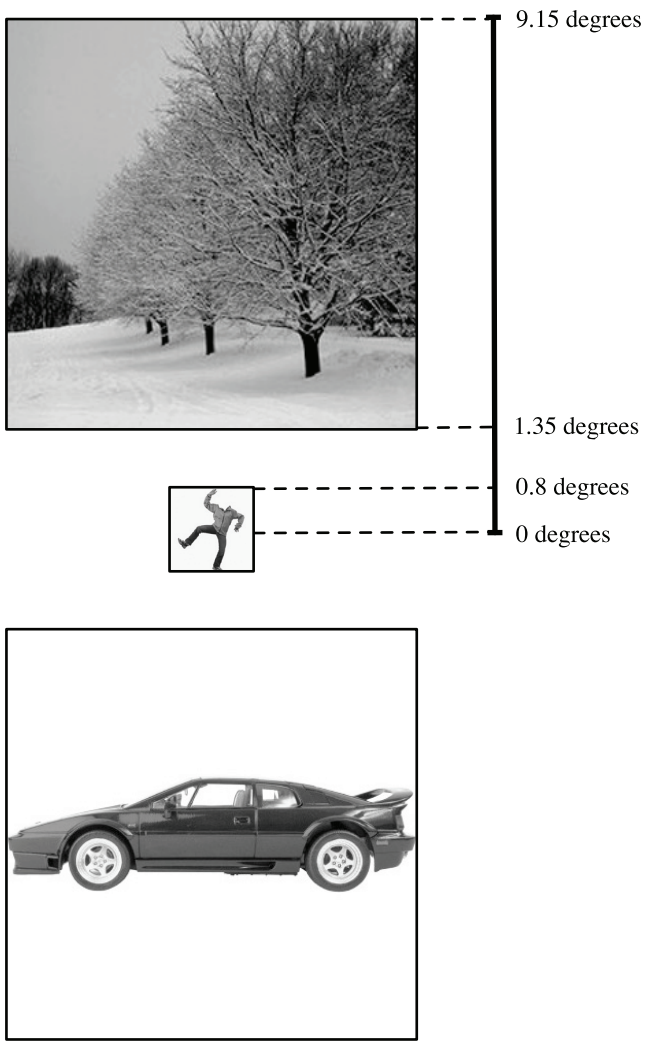

Figure 1: Schematic diagram of Experimental stimuli.

In the actual experiment, only one image was presented at a time. Stimuli were presented in blocks of twenty images, with object category and location kept constant within each block. All four object categories (bodies, cars, faces, or scenes) were presented equally often in all three retinal locations (above, below, or at fixation), to yield twelve Experimental conditions.

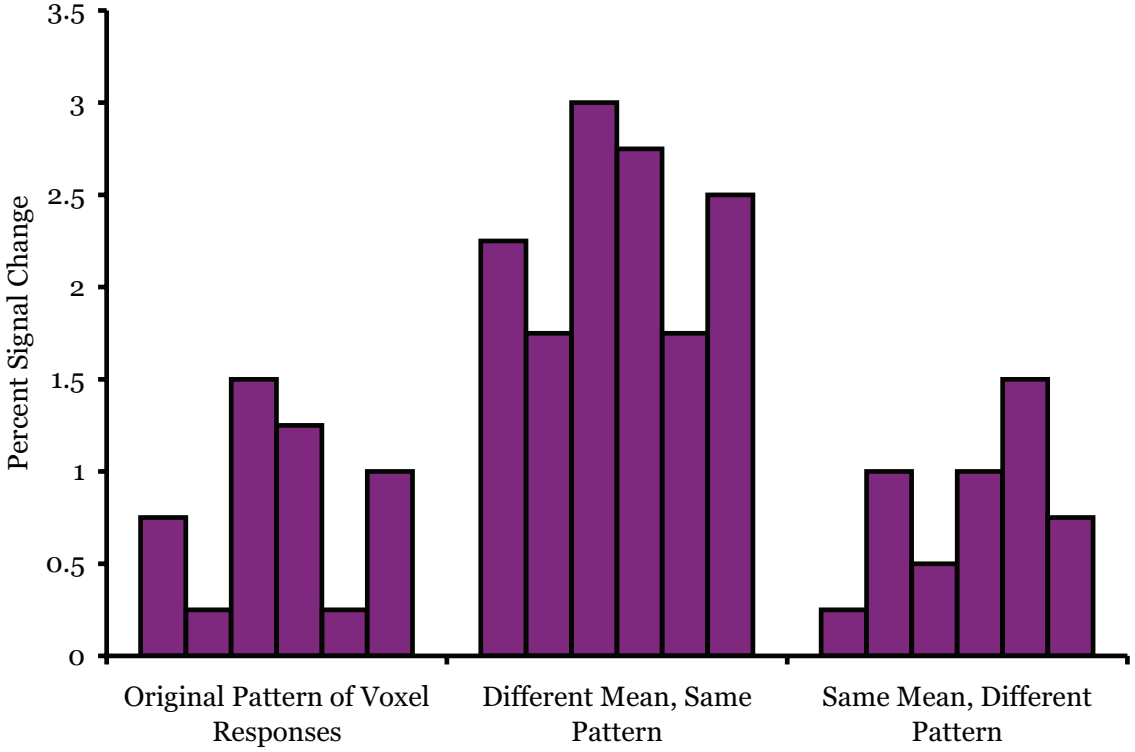

Figure 2: Schematic depiction of the independence between mean and pattern measures of information.

Each bar represents the magnitude of response of a single voxel in a fictitious ROI. The first set of bars depicts a hypothetical pattern of response across the voxels in the ROI to a given condition. The second set of bars depicts the pattern of response of those same voxels to a different condition in the case where the responses of all voxels are uniformly increased by the same amount. In this case the means will be substantially different for these two conditions, suggesting the presence of information discriminating the two using a population mean response, however the correlations between them will be high, demonstrating little discriminating information using pattern analyses. A comparison of the conditions of the first set of bars with the third demonstrates the opposite result; These two conditions have the same mean response, indicating a lack of information using means, however their patterns are substantially different, suggesting that information discriminating between the two conditions may exist in the patterns of response.

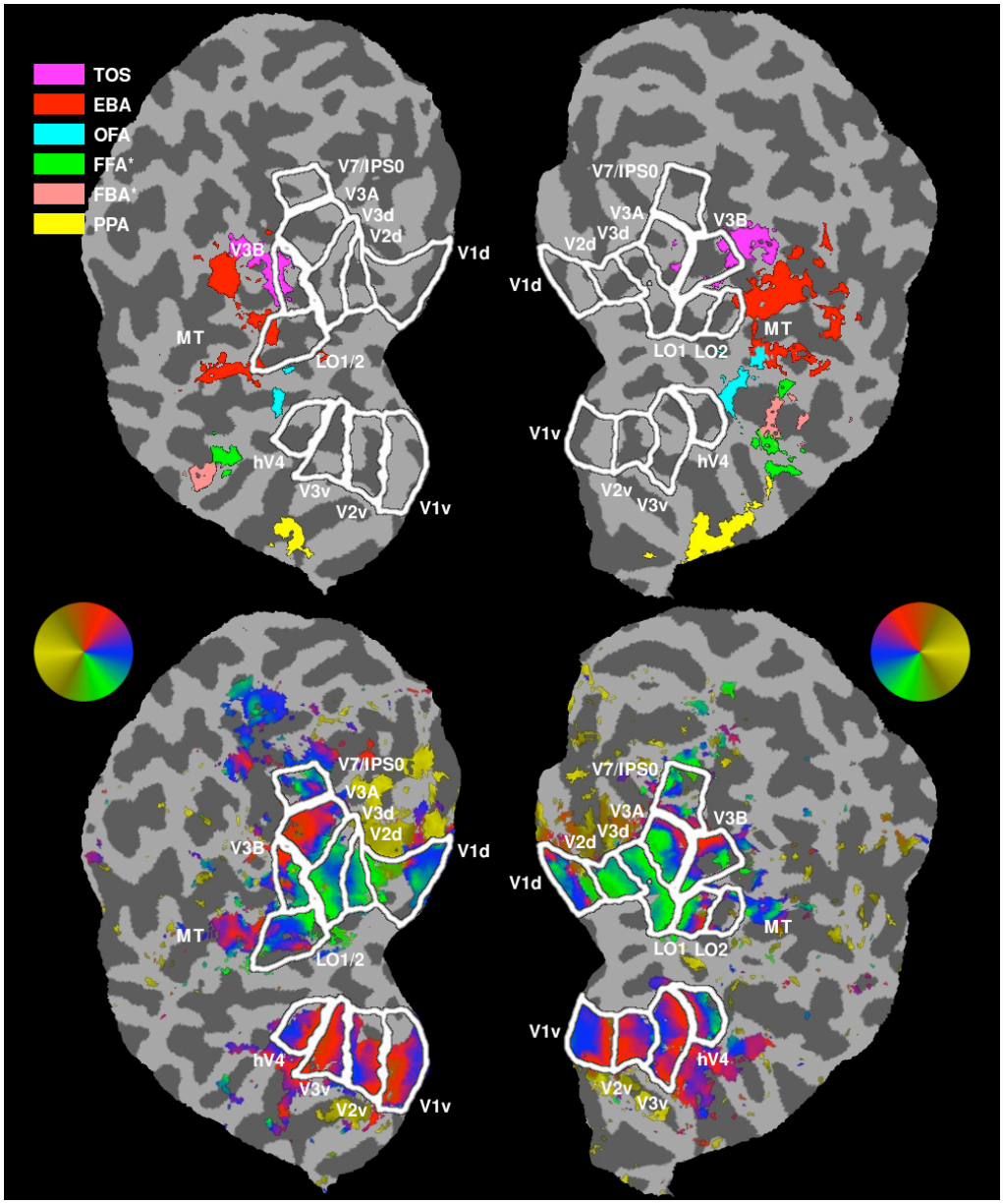

Figure 3: Retinotopic mapping and object-selective ROIs.

Activations are shown here on bilateral flattened occipital surfaces from a single representative participant. White outlines represent boundaries of identifiable retinotopic visual areas, which are shown overlaid upon object-selective ROIs (top), and retinotopic maps of polar angle (bottom). The occipital pole was not included in the slice prescription.

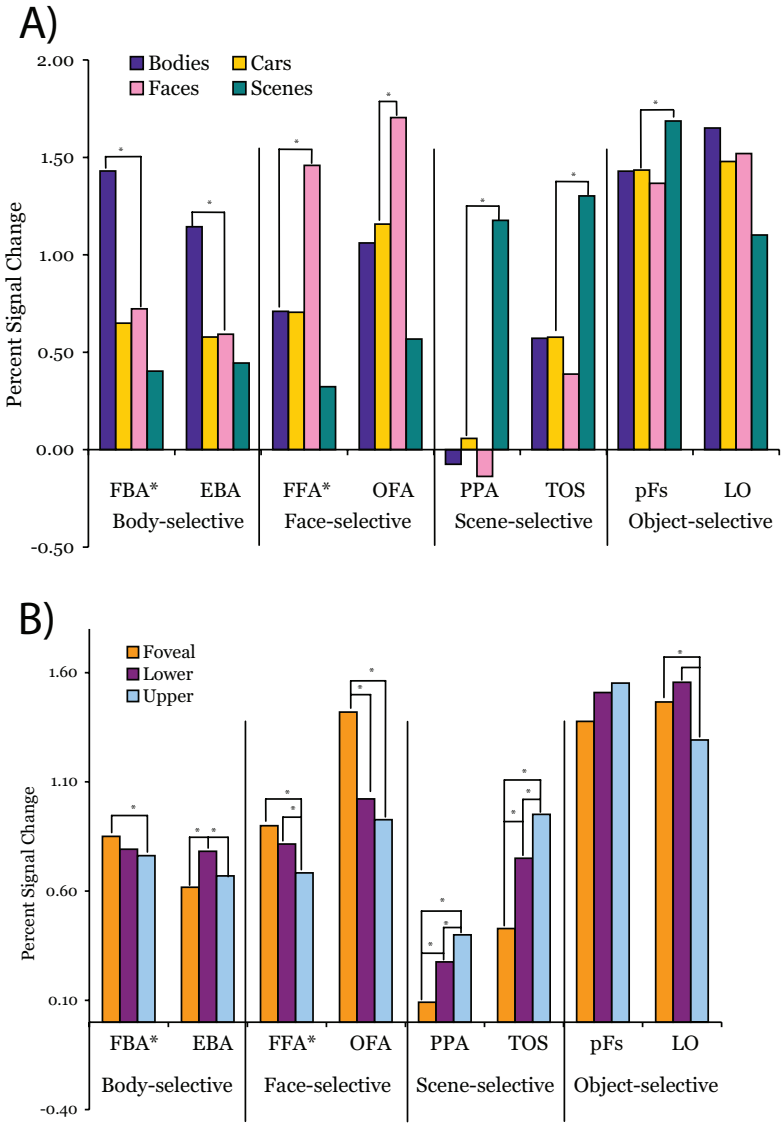

Figure 4: Mean response magnitude in each object-selective ROI.

A: Mean responses across each ROI to each of the four object categories averaged across the three retinal locations. Significant differences ($p < 0.05$) in individual contrasts are indicated with a star (here shown only between the highest and second highest responses). All ROIs defined by body-, face-, or scene-selectivity show significantly greater responses to their preferred category than to all other categories. B: Mean responses across each ROI to each of the three stimulus locations averaged across all four object categories. Significant differences in mean response to different locations were found in nearly all ROIs and demonstrate that these regions contain location information.

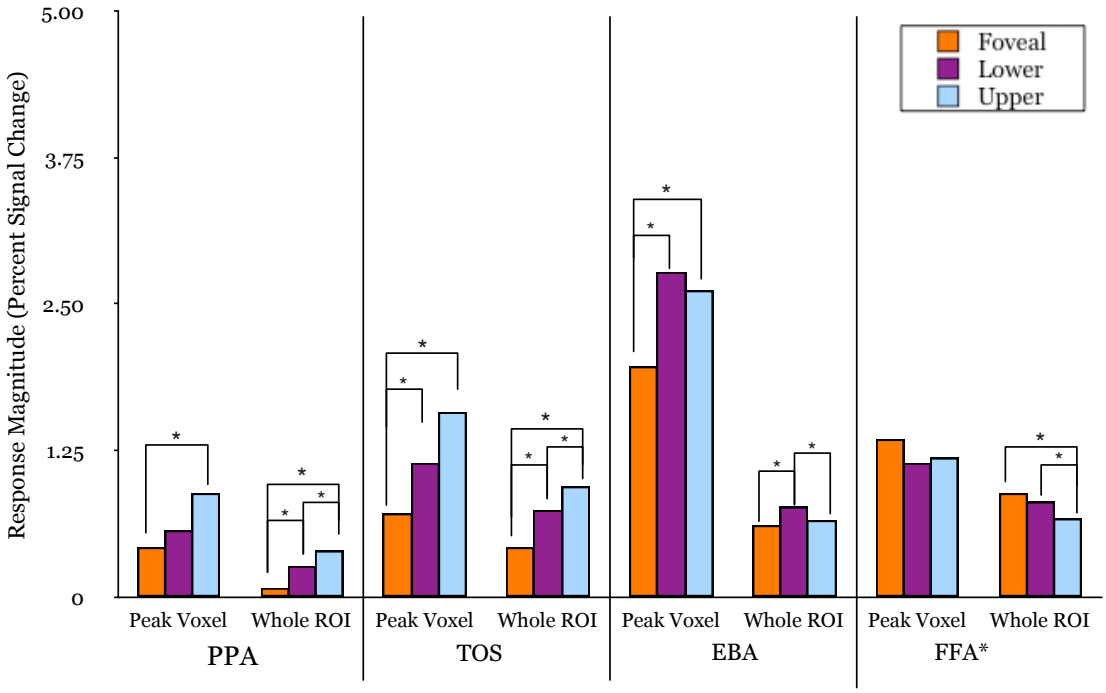

Figure 5: Response magnitude in peak voxels.

Responses in the single peak voxels of four ROIs to each of the stimulus locations are shown here after averaging across all object categories. The mean responses in the whole ROIs are shown next to the peak voxel responses for the sake of comparison. Brackets with stars indicate statistically significant differences at $p < 0.05$. In all four ROIs, the location eliciting the highest response is the same for the single peak voxel as for the whole ROI.

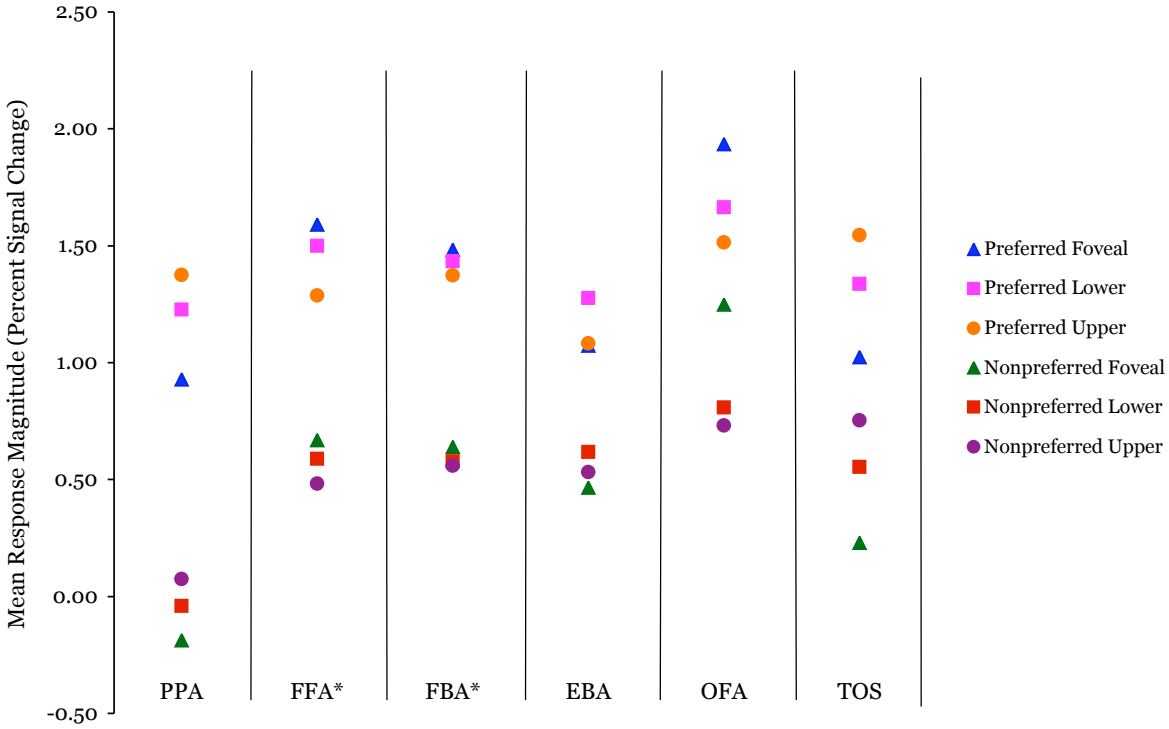

Figure 6: Location biases for preferred and nonpreferred categories.

Mean response magnitudes for the three stimulus locations are shown for all six category-selective ROIs. Mean responses by location are plotted for the preferred category and for the average of the three nonpreferred categories. Triangles represent the response to foveal stimuli, squares represent the response for lower visual field stimuli, and circles denote the response to upper field stimuli. For each of the six ROIs, the rank serial order of location biases is the same for both preferred and nonpreferred stimuli.

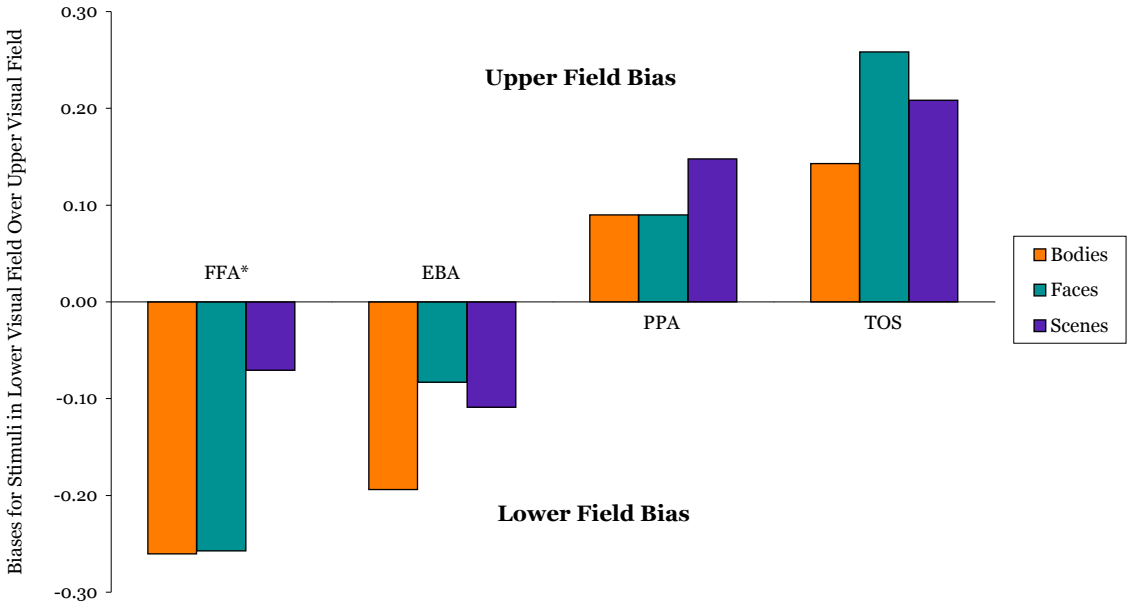

Figure 7: Direction of elevation biases by ROI and stimulus category.

An indicator of elevation bias direction, calculated as the difference between the mean ROI response to upper field stimuli and lower field stimuli, is plotted for each of the four ROIs that demonstrated significant elevation biases in the main analysis. A positive value indicates that a particular ROI produced a greater response to the stimulus category when it was in the upper visual field than in the lower field, while a negative result indicates the reverse. This figure demonstrates that elevation biases are properties inherent to ROIs, not to stimulus categories.

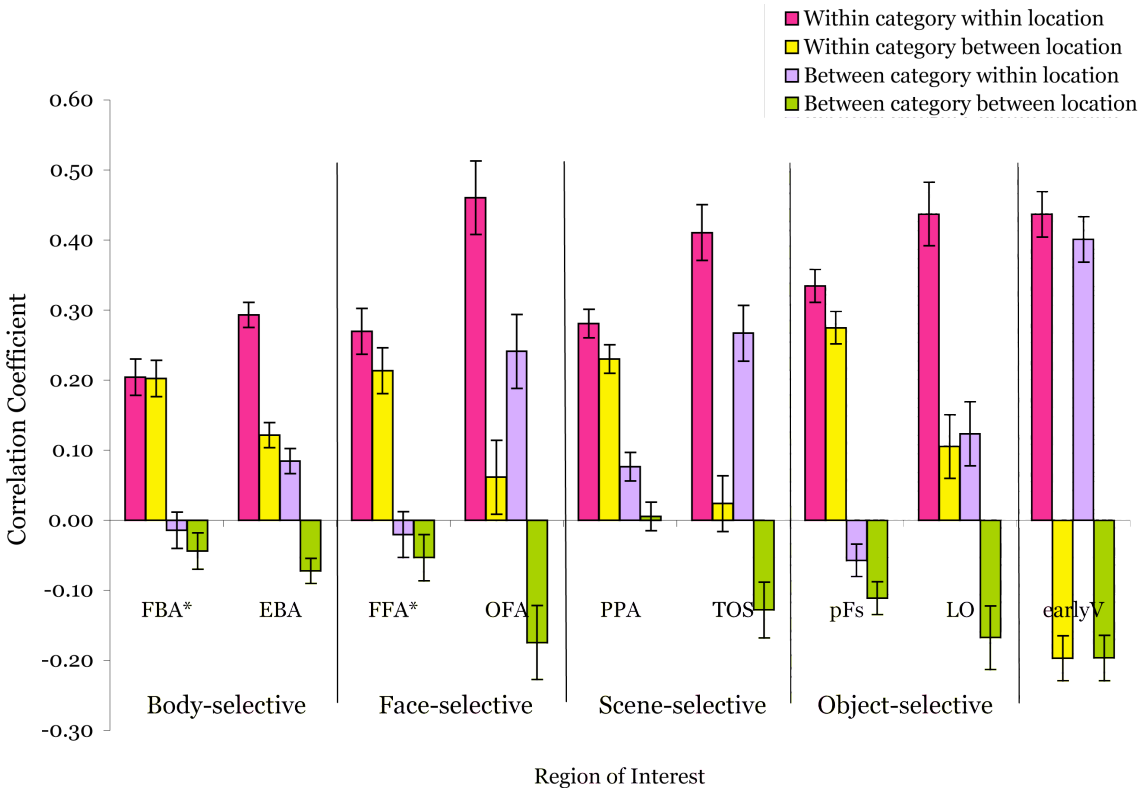

Figure 8: Voxelwise pattern information as demonstrated by average correlations across voxels in each ROI.

These correlations are plotted as a function of whether the two response patterns are from the same ("within") or different ("between") categories, and from the same ("within") or different ("between") locations. They reveal that nearly all ROIs demonstrate both category and location information.

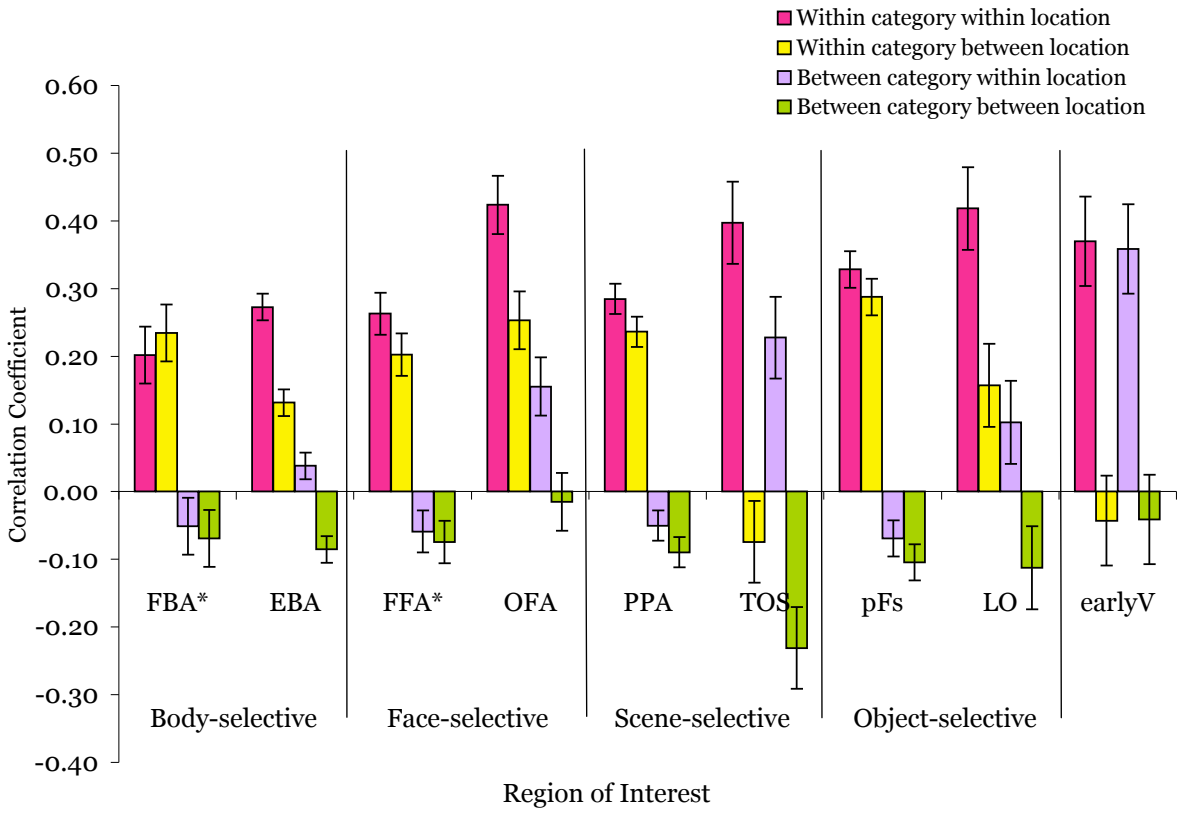

Figure 9: Voxelwise pattern information from only upper or lower stimulus locations. The corresponding statistical significances are listed in Table 3. This analysis demonstrates that location information persists after size confounds have been removed.

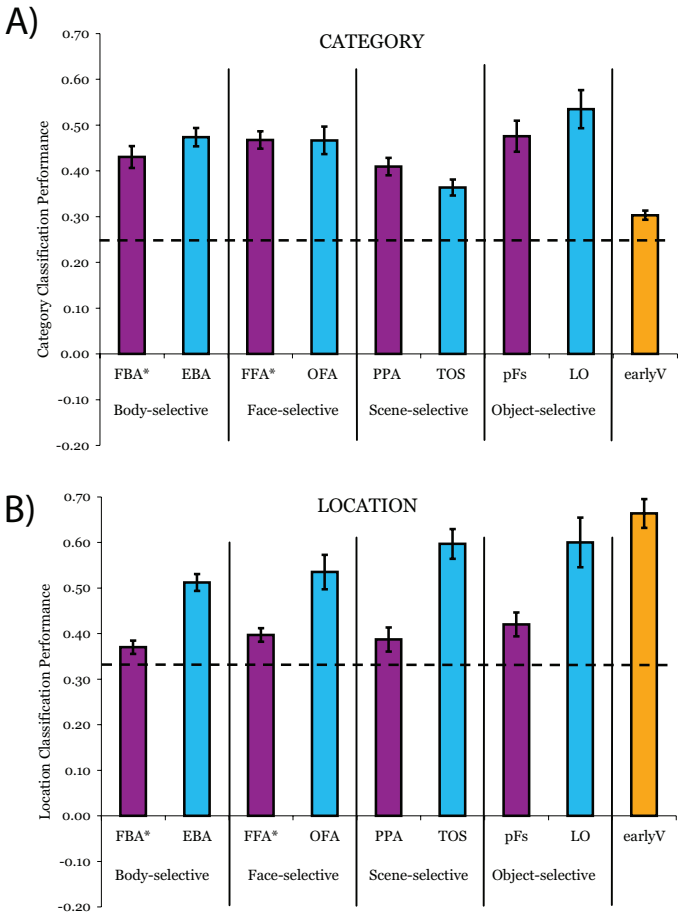

Figure 10: Support vector machine classification performance.

SVM classification accuracy on category (a) and location (b) are shown for each ROI. Chance performance is denoted with a dashed black line. Purple bars represent ventral ROIs while blue bars represent lateral ROIs. The retinotopic control region earlyV is shown in orange. These analyses demonstrate that both category and location information exist in object-selective ROIs, replicating the results obtained with correlations.

	Eccentricity		Elevation
	Foveal versus Lower	Foveal vs Upper	Lower vs Upper
FBA*	F>L (0.09)	**F>U (0.04)**	L>U (0.08)
EBA	**L>F (0.0004)**	U>F (0.29)	**L>U (0.0007)**
FFA*	F>L (0.07)	**F>U (0.00008)**	**L>U (0.003)**
OFA	**F>L (0.04)**	**F>U (0.007)**	L>U (0.15)
PPA	**L>F (0.005)**	**U>F (0.003)**	**U>L (0.02)**
TOS	**L>F (0.008)**	**U>F (0.0001)**	**U>L (0.03)**
pFs	L>F (0.19)	U>F (0.14)	U>L (0.78)
LO	L>F (0.31)	**F>U (0.02)**	**L>U (0.001)**

Table 1: Results of two-tailed paired t-tests on mean response magnitude in each object-selective ROI for each pair-wise comparison between locations that differ either in eccentricity or elevation. Directions of differences in means are indicated using the symbols F (foveal), L (lower), and U (upper), and p-values are shown in parentheses. Bold lettering indicates comparisons that show significant differences. Corresponding mean response magnitudes are shown in Figure 4.

	n	Category	Location	Category x Location
FBA*	7	39.9 (0.0007)	0.82 (0.40)	0.34 (0.58)
EBA	18	121.8 ($4x10^{-9}$)	44.5 ($4x10^{-6}$)	3.6 (0.08)
FFA*	12	28.5 (0.0002)	5.0 (0.047)	1.6 (0.24)
OFA	13	26.6 (0.0002)	26.5 (0.0002)	0.53 (0.48)
PPA	10	48.5 ($7x10^{-5}$)	10.8 (0.009)	2.1 (0.18)
TOS	12	41.4 ($5x10^{-5}$)	36.9 ($8x10^{-5}$)	0.2 (0.70)
pFs	11	79.8 ($4x10^{-6}$)	119.5 ($7x10^{-7}$)	0.32 (0.58)
LO	11	40.3 ($8x10^{-5}$)	23.5 (0.0007)	21.2 (0.001)
earlyV	7	2.9 (0.14)	136.3 ($2x10^{-5}$)	3.2 (0.13)

Table 2: Results of repeated-measures 2 x 2 ANOVAs on correlations across voxels in each ROI between odd and even data sets as a function of Category (within versus between) and Location (within versus between). Results are shown as F-values with p-values in parentheses. Nearly all ROIs show a significant amount of information about both category and location, yet fail to show a significant interaction between the two.

	n	Category	Location	Category x Location
FBA*	7	21.6 (0.004)	0.13 (0.73)	1.1 (0.33)
EBA	18	69.0 ($2x10^{-7}$)	23.3 (0.0002)	2.3 (0.15)
FFA*	12	34.9 (0.0001)	5.8 (0.04)	3.3 (0.10)
OFA	13	35.7 ($6x10^{-5}$)	10.0 (0.008)	0.004 (0.95)
PPA	10	90.5 ($5x10^{-6}$)	5.5 (0.04)	0.12 (0.74)
TOS	12	34.2 (0.0001)	21.8 (0.0007)	0.23 (0.64)
pFs	11	105.2 ($6x10^{-7}$)	11.5 (0.006)	0.35 (0.57)
LO	11	32.6 (0.0002)	7.2 (0.02)	1.5 (0.25)
earlyV	7	0.28 (0.62)	12.6 (0.01)	0.33 (0.59)

Table 3: Pattern analysis results after exclusion of foveal stimulus conditions.

Repeated-measures Category x Location ANOVAs were computed with correlations from only upper and lower stimulus locations to eliminate size confounds. The results of this analysis, shown here as F-values with p-values in parentheses, were similar to those from the prior analysis on all three locations, namely that most object-selective ROIs showed significant category and location information, but none showed a significant interaction between the two.

	Bodies FBA*/EBA	Faces FFA*/OFA	Scenes PPA/TOS	Objects pFs/LO
n	7	10	9	11
Surface	1.3 (0.30)	5.3 (0.047)	0.01 (0.93)	1.6 (0.24)
Category	62.1 (0.0002)	24.0 (0.0009)	54.6 ($8x10^{-5}$)	96.1 ($2x10^{-6}$)
Location	11.6 (0.02)	30.8 (0.0004)	45.1 (0.0002)	31.5 (0.0002)
Surface x Category	0.32 (0.59)	2.2 (0.17)	2.6 (0.14)	2.9 (0.12)
Surface x Location	8.9 (0.03)	21.3 (0.001)	17.7 (0.003)	16.0 (0.003)
Category x Location	0.14 (0.72)	0.02 (0.90)	1.9 (0.21)	12.6 (0.005)
Surface x Category x Location	.88 (0.38)	5.5 (0.04)	0.48 (0.51)	7.4 (0.02)

Table 4: Repeated-measures 2 x 2 x 2 ANOVA of pattern information with Surface (ventral or lateral), Category (within or between), and Location (within or between) as factors. Ventral and lateral ROIs were paired by their defining contrast (e.g., scenes > objects for PPA and TOS). Interactions between Surface and Location in all four ROI pairs demonstrate that lateral regions contain significantly more location information than do their ventral counterparts.

	Whole PPA/TOS	15-voxel PPA/TOS
n	9	7
Surface	0.01 (0.93)	7.6 (0.03)
Category	54.6 ($8x10^{-5}$)	196.3 ($8x10^{-6}$)
Location	45.1 (0.0002)	18.0 (0.005)
Surface x Category	2.6 (0.14)	1.9 (0.22)
Surface x Location	17.7 (0.003)	6.3 (0.046)
Category x Location	1.9 (0.21)	0.02 (0.90)
Surface x Category x Location	0.48 (0.51)	5.3 (0.06)

Table 5: Pattern analysis results after equating ROI size. Repeated-measures Surface x Category x Location ANOVAs were run on the standard ('whole') PPA and TOS ROIs used in the study, as well as on small 15-voxel control regions comprised of the peak voxel and fourteen contiguous voxels in each ROI. Results are shown here as F-values with p-values in parentheses. Both ANOVAs revealed a significant Surface x Location interaction, demonstrating that differences in ROI size do not account for the greater amount of location information found in lateral than ventral regions.

	Position-invariant Category Information	Category-invariant Position Information
FBA*	0.001	0.23
EBA	0.00000001	0.000007
FFA*	0.0001	0.07
OFA	0.00009	0.0002
PPA	0.00002	0.003
TOS	0.00003	0.00006
pFs	0.000001	0.00002
LO	0.0002	0.0009
earlyV	0.98	0.00004

Table 6: Results of two-tailed paired t-tests (shown here as p-values) from voxelwise pattern measures of position-invariant category information and category-invariant position information. Higher correlations within than between categories when both are between locations are used as a measure of position-invariant category information. Conversely, higher correlations within than between locations when both are between categories serve as a measure of category-invariant position information. The results show that all object-selective regions exhibit position-invariant category information and most also show category-invariant position information.

	Category Classification	Location Classification
FBA*	0.0002	0.048
EBA	$3x10^{-9}$	$3x10^{-8}$
FFA*	$2x10^{-7}$	0.002
OFA	$9x10^{-6}$	0.0002
PPA	$2x10^{-5}$	0.09
TOS	$5x10^{-5}$	$6x10^{-6}$
pFs	$4x10^{-5}$	0.007
LO	$4\ x10^{-5}$	0.0006
earlyV	0.002	$4x10^{-5}$

Table 7: SVM classification performance. Single-sample two-tailed t-tests were used to compare category classification to chance (0.25) and, separately, to compare location classification to chance (0.33). Results, shown here as p-values, support the prior findings with pattern analyses that nearly all regions contain information about both category and location.

6. Conclusion: Maps in Object-Selective Cortex

The experiments described in this thesis address the question of how category-selective regions are topographically organized in ventral visual cortex (VVC). In particular, we tested several hypotheses about what principles, if any, guide the layout of these regions: Is the spatial arrangement of selectivities in the VVC organized in terms of meaning, temporal association, amount of location information, object shape, computational demands, or the spatial location where stimuli are most frequently seen? Our findings support some of these hypotheses, provide evidence against others, and reveal unexpected findings that raise new questions for future study. Overall, these findings provide preliminary evidence that the layout of VVC may be determined by systematic organizing dimensions and suggest that cortical maps are found not only for low-level sensory and motor processing, but also for higher-level cognitive functions.

A Test of Modularity in VVC

In Chapter 2, we addressed the question of whether category-selective regions in VVC

are domain-specific using the test case of selectivities for faces and bodies on the fusiform gyrus. We scanned at high spatial resolution and found that these two selectivities could be fully dissociated, supporting the hypothesis that selectivities for specific object categories are distinct in VVC. Nevertheless, we found that these two selectivities were consistently located adjacent to one another across all participants, which also suggests that stable, systematic principles may guide the spatial layout of category-selective region in VVC.

Evidence Against Hypotheses of VVC Organization

In Chapters 3 and 4 we tested two hypotheses, either of which could account for the adjacency of face- and body selectivity in VVC. The first of these hypotheses claims that the VVC is organized based on the strength of temporal associations that develop between object categories as a result of the statistical properties of everyday visual experience (Chapter 3). Since eyeglasses should be temporally associated with faces, we tested whether pictures of glasses would activate cortex in or around the face-selective region FFA. Even though we did find cortical regions that responded quite strongly to glasses, these were not in or around the FFA, which argues against a robust, large-scale organization of VVC based on temporal association.

The second hypothesis, which was tested in Chapter 4, states that object-selective regions of VVC are organized according to an animate-inanimate distinction. We found that the body-selective region EBA produced a greater response to a subset of inanimate stimuli (glasses, bicycles, and shoes) than to another animate object category (faces). This example runs contrary to the hypothesis that object selectivities in VVC are strictly segregated based on animacy.

It is important to note that a failure to find evidence for an organizing dimension could be due to several factors, including insufficient sample sizes or limitations in the spatial

resolution or sensitivity of fMRI. Moreover, the questions addressed in this thesis relate to the large-scale organization in VVC, rather than the response properties of individual neurons, which cannot be measured directly using fMRI.

Evidence Supporting Hypotheses of VVC Organization

Although the results of Chapter 4 argued against a strict division of category selectivities based on the animacy of objects, an unexpected result in this study led us to propose a different hypothesis for the type of property that might be represented in object-selective extrastriate cortex. The overlapping selectivities we found for body parts, glasses, bicycles, and shoes in VVC could be explained if this region of cortex is sensitive to specific motor actions associated with objects. The hypothesis that motor actions associated with objects may affect the representation of those objects in VVC is consistent with prior studies that found that cortex in and around the EBA is selective for the performance of object-directed motor actions (Astafiev et al., 2004; Peelen and Downing, 2005), and that motion associated with objects can drive the motion-selective area MT to respond to stationary presentations of those objects as well (Kourtzi and Kanwisher, 2000; Schlack and Albright, 2007; Weisberg et al., 2007). Further studies will be necessary to determine whether one or more regions within VVC may be sensitive to learned motor-object associations.

In Chapter 5, we tested the established theory that the category-selective regions in VVC adhere to a center/periphery organization based on the information and computations required to recognize objects from those categories (i.e., large-scale integration for scenes and fine-grained acuity for faces) (Levy et al., 2001; Hasson et al., 2002; Malach et al., 2002). We tested this theory using different stimuli that allowed us to compare different retinal locations at equal eccentricities. Although our results replicated the eccentricity biases found in prior studies (Levy et al., 2001; Hasson et al., 2002; Hasson et al., 2003), we

also found that most category-selective regions also demonstrated biases for stimuli at equal eccentricities (specifically, locations above and below fixation.) These findings indicate that the location biases found in these regions may not be caused by computational demands. Instead, they are compatible with a different explanation: that category-selective regions also contain biases for stimuli in the retinal locations where they are typically seen in one's visual experience. For example, responses in the body-selective EBA were higher for stimuli presented in the lower than upper visual field, consistent with the likelihood that bodies are more frequently seen in the lower visual field because of the tendency to fixate faces (Kanwisher, 2001). This hypothesis will require further testing, as well as more complete information about where object categories are typically seen on the retina in everyday, non-experimental settings.

The amount of location information contained in object-selective regions could serve as another organizing dimension in the VVC. Early theories, including the Dual Pathway Model (Underleider and Mishkin, 1982), state that object-selective regions in VVC form part of a processing stream largely devoid of information about location. Subsequent studies have documented that some degree of location information is present in shape-selective neurons of monkey IT (Op De Beeck and Vogels, 2000; DiCarlo and Maunsell, 2003). However, in Chapter 5 we find evidence in humans, both by measures of mean location biases and by multivariate pattern analyses, that considerable location information is present in the category-selective regions of VVC. Moreover, we find that the amount of location information in cortex may serve as an organizing dimension that distinguishes category-selective regions on the lateral occipital cortical surface from those on the ventral temporal surface.

Overall, these studies have identified several promising dimensions along which object-

selective regions may be organized. These findings relate to a question put forth in the Introduction: does the organization of VVC constitute a map in higher-level cortex? To address this question, we turn to a brief discussion of the nature of maps and compare the layout of VVC with well-studied maps in early sensory cortices.

Do category-selective regions in VVC comprise a map?

Although the term 'map' has a straightforward meaning with respect to primary sensory and motor areas, it is less clear how a map would manifest itself in cortex that processes complex information. According to the wiring optimization principle, maps develop to minimize the axonal and dendritic 'wiring costs' associated with long-range connections (Chklovskii and Koulakov, 2004). Therefore, maps can arise in any area, regardless of the nature of its content or the complexity of its computations, so long as variation in response properties along one or more dimension creates differential requirements for communication between cells. This necessity for local connections could be as simple as needing 'nose cells' and 'mouth cells' to interact with each other in order to recognize a face, or 'face cells' and 'body cells' to interact in order to recognize a person.

Although the principles of wiring optimization should apply to all regions of cortex, the nature of the representations in VVC is drastically different from those in early sensory cortices, and it is not clear how maps in one would resemble maps in the other. In typical early sensory maps, neural tuning changes smoothly across the cortex along a principal organizing dimension, such as retinal location for the retinotopic map in primary visual cortex (Durbin and Mitchison, 1990). The VVC appears to be composed of relatively discrete subregions with relatively sharp boundaries (Spiridon et al., 2006), suggesting that it may not form a 'smooth' map like the one in primary visual cortex. However, Op de Beeck and colleagues (Op De Beeck et al., 2008) point out that some early sensory

maps are comprised of discontiguous representations, such as the somatotopic map in the barrel cortex of rodents, which is made up of modular representations of each individual whisker. Thus, the spatial discontinuity of a functional region per se does not necessarily argue against the idea that this region is part of a broader map.

In addition to being 'lumpy', the organization of VVC in humans appears to be incomplete. Large portions of cortex are devoted to a limited set of categories (faces, bodies, and scenes), while localized selectivities for other objects (e.g., cars or trees) are rarely seen with fMRI (Spiridon and Kanwisher, 2002; Downing et al., 2006). However, this property of VVC is not incongruent with early sensory maps in the brain, which magnify inputs that are most relevant to the organism (e.g., foveal representations in primary visual cortex and finger representations in primary somatosensory cortex). In the same way, those objects whose recognition is most survival-relevant may receive expanded cortical representation. Thus, the 'lumpy' distribution of category selectivity in extrastriate cortex, made up of few highly selective subregions, does not preclude the existence of a coherent, inclusive map in VVC; other object categories may claim one or more localized representations in VVC as well, though the amount of cortex dedicated to these objects may be too small to be detected at the level of fMRI.

In sum, the principles of wiring optimization predict that maps can form anywhere in the brain, and comparison of the organization of VVC with early sensory maps shows that similarities exist between the two, despite the substantial differences in the types of information they contain. Moreover, the experiments presented in this thesis provide evidence for several candidate dimensions that may guide the layout of object-selective cortex. As yet, the use of the term "map" to describe high-level cortical areas is provisional, since the dimensions of that map have yet to be identified. However, studies like the ones

described in this thesis can advance our understanding of the large-scale organization of VVC and, in doing so, may help us see what such a map would look like.

References

Astafiev SV, Stanley CM, Shulman GL, Corbetta M (2004) Extrastriate body area in human occipital cortex responds to the performance of motor actions. Nat Neurosci 7:542-548.

Chklovskii DB, Koulakov AA (2004) Maps in the brain: what can we learn from them? Annu Rev Neurosci 27:369-392.

DiCarlo JJ, Maunsell JH (2003) Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. J Neurophysiol 89:3264-3278.

Downing PE, Chan AW, Peelen MV, Dodds CM, Kanwisher N (2006) Domain specificity in visual cortex. Cereb Cortex 16:1453-1461.

Durbin R, Mitchison G (1990) A dimension reduction framework for understanding cortical maps. Nature 343:644-647.

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

Hasson U, Levy I, Behrmann M, Hendler T, Malach R (2002) Eccentricity bias as an organizing principle for human high-order object areas. Neuron 34:479-490.

Kanwisher N (2001) Faces and places: of central (and peripheral) interest. Nat Neurosci 4:455-456.

Kourtzi Z, Kanwisher N (2000) Activation in human MT/MST by static images with implied motion. J Cogn Neurosci 12:48-55.

Levy I, Hasson U, Avidan G, Hendler T, Malach R (2001) Center-periphery organization of human object areas. Nat Neurosci 4:533-539.

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

Op De Beeck H, Vogels R (2000) Spatial sensitivity of macaque inferior temporal neurons. J Comp Neurol 426:505-518.

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

Peelen MV, Downing PE (2005) Is the extrastriate body area involved in motor actions? Nat Neurosci 8:125-126.

Schlack A, Albright TD (2007) Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. Neuron 53:881-890.

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

Spiridon M, Fischl B, Kanwisher N (2006) Location and spatial profile of category-specific regions in human extrastriate cortex. Hum Brain Mapp 27:77-89.

Underleider LG, Mishkin M (1982) Two Cortical Visual Systems. In: Analysis of Visual Behavior (Ingle MA, Goodale MI, Masfield RJW, eds), pp 549-586. Cambridge, MA: MIT Press.

Weisberg J, van Turennout M, Martin A (2007) A neural system for learning about object function. Cereb Cortex 17:513-521.