DISTRIBUTED ACTIVITY PATTERNS FOR OBJECTS AND THEIR FEATURES: DECODING PERCEPTUAL AND CONCEPTUAL OBJECT PROCESSING IN INFORMATION NETWORKS OF THE HUMAN BRAIN

Marc N. Coutanche

A DISSERTATION

in

Psychology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2013

Supervisor of Dissertation

_____

Sharon L. Thompson-Schill, Professor of Psychology

Graduate Group Chairperson

_____

John C. Trueswell, Professor of Psychology

Dissertation Committee

Russell A. Epstein, Associate Professor of Psychology

Joseph W. Kable, Assistant Professor of Psychology

DISTRIBUTED ACTIVITY PATTERNS FOR OBJECTS AND THEIR FEATURES: DECODING PERCEPTUAL AND CONCEPTUAL OBJECT PROCESSING IN INFORMATION NETWORKS OF THE HUMAN BRAIN

ACKNOWLEDGMENTS

I would like to express my profound thanks and appreciation to my advisor, Sharon Thompson-Schill. Sharon's intellectual insights, support, professional guidance, sense of humor and fantastic taste in restaurants are just some of the many qualities I have benefited from over the last five years. Sharon's perfect balance of providing intellectual guidance, with intellectual freedom, has helped fuel the excitement I feel about my research. Anyone who has ever raised an eyebrow when I casually mentioned that being a Ph.D. student for the rest of my life sounds like fun can blame Sharon.

I wish to thank my committee for their valuable ideas, guidance and continued professional support. I am very grateful to the Howard Hughes Medical Institute for kindly awarding me an International Student Research Fellowship, and to the National Institutes of Health, National Science Foundation (for grants awarded to ST-S) and the American Psychological Foundation for research funding. I also want to thank the many members of the Penn Psychology community that I have interacted with during my graduate studies. Supportive faculty members, psychology graduate students, lab members, and the undergraduates I co-advised all have my heartfelt thanks.

Finally, I wish to thank my family for their continual love and support, and my wife, Lauren, for her emotional support, intellectual insights and so much more.

The findings from Chapter 2 are currently under review. The findings from Chapter 3 have been published as: Coutanche, M. N., & Thompson-Schill, S. L. (2013). Informational Connectivity: Identifying synchronized discriminability of multi-voxel patterns across the brain. *Frontiers in Human Neuroscience*, *7*(15), 1–14.

ABSTRACT

DISTRIBUTED ACTIVITY PATTERNS FOR OBJECTS AND THEIR FEATURES: DECODING PERCEPTUAL AND CONCEPTUAL OBJECT PROCESSING IN INFORMATION NETWORKS OF THE HUMAN BRAIN

Marc N. Coutanche

Sharon L. Thompson-Schill

How are object features and knowledge-fragments represented and bound together in the human brain? Distributed patterns of activity within brain regions can encode distinctions between perceptual and cognitive phenomena with impressive specificity. The research reported here investigated how the information within regions' multi-voxel patterns is combined in object-concept networks. Chapter 2 investigated how memory-driven activity patterns for an object's specific shape, color, and identity become active at different stages of the visual hierarchy. Brain activity patterns were recorded with functional magnetic resonance imaging (fMRI) as participants searched for specific fruits or vegetables within visual noise. During time-points in which participants were searching for an object, but viewing pure noise, the targeted object's identity could be decoded in the left anterior temporal lobe (ATL). In contrast, top-down generated patterns for the object's specific shape and color were decoded in early visual regions. The emergence of object-identity information in the left ATL was predicted by concurrent shape and color information in their respective featural regions. These findings are consistent with theories proposing that feature-fragments in sensory cortices converge to higher-level identity representations in convergence zones. Chapter 3 investigated whether brain regions share fluctuations in multi-voxel information across time. A new analysis method was first developed, to measure dynamic changes in distributed pattern information. This method, termed "informational connectivity" (IC), was then applied to data collected as participants viewed different types of

man-made objects. IC identified connectivity between object-processing regions that was not apparent from existing functional connectivity measures, which track fluctuating univariate signals. Collectively, this work suggests that networks of regions support perceptual and conceptual object processing through the convergence and synchrony of distributed pattern information.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1: GENERAL INTRODUCTION

> *"When a collection of musical instruments is played in a coordinated manner, the aggregate function is the symphony that flows and has a recognizable structure and coherence … the aggregate is much richer than a particular element."*
>
> – McIntosh (2000, p. 863)

When you see or think about a lime, you can recognize or imagine its round shape and green color. You can anticipate how to hold it, and expect a zingy and sour taste from eating one. Each of these individual features is shared with many other objects, but their unique combination defines our concept of a lime. How does the human brain represent and integrate perceptual and knowledge fragments of objects? This dissertation investigates how distributed patterns of brain activity encode different components of object-related information, and how brain regions carrying these distributed codes interact during object perception and conception.

A widespread array of brain regions becomes active as we perceive an object or retrieve object knowledge from semantic memory (Binder & Desai, 2011). Some theories propose that our object knowledge is supported by the same neural systems that underlie perception and action; supported by findings that sensorimotor cortex becomes active when object knowledge is retrieved (Binder & Desai, 2011; Martin & Chao, 2001; Martin, 2007). Alternatively, one or more integration areas might bind together an object's properties, to form its identity elsewhere in cortex (Binder & Desai, 2011; Damasio, 1989; Lambon Ralph, Sage, Jones, & Mayberry, 2010). Such integration theories differ in the hypothesized location, and number, of integration zones (Binder & Desai, 2011; Simmons & Barsalou, 2003). Some theories propose that different object features are brought together in distinct convergence zones, which in turn feed into higher-level convergence zones (Binder & Desai, 2011; Damasio, 1989; Meyer & Damasio, 2009), while others propose that one central "hub" encodes object-concepts in an amodal format (Patterson, Nestor, & Rogers, 2007). A recent version of this theory, the "hub-and-spoke" model, retains a role for sensorimotor regions (the spokes) in supporting featural knowledge, while hypothesizing integration in a central hub (Lambon Ralph et al., 2010). While integration theories differ in their

details, they all propose the existence of at least one integration site for moving beyond single-feature processing.

The neural systems underlying object processing and semantic memory can be investigated in healthy humans using functional magnetic resonance imaging (fMRI). The blood-oxygenation level dependent (BOLD) signal recorded during an fMRI scan reflects the activity of large populations of neurons, providing a valuable window into a brain region's neural processing. Most commonly, the average BOLD response in a voxel or larger region is compared between conditions (commonly using the General Linear Model; Friston et al., 1994), where a greater average BOLD response is interpreted as reflecting increased neural processing. This overall response, however, is only one way in which information can be contained in a region's activity, and as will be discussed, an alternative form is more closely linked to the fine-grained distinctions in features and objects.

In addition to a region's overall response, information is now known to exist in the patterns of responses that are distributed across multiple voxels (Haxby et al., 2001; Norman, Polyn, Detre, & Haxby, 2006). Distinct percepts and cognitive states can be encoded within these unique combinations of voxel responses ("multi-voxel patterns"), allowing a large number of potential discriminable patterns within a voxel population, even when the region's mean response is similar (Coutanche, 2013). Multi-voxel patterns are particularly important to the neural underpinnings of object knowledge. Because distributed patterns are multi-dimensional (i.e., each voxel is one dimension), multi-voxel patterns can encode distinctions that are not as easily represented by differences in a mean response. For example, although a greater regional mean response to viewing shapes, compared to colors, could reflect shape-related processing, it is less straightforward to predict how distinctions *between* shapes, such as a cube, sphere and pyramid, would map onto a univariate mean response. This level of specificity is, however, central to object knowledge. It is not enough to know that limes have *a* shape and *a* color: our knowledge is based on their specific shape and particular color. Investigations have supported the proposal

that multi-voxel patterns encode finer distinctions between objects, while univariate responses represent broader categorical differences (Brants, Baeck, Wagemans, & Op de Beeck, 2011).

Multi-voxel patterns in ventral temporal (VT) cortex contain information about a visually presented object's category (e.g., Coutanche, Thompson-Schill, & Schultz, 2011; Coutanche & Thompson-Schill, 2012; Haxby et al., 2001; O'Toole, Jiang, Abdi, & Haxby, 2005; Spiridon & Kanwisher, 2002) and identity (e.g., Eger, Ashburner, Haynes, Dolan, & Rees, 2008). Imagery for items has also been decoded from this area of cortex, although the imagined items typically differ greatly in visual appearance (Lee, Kravitz, & Baker, 2012) and/or semantic category (e.g., people versus cars: Peelen & Kastner, 2011). Multi-voxel patterns for perceptual features have been decoded in early visual cortex, including for shape (Stokes, Thompson, Cusack, & Duncan, 2009), orientation (Kamitani & Tong, 2005) and motion-direction (Kamitani & Tong, 2006). A particularly powerful test of whether a hypothesized principle or dimension (e.g., shape) underlies a region's neural processing is to evaluate whether a model trained to distinguish items with a hypothesized distinction can successfully decode new items that vary in other dimensions (e.g., color). This type of strict "generalization test" (Tong & Pratte, 2012) is applied in Chapter 2 of this dissertation.

As discussed above, many regions of human cortex have been implicated in storing and retrieving object knowledge (Martin & Chao, 2001; Martin, 2007; Thompson-Schill, 2003). Multi-voxel patterns are almost exclusively investigated in isolated regions-of-interest (ROIs) to determine the information that is present within their activity patterns. While studies of multi-voxel patterns have examined more than one region (e.g., Kriegeskorte, Goebel, & Bandettini, 2006), each voxel population is typically analyzed separately to compare their relative discriminability (e.g., Walther, Caddigan, Fei-Fei, & Beck, 2009). However, brain regions operate within coordinated networks (Fox et al., 2005; McIntosh, 2000), and understanding a region's role within a network may be crucial to fully understanding its function and neural operations. For example, a recent review of investigations of the angular gyrus (AG) noted that "the exact role of the AG critically depends on the set of regions it is interacting with during a given task/process. This

implies that the role of the AG cannot comprehensibly be identified in isolation but ideally needs to be understood in parallel with the influence from other regions" (Seghier, 2013, p. 48). Considering the importance of inter-regional networks, and the increasingly apparent role of multi-voxel information in many perceptual and cognitive functions (Tong & Pratte, 2012), understanding how multi-voxel information is operationalized and integrated at the network-level has great potential for advancing our understanding of how the brain integrates information across systems.

The work in this dissertation addresses how the multi-voxel patterns that underlie features and objects are brought together at the network level. To address this issue, the work here raises the novel question of how feature and object patterns emerge synchronously across regions that operate in networks. Whereas investigations employing MVPA have predominantly ignored region-to-region relationships, the studies described here examine synchronous decoding through new approaches, which allow questions to be raised that could not otherwise be addressed, including: 1) Does the decoding of higher-level objects depend on the synchronous decoding of lower-level features? 2) Is the level of an object's multi-voxel information correlated between regions across time? Such questions cannot be answered with typical MVPA approaches, but the approaches in this work allow these ideas, and others, to be tested.

Chapter 2 reports an investigation of distributed activity patterns underlying our knowledge of objects, including their features and identity. The question of how knowledge of an object's features and identity is instantiated in the human brain remains a key question of modern cognitive neuroscience (Meyer & Damasio, 2009). This study examines object knowledge at multiple levels of the visual hierarchy, and investigates links between these levels using a novel analysis of dependencies between multi-voxel codes. Chapter 3 further investigates how object processing engages multi-voxel information across the cortex, by identifying networks with similar across-time profiles of multi-voxel information. Functional networks are frequently investigated using univariate activation, where synchronous fluctuations in univariate responses are identified across voxels or regions (e.g., functional connectivity; Biswal, Zerrin Yetkin, Haughton, & Hyde,

1995). In contrast, there has been little characterization of networks of synchronized *multi-voxel* information. Chapter 3 first describes the development of a novel analysis method that allows this type of investigation. By using a metric of multi-voxel information that is deployed across a timeseries, this method – "informational connectivity" (IC) – can identify relationships between regions based on the ebb and flow of multi-voxel information. The chapter goes on to apply this method to brain activity recorded as individuals viewed different types of man-made objects. The method identified object-processing networks on the basis of their synchronized multi-voxel information. Regions are identified that have been hypothesized to play key roles in object processing, and the results contribute to current theoretical debates.

Multi-voxel patterns encode information at a level of specificity that is centrally relevant to the convergence of features into objects, and to distinctions between objects, during perceptual and conceptual processing. Crucial to understanding these cognitive systems is discovering how brain regions operate within connected information networks. Together, the chapters here seek to address how the neural signatures of these fine-grained perceptual and conceptual items are represented within networks of information-rich brain areas.

CHAPTER 2: CREATING CONCEPTS FROM CONVERGING FEATURES IN HUMAN CORTEX

Abstract

To make sense of the world around us, our brain must remember the overlapping features of millions of objects. Crucially, it must also represent the unique feature-convergence that defines every object. We know very little about how the brain binds feature knowledge fragments into identity. Here we describe a functional magnetic resonance imaging study of neural information for features and identity in humans searching for fruits and vegetables within random visual noise. A pattern-classification algorithm could decode a (unseen but anticipated) target's identity within the anterior temporal lobe, and its specific shape and color within early visual regions. A novel analysis revealed that converging shape and color codes predict emerging identity information. People with stronger dependencies between featural-convergence and identity had memory-generated activity that more closely resembled visually generated codes. These results support theories proposing that convergence zones bind feature knowledge fragments together to form an object's identity.

Introduction

We encounter millions of objects during our lifetime that we recognize effortlessly. We know that a lime is green, round and tart, whereas a carrot is orange, elongated and sweet, helping us to never confuse the wedge on our margarita glass with our rabbit's favorite treat. One property (or 'feature') alone is typically insufficient for identification: celery can also be green and tangerines are also orange. Instead, we draw-on the unique convergence of features that defines an object. How does our brain bind together the many possible sensorimotor features to form a unique memory representation?

Several theories have been proposed to address this challenge. One theory proposes that knowledge of objects resides in the very sensorimotor cortices that process their features during perception or use (Kiefer & Pulvermüller, 2012; Martin, 2007). An alternative set of theories suggest that objects become represented in one or more central cortical hubs or integration areas (Lambon Ralph, Sage, Jones, & Mayberry, 2010; Patterson, Nestor, & Rogers, 2007; Pobric, Jefferies, & Lambon Ralph, 2010; Simmons & Barsalou, 2003). The anterior temporal lobe (ATL) – an area well connected to temporal, parietal and frontal cortices – has been proposed as a candidate hub, supported by evidence of conceptual impairments that can accompany ATL atrophy during semantic dementia (Hodges, Patterson, Oxbury, & Funnell, 1992; Rogers, Patterson, & Graham, 2007). Hub-based theories have proposed that a hub enables concepts to be re-representation in a high-dimensional semantic space, enabling concepts that have very different features (such as a lime and carrot) to be semantically close and *vice versa* (Lambon Ralph & Patterson, 2008).

Damasio's first proposal for the existence of an integration zone suggested that convergence zones hold a binding code, or combinatorial record, for "feature fragments" (coded in sensory cortex) that form our knowledge of objects when successfully combined (Damasio, 1989; Meyer & Damasio, 2009; Simmons & Barsalou, 2003). Little direct evidence exists for the convergence zone hypothesis (Simmons & Barsalou, 2003), not least because a crucial information link – between identity in a potential convergence zone and the object's specific

feature fragments – has yet to be found in the human brain. A putative convergence zone should show certain characteristics that are testable using novel fMRI analysis techniques. Specifically, the convergence account leads to these three key predictions: i) Thinking about an object should evoke a pattern of brain activity coding its identity in a purported convergence zone (the result of convergence). ii) Early visual regions that specialize in processing the features of a retrieved object should activate specific feature fragments (the substrates for convergence). A given feature's code will be shared among objects that share this feature. Notably, we currently do not know the degree to which a convergence zone would reactivate features (Simmons & Barsalou, 2003), which could range from general shape processing for all objects, to the specific neural activity specifying a sphere rather than a cube. The strictest form of "feature fragment" would predict the latter. iii) Successful convergence (marked by successful identity decoding) should be linked to the simultaneous presence of the convergence substrates. One of the first theoretical proposals of a convergence zone set the specific requirement that activation of "convergence zones would produce synchronous activity in separate cortical sites presumed to contain feature-fragments related to the convergence zone" (Damasio, 1989, p. 56). In contrast, a rival theory, in which object knowledge is contained solely within sensorimotor regions, does not predict this link. We developed a novel analysis to test the relationships present between different types of distributed information, allowing us, for the first time, to detect a relationship between specific feature fragments and the result of their convergence to object identity, in the human brain.

In the present study we examined top-down generated activity patterns for fruits and vegetables that varied orthogonally by color, shape and identity. Top-down processes can influence neural activity in cortical regions that respond to visual stimulation (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990; Maunsell & Treue, 2006). We employ a task that engages top-down influences with no visual information on-screen, allowing us to investigate neural signatures for retrieved object knowledge rather than visual inputs.

Method

Subjects

Data from eleven participants (3 females, 18-35 years old) are analyzed (a twelfth participant's fMRI data were not analyzed due to abnormal behavioral responses during the task). All participants were right-handed with normal or corrected-to-normal vision and reported no history of neurological problems. Participants provided written informed consent and received monetary compensation for their participation. The human subjects review board at the University of Pennsylvania approved all experimental procedures.

Magnetic resonance imaging acquisition

Subjects were scanned with a 3-T Siemens Trio system equipped with an eight-channel head coil and foam padding for stabilizing the head. T1-weighted anatomical images were acquired at the beginning of each session (TR = 1,620 ms, TE = 3 ms, TI = 950 ms, voxel size = 0.977 mm × 0.977 mm × 1.000 mm). T2*-weighted scans recorded blood oxygenation level–dependent (BOLD) response using interleaved gradient-echo EPI (TR = 3,000 ms, TE = 30 ms, field of view = 19.2 cm × 19.2 cm, voxel size = 3.0 mm × 3.0 mm × 3.0 mm, 42 slices).

Experimental procedure

Prior to fMRI scanning, participants completed a behavioral staircasing procedure to determine the level of visual noise that was later applied to images during the fMRI scan. This ensured the in-scan detection task would be challenging enough to engage each subject. On each trial of this staircasing behavioral task, subjects indicated with a button-press when they could identify a fruit (bananas and tomatoes, two fruits not used for the primary task) displayed in a field of Gaussian visual noise. After each behavioral response, the variance of the noise added to the next image was increased or decreased, bringing the subject's final detection level to 75% accuracy by the end of the procedure. The particular noise variance producing this detection-level for each subject was then applied to that participant's images during their scan.

Participants' detection accuracies during the scanning session (M = 73%, s.d. = 11%) were very similar to the 75% staircasing target, suggesting the procedure was successful.

At the beginning of the scanning session, participants passively viewed images of exemplars of the four types of fruit and vegetables (carrots, celery, limes and tangerines) that would later act as targets, centrally placed on a white background. Six images of each type of fruit and vegetable were presented in a random order in each 18 sec block, with each image shown for 3 secs. Blocks were separated by 12 secs of fixation. In total, twelve exemplars of each type of fruit and vegetable (later hidden in the detection task) were presented, split across two blocks.

During the next four scanning runs, participants were instructed to respond with a button-press when they detected a cued fruit or vegetable within visual noise. Word cues were presented to indicate a type of fruit or vegetable that should be detected (e.g., 'carrot'). A variable number of images then followed, each with Gaussian white noise based on a mean of zero and variance-level determined by the subject's prior staircasing (see Figure 1). In every run, each of the four fruit and vegetable cues was cumulatively followed by the same total amount of visual noise. Each fruit and vegetable cue (e.g., 'carrot') occurred three times in a run (giving 12 blocks for each fruit and vegetable across the experiment). Blocks were presented in a pseudo-randomized order so that the same cue did not immediately repeat. Within the blocks, pure-noise images were each shown for 3 secs. Following a variable length of time (between 12 and 24 secs after the initial cue), a fruit or vegetable was presented, hidden within Gaussian noise. The block automatically ended after this trial. This design afforded us the high signal sensitivity found with block designs, combined with unpredictability to keep participants cognitively engaged. Two of the fruit-in-noise images ending the noise blocks contained a fruit or vegetable that did not match the preceding cue (i.e., 2 out of 12 were foils), to focus participants on detecting the specific target. The two foils for each kind of cue were other fruits / vegetables with the same color but different shape, or same shape but different color (e.g., for carrot: tangerine and celery), ensuring that the four objects acted as foils with the same frequency. To encourage participants

to search for the cued target from the very start of every block, the beginning of each run (during the four beginning time-points routinely removed in preprocessing) included a short block in which a cued fruit or vegetable appeared after only 3 or 6 secs, followed by 12 s of fixation. The hidden fruits and vegetables that ended each short block were not repeated in the main blocks and their BOLD signal did not contribute to any analyses.

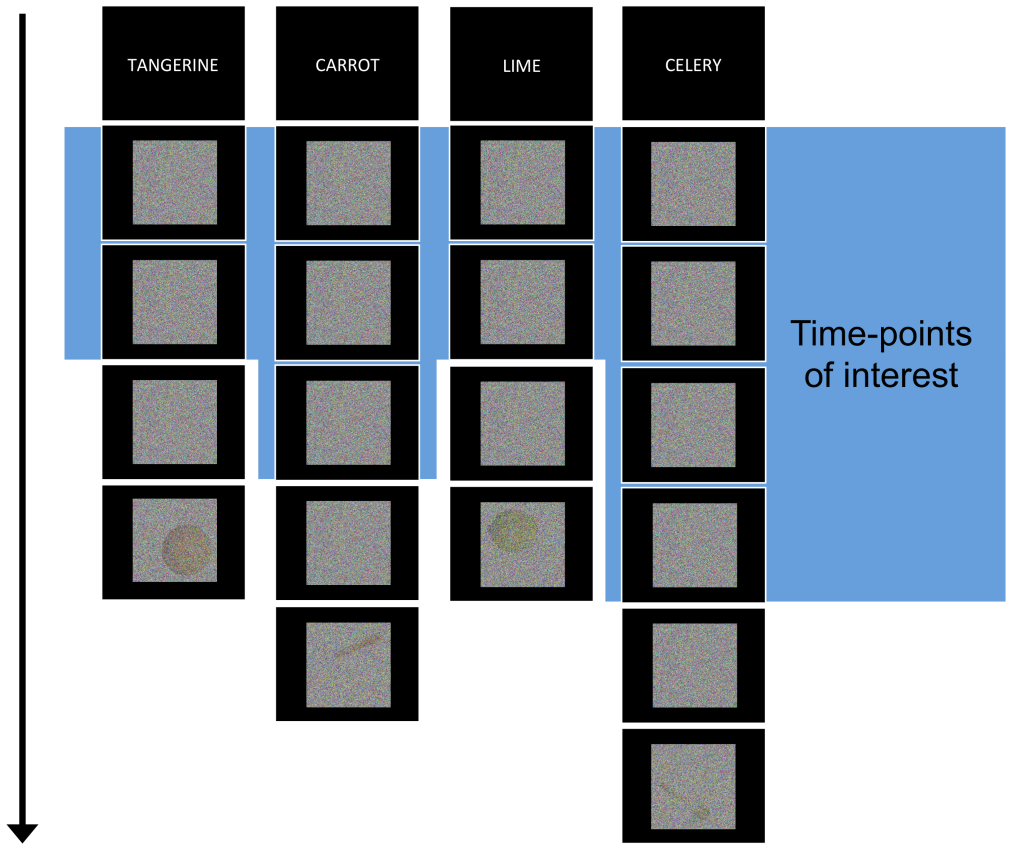


Figure 1: Experimental design. Participants were presented with cues of items to detect, followed by blocks of visual noise. Each block ended with an actual image embedded in noise at a threshold determined for each participant prior to their scan (shown here at a low threshold for visualization purposes). Blocks contained an unpredictable amount of pure noise and occasionally ended with an incorrect (non-cued) fruit or vegetable to keep participants on-task. The objects in the final trial are displayed here in each corner although they could appear in any

corner in the actual experiment. Every block ended with a unique instance of that kind of fruit or vegetable (e.g., no particular tangerine appeared more than once). Data from the last noise time-point was discarded to ensure the signal-ascent from viewing the image-in-noise did not influence the analyzed data.

The 12 images of each hidden fruit / vegetable (10 cued, 2 foils) ending the blocks were photographic examples in various orientations on a white background. The objects were all adjusted to have the same height. The hidden items appeared in four possible locations: top-left, top-right, bottom-left and bottom-right (see Figure 1 for an example). The objects appeared in each location 3 times across the experiment, in a randomized order. The objects appearing in each location were preceded by the same cumulative amount of Gaussian noise across the experiment (i.e., there was no contingency between the amount of noise and final stimulus location).

Magnetic resonance imaging preprocessing

Imaging data were preprocessed using the Analysis of Functional NeuroImages (AFNI) software package (Cox, 1996). The first four volumes of each functional run were removed to allow the signal to reach steady-state magnetization. All functional images were slice time corrected and a motion correction algorithm registered all volumes to a mean functional volume. Low frequency trends were removed from all runs using a high-pass filter threshold of 0.01 Hz. Voxel activation was scaled to have a mean of 100, with a maximum limit of 200. The data were not smoothed.

Decoding analysis

Pattern decoding was conducted within the MATLAB environment. The functional data were first z-scored within each run. Data from repetition times (TRs) corresponding to pure visual

noise trials (i.e., the TRs of data collected before the participant encountered a concealed fruit or vegetable) were isolated in the following manner: The pure-noise TRs were assigned binarized labels of the search target. They were then convolved with a time-shifted model of the hemodynamic response and thresholded at 0.8, to identify the events predominantly affecting each time-point. This gave a vector of activity values n-voxels long for each pure-noise TR, which were averaged by block. To ensure that the block average was not influenced by the signal-ascent of the block's final trial (in which a fruit or vegetable was actually present), we removed the last pure-noise TR of each block before averaging.

We conducted an information brain mapping 'roaming searchlight' analysis in each participant by centering a sphere (3-voxel radius) on each voxel in turn (Kriegeskorte, Goebel, & Bandettini, 2006). MVPA was conducted with the voxels in each searchlight volume (123 when not restricted by the brain's boundary) and performance was allocated to the central voxel. For each searchlight, 4-fold cross-validation was conducted (training on three runs; testing on the fourth) with a Gaussian Naïve Bayes (GNB) classifier (implemented through the MATLAB Statistics toolbox) to classify activity to noise trials according to the search target (carrot, celery, lime or tangerine). The classifier was trained and tested on the vectors of BOLD activity values that were averaged for each block in the manner described above. GNB classifiers have been shown to have particular success for datasets with small numbers of training samples (Mitchell et al., 2004; Ng & Jordan, 2002; Singh, Miyapuram, & Bapi, 2005), such as here where each block contributes one datapoint. It is also fast for searchlight analyses (Pereira, Mitchell, & Botvinick, 2009).

Each participant's map of searchlight accuracies was brought to standardized space (with the same resolution as the functional data) and spatially smoothed with a 6 mm FWHM kernel. The 11 searchlight maps were submitted to a group analysis to test whether the accuracy at each voxel was greater than 0.25 (chance), with familywise error correction for multiple comparisons (corrected to $p < 0.05$, with a 26-voxel cluster threshold estimated with AlphaSim; Cox, 1996).

The next analysis tested whether a model trained on the cued visual noise would generalize to activity patterns (also block averages) from the separate passive-viewing run. A classifier was trained on all the pure-noise trials labeled by cue, and tested on data from the passive-viewing run labeled by the fruit or vegetable on-screen. This 4-way classification was performed with the voxels of each searchlight volume identified in the prior analysis (transformed back into each participant's original space), with the searchlights' performances averaged. We could not train on the passive-viewing data due to an insufficient amount of training data. To assess statistical significance, we conducted permutation testing. First, each participant's classifier testing labels were scrambled 1,000 times and the classification was repeated for each new set of labels. This produced 1,000 permutation-generated classification accuracies for each participant. To obtain a group p-value, a null distribution was created by randomly sampling a classification accuracy value from every subject's 1,001 classification scores (1,000 permutations + 1 real order) and calculating the group mean. This was performed 10,000 times, giving 10,000 permuted group means. The real group mean was compared to this null distribution to identify the p-value.

To conduct color and shape generalization tests, we trained classifiers to distinguish two items differing in one dimension (e.g., carrot versus celery for color) and tested the model on the unused items that varied in the same way (e.g., tangerine versus lime). This was performed on data from lateral occipital cortex, V4 and the left ATL region identified from the searchlight analysis. A 4-fold leave-one-run-out cross-validation procedure was conducted twice: alternating which items were used for training. Each pair of scores was averaged. To assess statistical significance, we conducted the permutation testing procedure described above, with each set of randomized labels held constant for the two combinations of training and testing. The null distribution was generated by sampling 1,000 group means by randomly selecting from each participant's 100 permutations of classification scores. The p-value was then calculated from this distribution.

The shape and color convergence analyses were conducted by extracting classification accuracy vectors (i.e., 1 versus 0 for each block) for color classification in the color region, shape

classification in the shape region, and identity classification in the identified temporal lobe searchlights. A logistic model [with quadratic penalty determined by marginal likelihood maximization for convergence (Zhao & Lyengar, 2010) and coefficient stability] predicted object identity decoding success (48 values; one for each block) for each of the identified temporal lobe searchlights, with predicting variables for: block-by-block success in color decoding, block-by-block success in shape decoding and block-by-block color-shape conjunction (color decoding x shape decoding). Odds ratios were calculated for the models' coefficients (eB) and averaged across the identified searchlights for each subject.

Regions of interest

The color and shape across-item generalization tests were conducted using voxels in regions involved in shape and color processing. The shape-relevant region was based in lateral occipital cortex, an area with location-tolerant shape information (Carlson, Hogendoorn, Hubert Fonteijn, & Verstraten, 2011; Eger, Ashburner, Haynes, Dolan, & Rees, 2008). Previous research has shown that this region is modulated by top-down processing (Reddy, Tsuchiya, & Serre, 2010; Stokes, Thompson, Nobre, & Duncan, 2009). We extracted standard space coordinates from a highly cited study of shape processing (Grill-Spector et al., 1999). The lateral occipital shape region can be characterized by three vertices (dorsal, posterior and anterior), so we placed three spheres (6 mm radius -the reported extent of activation) against the vertex coordinates from the object > texture contrast, in both hemispheres (coordinates in Table 1). This successfully encompassed lateral and ventral regions of the LOC.

The color-processing region was based on a seminal color-processing study (McKeefry & Zeki, 1997). The coordinates for maximum activation in a chromatic versus achromatic contrast were extracted from this study and a sphere (radius 6 mm -the listed standard deviation of the extent of activation) was placed at the right hemisphere coordinates. Investigations have suggested that right V4 is particularly modulated by top-down control of color processing (Bramão, Faísca, Forkstam, Reis, & Petersson, 2010; Kosslyn, Thompson, Costantini-Ferrando,

Alpert, & Spiegel, 2000; Morita et al., 2004) and achromatopsia is differentially associated with right V4 damage (Bouvier & Engel, 2006), so we focused on the right region (coordinates in Table 1), although also examined left V4 from the same study.

Shape	-41, -77, 3	-36, -71, -13	-38, -50, -17
	40, -72, 2	37, -69, -10	33, -47, -14
Color	30, -78, -18		
	-26, -80, -14		

Table 1. Coordinates for feature ROIs. Talairach coordinates for shape (extracted from Grill-Spector et al., 1999) and color (extracted from McKeefry & Zeki, 1997) regions. The shape coordinates reference the three vertices in each hemisphere that characterize the lateral occipital shape region. Spheres (6 mm radius) were positioned to border each vertex. The color coordinates reflect the center of right and left placed spheres (6 mm radius).

Results

We presented participants with images of colored random noise and directed subjects to detect one of four types of fruits and vegetables – carrot, celery, lime and tangerine – that vary systematically by shape and color (Figure 1). These blocks of pure-noise ended after an unpredicable amount of time with the cued fruit or vegetable, or a foil, embedded within noise. We analyzed data from time-points before any fruit or vegetable was revealed, in order to examine top-down driven activity. Prior to the main task, participants passively viewed exemplars of the four types of fruit and vegetable, giving us examples of visually-generated activity patterns for these items.

Decoding object identity from anticipatory visual activity

To test the first requirement of a convergence zone – that a brain region contains a memory-evoked code for object identity – we asked if the identity (carrot, celery, lime or tangerine) of the searched-for object could be decoded as participants viewed visual noise. The location (or even existence) of a convergence zone has not been established, so we used a searchlight analysis to analyze sequential clusters of voxels. The functional data recorded during visual noise time-points were labeled by the participant's current detection target (given to participants through a preceding cue) and then submitted to a 4-way machine learning classifier. The classifier was able to decode (at $p < 0.05$ corrected) the identity of the anticipated-but-unseen targets in a cluster of 64 searchlights in the left ATL (in which the volume includes the left fusiform gyrus, interior temporal, middle temporal and superior temporal cortex, verified by cortical segmentation and automated labeling through FreeSurfer; Fischl et al., 2002). The region was centered at -41x, -8y, 17z and is shown in Figure 2.

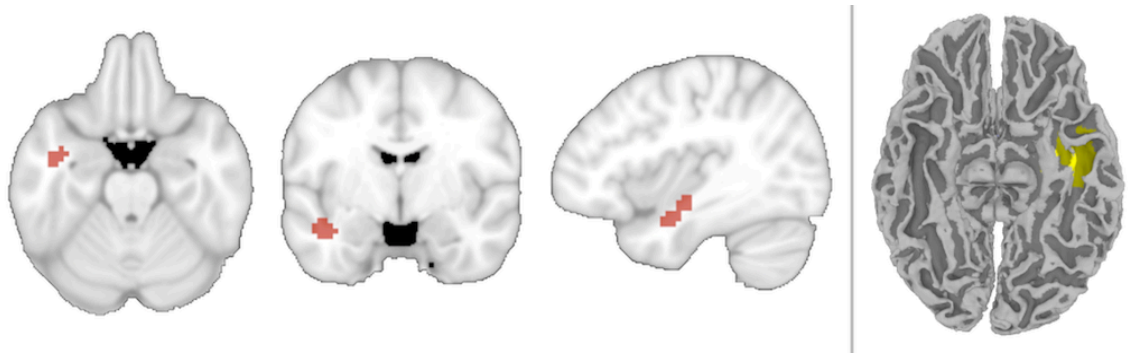


Figure 2: Location for searchlights with above-chance decoding of object identity while participants viewed visual noise and attempted to detect one of four kinds of fruit and vegetable. Left: A 4-way searchlight analysis revealed a region within the left ATL capable of decoding the target. Searchlight centers are shown in red. Right: The searchlights' volume displayed in one participant's original space, shown on their T1 anatomical image after automated cortical reconstruction and volumetric segmentation using the FreeSurfer image analysis package (Fischl et al., 2002).

We verified that this significant decoding was not based purely on a sub-categorical distinction between 'fruits' and 'vegetables' by successfully classifying items that do not cross this fruit / vegetable boundary (i.e., carrot vs. celery and lime vs. tangerine) at a level significantly above chance (permutation testing: $p = 0.025$). We also confirmed that time-points from each of the four fruits and vegetables had above-chance accuracies ($p < 0.05$). Although unlikely that motor responses could account for temporal lobe performance, we also confirmed that participants' numbers of motor responses did not differ significantly between the different targets ($F(3,30) = 1.62$, $p = 0.23$).

We performed further analyses to test the specificity of the left lateralization of the identified region by analyzing an ROI in the right hemisphere at the same y and z coordinates as the left ATL region. Successful decoding was specific to the left ATL: the right ATL's performance was not significant ($M = 0.26$ where chance $= 0.25$; $p = 0.30$), with the left ATL having significantly greater performance in a two-tailed paired t-test: $t(10) = 3.64$, $p = 0.005$.

Because of the known signal issues in the ATL, we measured the temporal signal-to-noise ratio (tSNR) of the left and right ATL regions, to assess signal quality, and to ask if tSNR differences could account for the lateralization. The tSNR is calculated by dividing each voxel's mean signal with its standard deviation over the time-course of each run. The tSNR values of the searchlight centers were high for both ATL regions (mean left = 77.4; mean right = 77.5) and well above levels that are considered suitable for signal detection (e.g., 20 in Binder et al., 2011). This indicated that the signal was strong in both regions, which additionally did not differ (t_{10} = 0.01, p = 0.99).

Are multi-voxel patterns necessary to distinguish object identity? A direct and comparable way to examine if univariate differences can distinguish the objects is to re-run the classification, but replacing the multi-voxel patterns with the univariate mean of each block (Coutanche, 2013). If conditions are separable by any univariate differences, this approach will produce above-chance classification performance. We ran this analysis within the left ATL and found that univariate activation cannot separate the conditions (mean classification performance = 0.24; chance = 0.25; p = 0.70 from permutation testing). The importance of multi-voxel coding to the investigated contrast is expected, given the necessity of multi-voxel patterns in successfully decoding information about different objects (Eger et al., 2008; Haxby et al., 2001).

We next examined the nature of the top-down generated identity code by asking whether it would generalize to activity that had been recorded while subjects viewed examples of the fruits and vegetables. We trained a classifier on the noise-only trials in the searchlights identified above (transformed back to each participant's original space), with each trial labeled according to the search target. We tested the trained models on data from a separate run in which participants had viewed blocked images of each kind of fruit or vegetable. The model trained on preparatory activity in the ATL was able to successfully classify the type of fruits and vegetables viewed in the separate passive-viewing run (M = 0.30, s.d. = 0.08; chance = 0.25; permutation testing: p = 0.037), revealing that the memory-generated and visually generated patterns were similarly structured (Figure 3).

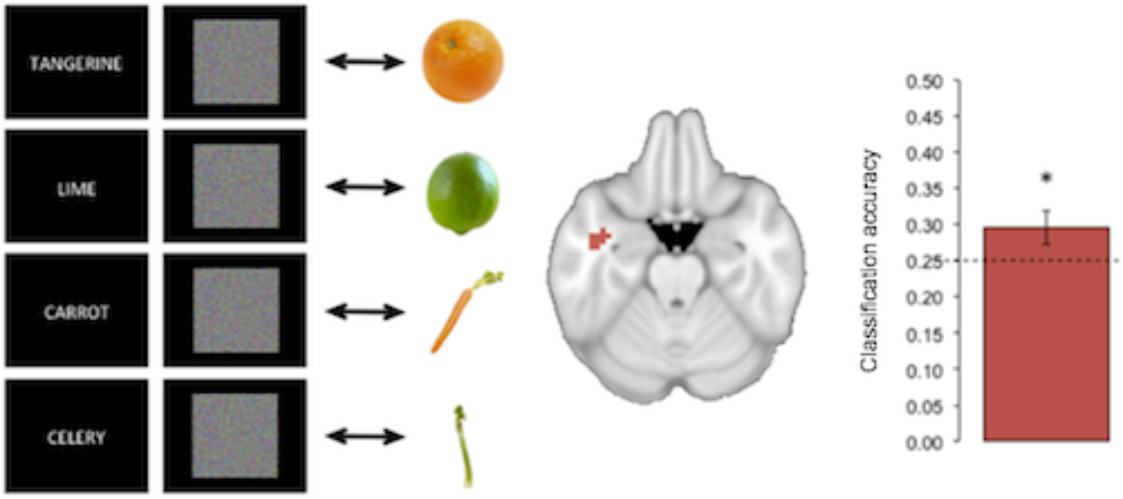


Figure 3: Generalizing from top-down activity to visual perception. Left: A classifier was trained on activity patterns recorded as participants viewed visual noise and sought to detect a cued fruit or vegetable. The classifier model was then tested on activity recorded as participants viewed real images of category examples in a separate run. Center: Activity patterns in this analysis were extracted from the left temporal lobe searchlights identified in the prior analysis of noise trials alone. Right: Classification accuracy significantly exceeded chance performance, reflecting successful generalization from anticipatory activity to visual perception. The dashed line reflects the level of chance and the error bar shows the standard error of the mean. The asterisk signifies above-chance classification performance ($p < 0.05$).

Decoding object features

The four targets in this study differed orthogonally by shape (two elongated, two spherical) and color (two orange, two green), allowing us to decode each feature independently, and test the second prediction of the convergence zone theory: that specific feature knowledge fragments become active in sensory regions (Figure 4). We examined this by asking whether a model trained to distinguish different shapes or colors could generalize to another pair of objects with this same distinction, but variation in other dimensions. We investigated shape and color

generalization in: i) a bilateral region of lateral occipital cortex that is associated with shape processing, and ii) an occipital area (right V4) associated with color processing (see Method for full details).

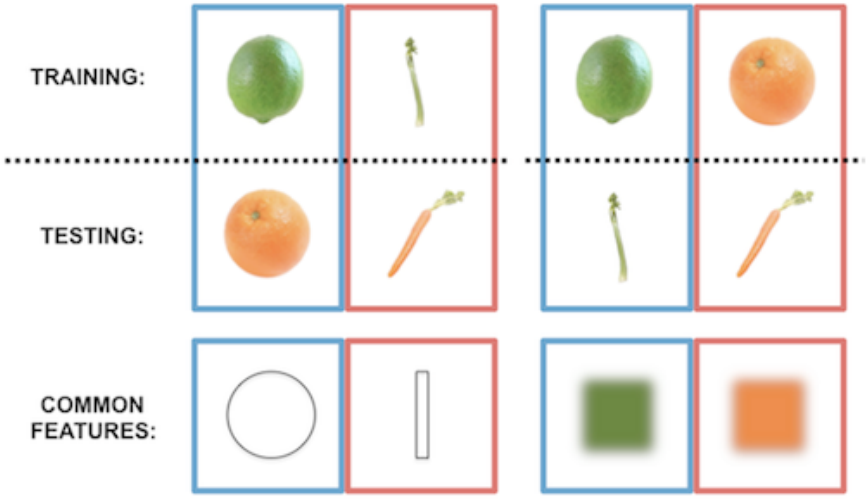


Figure 4: Feature-based generalization. Classifiers were trained to distinguish noise trials in which participants were searching for fruits and vegetables differing by shape or color. The classifiers were then tested on noise trials with the other pair of targets that differed in the same way. In the first example (left), classifiers are trained and tested based on shape (trained on lime versus celery, tested on tangerine versus carrot). In the second example (right), classifiers are trained and tested based on color (trained on lime versus tangerine, tested on celery versus carrot). The items took turns to act as the training data and the results of both comparisons were then averaged.

A classifier model that was trained on data from when participants were searching for two fruits and vegetables that differed by shape (e.g., lime vs. celery), could decode the remaining fruits and vegetables with similar shapes (tangerine vs. carrot), using activity from the bilateral lateral occipital cortex ($M = 0.55$, s.d. $= 0.06$; $p = 0.01$). Successfully decoding specific shapes across different colors (training on green, testing on orange) provides strong evidence that feature fragments were specific to the features of each activated object. The same region could

not decode the color of the targets (M = 0.51, s.d. = 0.07; p = 0.43), with higher performance for shape (p = 0.05). The right V4 region contained activity patterns that, using the same approach as above, decoded colors (M = 0.56, s.d. = 0.08; p = 0.01), and also shape at a trend-level (M = 0.54, s.d. = 0.09; p = 0.09, no significant difference: p = 0.17). This was specific to the right V4 region: left V4 could decode neither (p > 0.4; lower accuracies than the right region for color: p = 0.04, although not shape: p = 0.15). There was a significant interaction for greater shape decoding in lateral occipital cortex and greater color decoding in right V4; p = 0.03; Figure 5). As expected for regions coding feature fragments, identity decoding was unsuccessful in both regions (p > 0.46). We also returned to the identity-decoding ATL region and applied the feature generalization tests to activity from this region. Consistent with the ATL region containing identity information that is transformed away from features, neither shape (M = 0.49, s.d. = 0.04; p = 0.61) nor color (M = 0.50, s.d. = 0.05; p = 0.46) were decodable (significantly lower decoding than in V4 for color: p = 0.01, and lateral occipital cortex for shape at a trend: p = 0.09).

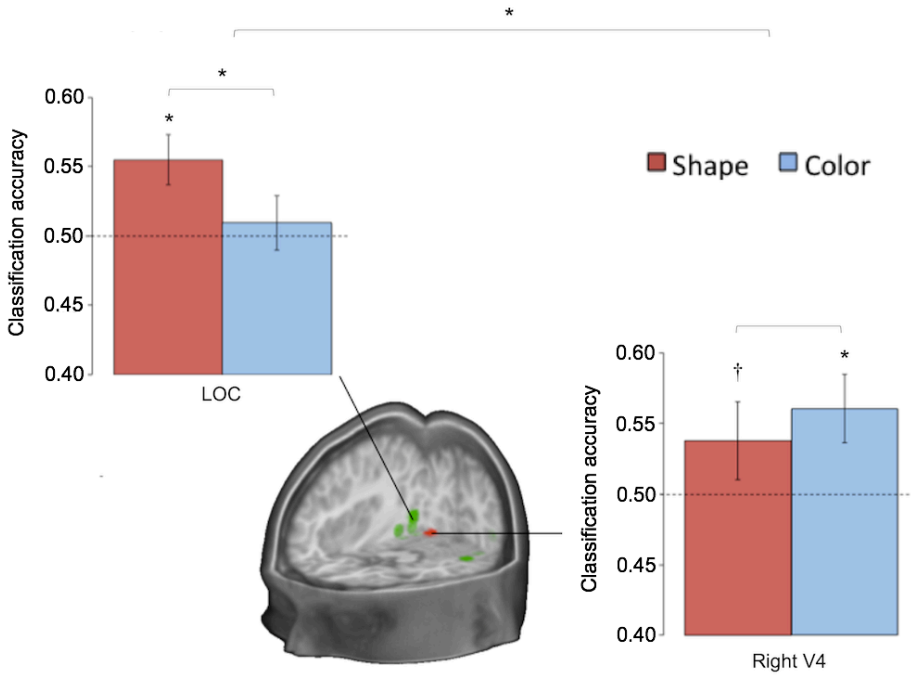


Figure 5: Classification results from the shape and color decoding analyses. Results are displayed from training a classifier on data from noise trials when participants were attempting to

detect targets that differed by shape or color, and tested on data with other targets that varied in the same way. The shape results (e.g., training: lime versus celery, testing: tangerine versus carrot) are shown in red. The color results (e.g., training: tangerine versus lime, testing: carrot versus celery) are shown in blue. The dashed lines reflect the level of chance and the error bars show the standard error of the mean. Asterisks signify above-chance classification performance ($p < 0.05$). The cross signifies trend-level performance ($p < 0.1$). The green region displayed in the cross-section is the area of the lateral occipital complex. The red region is based on the color-responsive area - right V4 (Methods).

Shape and color conjunction predicts the left ATL's identity code

The third and final convergence zone prediction was that the converged (identity) code should occur with converging activation of the specific shape and color feature fragments for the object. We employed a novel decoding-dependency analysis to examine this. We first coded each classified block of every participant for whether its neural activity contained decodable object identity in the ATL, color in V4 and shape in lateral occipital cortex. We then created a logistic regression model (full details in Method) of identity-decoding success (1 versus 0) in each block for the discovered ATL region, with predictors for: i) color decoding success in right V4, ii) shape decoding success in lateral occipital cortex, and iii) simultaneous shape and color decoding (i.e., i x ii). The odds ratios of this model reflect dependencies between the feature fragments and converged-upon identity.

The conjunction (i.e., convergence) of V4 color decoding and lateral occipital shape decoding was specifically predictive of successful ATL identity decoding in the model (M odds ratio = 2.64; odds ratio > 1: $t(10) = 4.08$, $p = 0.002$) unlike each feature alone (M odds ratio for color: 0.76; M odds ratio for shape: 0.66). This relationship is also apparent from directly comparing the degree of converging shape and color decoding in blocks with, and without, correct identity classification in the ATL: blocks with successful identity decoding had higher proportions of conjunctive shape and color decoding (M = 0.35, s.d. = 0.10; paired two-tailed t-

test: t (10) = 11.63, p < 0.001) than blocks that were misclassified (M = 0.15, s.d. = 0.05). These types of blocks did not differ in proportions of successful color (paired two-tailed t-test: t (10) = -0.14, p = 0.89) or shape (paired two-tailed t-test: t (10) = 1.37, p = 0.20) decoding alone. How do these rates of conjunctive color-and-shape decoding compare to a model of independence (i.e., where $P_{color\text{-}and\text{-}shape} = P_{color} \times P_{shape}$)? The degree of shape-color conjunction was greater in ATL identity-decoded blocks than would be predicted under a model of independence (trend level: two-tailed paired t-test: t (10) = 1.93, p = 0.08), and was substantially lower in identity-misclassified blocks (two-tailed paired t-test: t (10) = -7.74, p < 0.001), showing that the concurrent color and shape decoding co-occurs with successful object-identity decoding to a greater degree than expected from their baseline occurrences. The feature decoding results for blocks with successful and unsuccessful identity decoding are shown in Figure 6.

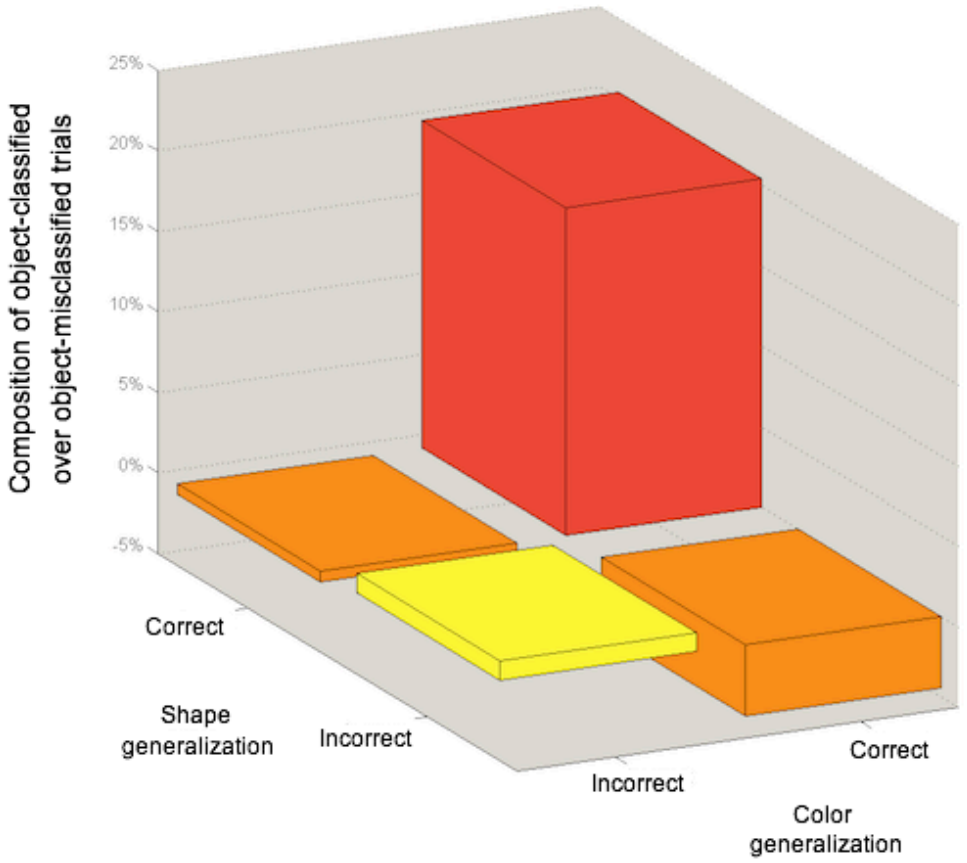


Figure 6: The properties of blocks with successful ATL object-identity decoding, compared to those blocks with unsuccessful identity decoding. Positive values indicate that greater shape and/or color decoding was found in blocks having correct identity decoding.

Finally, if the relationship between color–shape conjunction and ATL identity decoding plays an active role in evoking a concept, we might expect subjects with stronger relationships to have conceptually driven ATL-codes that more closely match codes evoked from viewing exemplars of a concept. This correspondence between conceptual and perceptual codes reflects the extent that top-down processes activate generalizable patterns. A relationship between the strength of the feature-to-identity relationship, and percept / concept generalizability, would suggest that this convergence is directly tied to the character of the activated concept. We confirmed this relationship. Subjects with a stronger link between feature fragments and their ATL (indicated by a higher odds ratio for shape-and-color convergence predicting ATL identity in the previous logistic regression) had top-down ATL identity codes that more closely resembled visually driven activity patterns ($r = 0.67$, $p = 0.02$; Figure 7). This was not simply due to differences in the robustness of the noise-related activity: the strength of the features-to-ATL link was not related to decoding success when a model was trained and tested on noise only ($p = 0.72$), suggesting it was specific to memory-generated patterns being more similar to visually generated codes.

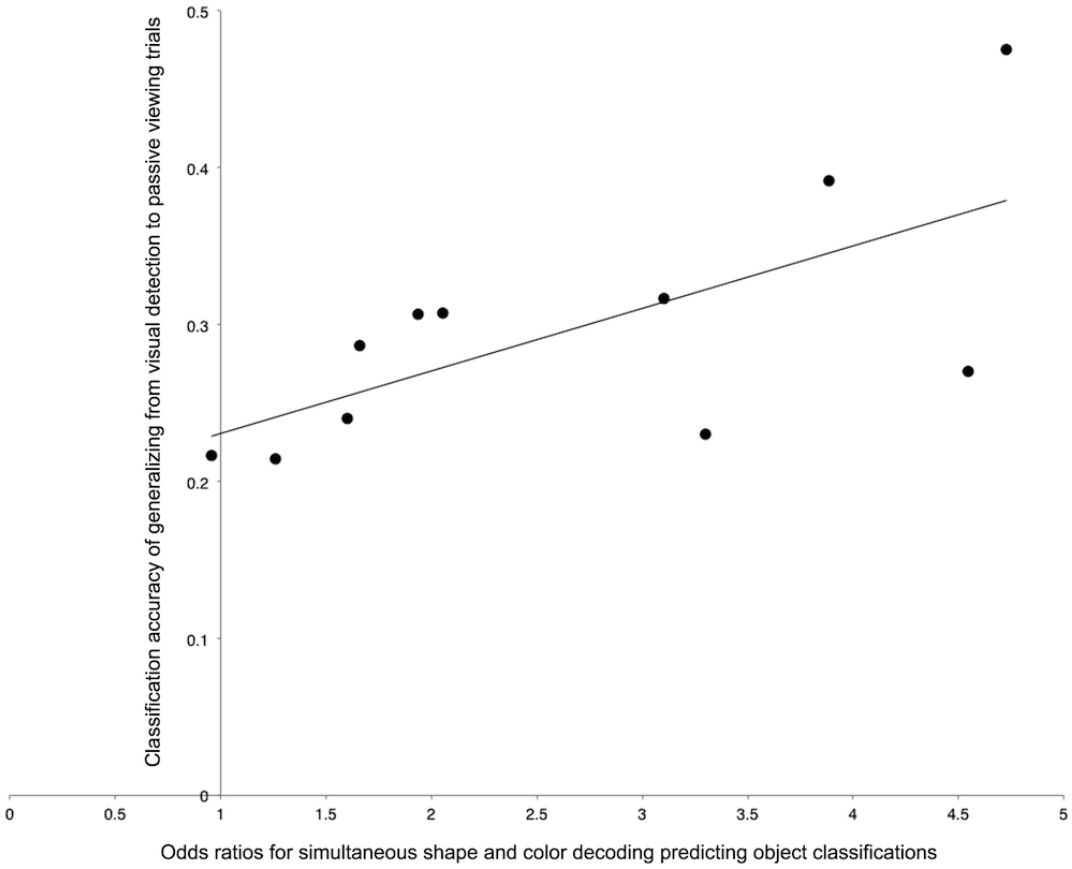


Figure 7: Across-subject differences in noise-to-visual generalization against the strength of the relationship between featural and object-identity decoding. The y-axis represents each subject's classification performance from training on cued noise and testing on visual presentations of each fruit and vegetable in the ATL. The x-axis reflects each participant's odds ratio for the conjunction of color-and-shape decoding (in relevant feature regions) predicting noise-trial identity classifications in the ATL. A logistic regression model generated these odd ratios (details in Method).

Discussion

The results described here provide evidence for theories of memory that posit the existence of a region of integration, such a convergence zone (Damasio, 1989; Meyer & Damasio, 2009; Simmons & Barsalou, 2003). Activity patterns representing an object's identity were present in a region of left ATL, and could generalize to activity patterns produced from passive viewing. Posterior featural regions encoded the anticipated objects' specific shapes and colors. Importantly, these levels of representation were closely linked: ATL-decoding of object identity was more likely when *both* color and shape could be decoded from featural regions. Further, the stronger this relationship across subjects, the more that participants' top-down generated patterns matched visually generated patterns.

The particular cortical site identified here as encoding object identity information is consistent with a variety of patient work that points to the ATL's role in conceptual knowledge (Patterson et al., 2007; Rogers et al., 2007). The findings also support theories that the ATL contains convergence zones between visual components of objects (Damasio, 1989). The results are additionally consistent with proposals that the ATL acts as a central 'semantic hub' (Patterson et al., 2007). Unlike convergence zone proposals, hub-based theories suggest just one central integration zone (in the ATL). As well as a role for the ATL, however, we have found that visual featural regions: i) evoke specific feature-relevant activity patterns from top-down influences, and ii) have a functionally relevant relationship with anterior regions. Shape and color processing are known to be neurally dissociable (e.g., agnosia patients can show impairments in one but not the other: Cavina-Pratesi et al., 2010), arguing that distinct feature systems are sharing information with the ATL and contributing different types of conceptual information (Lambon Ralph et al., 2010). Of the hub-based models, the findings here are most supportive of so-called "hybrid" models, which incorporate important roles for both sensorimotor cortices and a central hub. One such theory, the "hub-and-spoke" model, proposes that modality-specialized regions (V4 and lateral occipital cortex in this study) provide sensory and motor substrates that are combined into an independent high-dimensional representational space in a central hub

(Pobric et al., 2010; Lambon Ralph et al., 2010). The strong and predictive link found between feature and identity-coding regions in these results give weight to a significant role for sensory regions, such as that proposed by the hub-and-spoke model.

The findings of this study may be helpful in interpreting some recent semantic dementia patient findings. A recent study of semantic dementia reported that processing items rich in visual color and form was disproportionately impaired in patients with severe dementia, unlike items with other features such as sound/motion and tactile/action (Hoffman, Jones, & Lambon Ralph, 2012). The authors speculated that temporal lobe atrophy may have spread more posteriorly to affect basic featural regions in these severe cases. Our results suggest a new possible explanation: the patients may have experienced disruption to a key shape and color convergence zone.

Our between-subject analyses showed that participants differed in the resemblance between their top-down generated, and their visually generated, ATL patterns, according to the strength of the ATL's dependence on color-and-shape decoding synchrony. This raises several fascinating questions for future research: Do individual differences in the link between object identity and feature synchrony produce differences in people's phenomenological experiences during processes such as imagery? Are time-points with synchronous color and shape decoding more likely to be accompanied by particularly vivid imagery? The relationship between our findings and peoples' inner visual experiences is an exciting topic for future research.

There are a number of reasons to be confident that the object-identity decoding reported here reflects visual processes, rather than others such as verbal rehearsal. The link between identity and shape-and-color decoding argues strongly for a perceptual basis for the identity decoding, rather than other semantic features such as taste. Further, the ability to generalize decoding from top-down to visually presented objects (where no task was required) supports a visual account.

In our featural analyses, we found shape, but not color, information in the shape region. In contrast, the color region had decodable color information and also shape information at a

trend level. Interestingly, this asymmetry was also reported in a recent meta-analysis of modality-specific imagery, where shape-related activity overlapped with color regions, but not *vice versa* (McNorgan, 2012). Prior work has also suggested that "V4 neurons are at least as selective for shape as they are for color" (Roe et al., 2012, p. 17). Shape curvature is particularly represented in V4 (Roe et al., 2012), which would account for V4 decoding spherical versus elongated shapes at a trend level here. We note that our interpretation of the convergence pattern does not require that the two regions are uniquely selective to color or shape; only that they contain different (i.e., non-redundant) patterns of information.

We employed a novel analysis in this work, enabling us to identify a link between ATL's object-identity code and a conjunction of visual feature decoding in occipital regions. This type of analysis has great potential for future investigations of other configural stimuli, such as multi-sensory interplay (Driver & Noesselt, 2008), to test whether the synchronous emergence of composing features co-occurs with the generation of a higher-level code. Relating this measure to between-subject differences, as we have done here, can help elucidate the behavioral and neural consequences of connected lower-level conjunctions and higher-level representations.

In summary, this study has demonstrated that top-down retrieval of object knowledge leads to activation of shape-specific and color-specific codes in relevant specialized visual areas, as well as an object-identity code within left ATL. Moreover, the presence of identity information in left ATL was more likely when shape and color information was simultaneously detectable in their respective feature regions. The strength of this relationship predicted the correspondence between top-down and bottom-up generated identity activity patterns. These findings support proposals that convergence zones integrate converging featural information into a less feature-dependent representation of identity.

CHAPTER 3: INFORMATIONAL CONNECTIVITY: IDENTIFYING SYNCHRONIZED DISCRIMINABILITY OF MULTI-VOXEL PATTERNS ACROSS THE BRAIN

Abstract

The fluctuations in a brain region's activation levels over a functional magnetic resonance imaging (fMRI) time-course are used in functional connectivity to identify networks with synchronous responses. It is increasingly recognized that multi-voxel activity patterns contain information that cannot be extracted from univariate activation levels. Here we present a novel analysis method that quantifies regions' synchrony in multi-voxel activity pattern discriminability, rather than univariate activation, across a timeseries. We introduce a measure of multi-voxel pattern discriminability at each time-point, which is then used to identify regions that share synchronous time-courses of condition-specific multi-voxel information. This method has the sensitivity and access to distributed information that multi-voxel pattern analysis enjoys, allowing it to be applied to data from conditions not separable by univariate responses. We demonstrate this by analyzing data collected while people viewed four different types of man-made objects (typically not separable by univariate analyses) using both functional connectivity and informational connectivity methods. Informational connectivity reveals networks of object-processing regions that are not detectable with functional connectivity. The informational connectivity results support prior findings and hypotheses about object processing. This new method allows investigators to ask questions that are not addressable through typical functional connectivity, just as MVPA has added new research avenues to those addressable with the general linear model.

Introduction

The enormous wealth of data generated by functional magnetic resonance imaging (fMRI) has driven the continual development of new analytical methods to understand the brain's functions and processes. For many years, a predominant analysis approach has applied the general linear model (GLM) to compare blood oxygenation level dependent (BOLD) univariate activation levels across conditions, regions and subject groups (Friston et al., 1994). The last ten years, however, have seen increased recognition within the fMRI community that information can also be encoded in the activity patterns of populations of voxels. A multitude of studies have now successfully employed multi-voxel pattern analysis (MVPA) techniques to decode information contained within multi-voxel activity patterns (Haynes and Rees 2006; Norman, Polyn, Detre, & Haxby, 2006; O'Toole et al., 2007). Many such studies have reported that their conditions of interest could not be distinguished by the mean voxel-response differences that are assessed in a univariate GLM approach (e.g., Haxby et al., 2001).

In this study, we introduce an analysis method that combines MVPA's access to distributed encoding, with connectivity analyses. Functional connectivity (FC) techniques measure the degree of response-level synchrony between different brain regions or voxels (Biswal, Zerrin Yetkin, Haughton, & Hyde, 1995). The particular measures used to index connectivity (during rest or while performing a task) vary with different approaches (e.g., Friston et al., 1997), but a frequent goal is to identify regions with response levels that fluctuate in a synchronized manner. Just as univariate analyses have led to numerous findings, GLM's cousin – the analysis of *fluctuating* univariate responses of voxels or regions (FC) – has led to results in a wide spectrum of research fields. In this paper, we introduce a method – Informational Connectivity (IC) – that could analogously be considered a cousin of MVPA.

As discussed above, multi-voxel pattern investigations have revealed that one voxel's response magnitude is frequently insensitive to information encoded across a pattern of voxels. Instead of comparing the magnitude of activation levels, multi-voxel analyses frequently employ a machine learning classifier to assess the multivariate discriminability of conditions. While GLM

investigations look to increased or decreased response levels as an indication of relevant neural activity, studies using MVPA often consider the successful separation of conditions as being an indicator of relevant neural information. In this paper, we introduce a method that quantifies the discriminability of multi-voxel patterns in a seed region and identifies regions of the brain that show synchronized discriminability over time.

Whereas FC is frequently applied to measure connectivity between a seed and individual brain voxels, it is (by definition) not possible to measure multi-voxel patterns in single voxels. Instead, we quantify how well a condition can be discriminated from other conditions in the *multi-*voxel patterns at each time-point in a scanning session. We measure the time-course of discriminability for a seed region and for three-dimensional spheres ('searchlights') placed at every location in the brain. We correlate the seed region's discriminability time-course with the equivalent time-course of each searchlight: measuring the simultaneous ebb and flow of multi-voxel distributed information across regions (compared to FC in Figure 1).

	GLM	MVPA
Hypothesized neural target:	Increased or decreased neural processing	Information encoded within neural activity
Target for analysis:	Differences in BOLD univariate responses recorded for conditions	Differences in the distributed multi-voxel activity patterns recorded for different conditions
Measures submitted to analysis:	Increases or decreases in voxel / ROI levels of univariate activation to one condition versus another (typically submitted to an ANOVA)	Discriminability or non-discriminability of an ROI's multi-voxel patterns representing one condition versus another (typically binarized through a classifier)

	FUNCTIONAL CONNECTIVITY	INFORMATIONAL CONNECTIVITY
Hypothesized neural target:	Shared increases or decreases in neural processing across regions	Shared presence of information within neural activity across regions
Target for analysis:	Synchronized changes in levels of univariate activation over a timecourse	Synchronized changes in the presence of multi-voxel patterns over a timecourse
Measures submitted to analysis:	ROIs' / voxels' changes in univariate activation over a timecourse (typically correlated)	ROIs' changes in multi-voxel pattern discriminability over a timecourse (correlated)

Figure 1: The relationship between Informational Connectivity and other fMRI measures.

Since the conference presentation of an earlier version of this work (Coutanche and Thompson-Schill, 2011), Chiu and colleagues (2012) have employed a functional connectivity framework to identify voxels that vary in univariate responses for two cognitive states that were identified by a multivariate classifier in a region-of-interest (ROI). Our approach contrasts with this by identifying regions that have synchronized discriminability of multi-voxel information (rather than changing univariate activation). This makes our technique available for examining conditions that are not accompanied by differing univariate responses. Multivariate techniques

have previously been applied in alternative connectivity approaches (such as the application of information-theoretical measures; Chai, Walther, Beck, & Fei-Fei, 2009; Lizier, Heinzle, Horstmann, Haynes, & Prokopenko, 2011). Our approach contrasts with prior work that has applied multivariate analyses to functional connectivity results (e.g., Welchew et al., 2005), by employing its own metric (instead of analyzing univariate change) to track multi-voxel pattern discriminability, building on the success of MVPA at detecting information inaccessible to univariate measures. This distinction is analogous to the difference between using MVPA and applying multivariate analyses to a GLM map. Although both approaches might yield interesting results, MVPA is specifically sensitive to the distributed condition-information in populations of voxels.

Here, we describe our method by example, and examine its effectiveness by applying it to a classic dataset from Haxby and colleagues (Haxby et al., 2001; later analyzed in: Hanson, Matsuka, & Haxby, 2004; O'Toole, Jiang, Abdi, & Haxby, 2005; Raizada and Connolly, 2012). For simplicity, and to test our technique's sensitivity to conditions that are distinguishable by potentially subtle differences in activity patterns, we restrict our analyses to time-points associated with presentations of four man-made object categories. We select six seed regions and identify brain areas that are informationally connected to each. We compare these results to a conventional FC analysis. The possible differences between these two methods include IC revealing: a subset of FC (selectivity), a superset of FC (sensitivity), a different set of regions, or no regions. We predicted that IC would identify more areas of cortex than FC, based on findings that multivariate decoding can detect information that a typical GLM cannot (Haxby et al., 2001) and a recent direct comparison of MVPA and GLM showing that MVPA can identify more areas of relevant cortex (Jimura and Poldrack, 2012). IC has MVPA's sensitivity and access to distributed information that is not obtainable from univariate responses. The larger IC networks might include the FC regions (i.e., a superset) or there may be little overlap. In their comparison of MVPA and GLM results, Jimura and Poldrack (2012) noted that a "conjunction of the two analyses revealed relatively small commonality in significant results across the brain" (p. 549),

leading us to predict that informational and functional networks may be largely distinct. One consequence of this predition is an expectation that some regions will be identified based on common univariate synchrony (FC) but not multivariate synchrony (IC). This hypothesis is supported by prior findings that univariate differences can sometimes identify regions that are not identified from MVPA (Jimura and Poldrack, 2012; Quamme, Weiss, & Norman, 2010).

Method

Stimuli and experimental design

Full experimental details are available from the original manuscript employing this data (Haxby et al., 2001), but the relevant details are as follows. Participants were presented with 24-second blocks (separated by 12 s of rest) of gray-scale photographic images belonging to one of eight categories: faces, houses, cats, scrambled images, bottles, chairs, shoes and scissors. For these analyses, we focused on the latter four categories (all man-made objects). Within blocks, stimuli were presented for 500 ms with an interstimulus interval of 1,500 ms. Participants identified object repeats (1-back) with a button-press. One block of every category appeared in each of twelve runs (excepting one participant where eleven runs were available). Analyses were performed on data for all runs from the five participants with anatomical T1-images and functional datasets available. The condition-labels for the time-points were shifted by two TRs for the multi-voxel pattern and IC analyses to account for the hemodynamic delay, giving nine TRs for each block and 108 for each condition across the experiment.

Imaging pre-processing

Hemodynamic changes were recorded with gradient echo echo-planar imaging with a 3T scanner (repetition time (TR) = 2.5 s, forty 3.5 mm thick sagittal slices, TE = 30 ms, flip angle = 90; Haxby et al., 2001). The functional data were slice-time corrected, motion-corrected, aligned to the subject's anatomical image and detrended with a second order polynomial. The anatomical image and functional data were transformed into standardized Talairach space with unchanged voxel resolution (3.5 x 3.75 x 3.75 mm for functional data). For the IC analyses, the effects of motion and global signal were removed from the data by modeling six motion parameters (pitch, roll, yaw, x, y, z) and mean white matter signal, and then using the residuals for subsequent analyses. This is equivalent to including motion and white matter signal as covariates in a FC model. The white matter signal was extracted using SPM8's segmentation procedure, which classifies voxels into gray matter, white matter and cerebrospinal fluid based on

image intensity and prior probabilities of the distribution of tissue types. A threshold of 0.75 was employed to select white matter voxels. The Analysis of Functional NeuroImages (AFNI) software package was used for preprocessing and relevant univariate analyses (Cox, 1996). Prior to MVPA and IC analyses, each voxel's task and rest data were z-scored within each run; normalizing the run's timeseries to have a mean of zero and unit variance.

Seed regions

We examined IC and FC for six empirically determined seed regions: two regions identified by both an MVPA searchlight and GLM group map; two regions found from the MVPA searchlight but not the GLM; two regions found in the GLM but not the searchlight. To create the relevant group MVPA searchlight map, each individual's dataset was submitted to a 4-way correlation-based classifier (a popular classification approach) to separate activity patterns from the four types of man-made objects. We implemented a roaming searchlight analysis (Kriegeskorte, Goebel, & Bandettini, 2006), where a spherical volume (3-voxel radius) is centered on each brain voxel in turn and an analysis (in this case, classification) is conducted using data from the voxels included within the searchlight volume. For each searchlight, a leave-one-run-out cross-validation procedure trained on eleven runs and tested on the twelfth. Each testing TR's vector of activity values was correlated with the mean activity pattern for each of the four conditions in the training set. The condition that was most strongly correlated with the testing time-point determined the classifier's prediction for that TR. Classifier performance was calculated as the proportion of correctly predicted time-points (chance = 25%). The classification accuracy from each searchlight was allocated to its central voxel for mapping purposes. Individual searchlight maps were smoothed (9 mm Full-Width at Half Maximum; FWHM) and subjected to a one-way group *t*-test for performance above chance. As this was performed purely to identify seeds, we adopted a liberal threshold of $p<0.005$ and cluster size of at least 5 voxels.

To create a group GLM map, each individual's dataset was submitted to a typical univariate analysis with six motion parameters as covariates. As the above searchlight analysis

attempted to distinguish the four man-made objects, we ran a similar analysis with the GLM: running six pairwise comparisons, smoothing each individual's pairwise maps (9 mm FWHM) and submitting the maps for each comparison to a group analysis. The six group maps were then thresholded at $p < 0.005$ and a union of the six maps was created. A 5-voxel cluster threshold was then applied. Relatively few voxels survived even this liberal threshold, as expected from prior literature showing that object identity is typically not identifiable from univariate differences (Haxby et al., 2001).

The six seeds were created by selecting the central voxels of the two largest cluster volumes found only in the searchlight map, the two largest found only in the GLM map (although as discussed above, this was at a sub-significant level), and the two largest found in both maps. Selecting the seed locations based on the largest clusters (rather than statistical peaks) gave confidence that the majority of voxels in the seeds had the desired characteristic (e.g., condition-differences in a GLM), and is also consistent with findings of greater reliability from cluster-based statistical thresholds (e.g., Thirion et al., 2007). The seeds were located in the right inferior occipital gyrus, left inferior occipital gyrus, left fusiform gyrus, left superior temporal sulcus, right supramarginal gyrus and right postcentral sulcus (coordinates in Table 1). A 3-voxel radius sphere (with a volume of 123 voxels) was placed at each central voxel to create each seed.

Informational connectivity

The metric underlying informational connectivity quantifies how robustly the real class's activity pattern (versus the alternative classes) becomes discriminable at points along the timeseries. During correlation-based MVPA, the activity pattern at a time-point (i.e., a vector of voxel activations *m*-voxels long, recorded at that time) is compared to the mean voxel activity pattern corresponding to each condition in the held-out training set (i.e., the mean vector of each condition that is calculated by averaging the condition's time-points). We quantified multi-voxel pattern discriminability for each time-point with the following procedure (also captured in the formulae below):

1. Calculate the Pearson correlation coefficient between the i) vector of voxel activation values for that time-point (i.e., its activity pattern) and ii) vector of mean voxel activation values for the time-point's condition in the training data (i.e., the prototypical activity pattern for the condition). Fisher-transform to z-score.
2. Calculate the Pearson correlation coefficient between the i) vector of voxel activation values for that time-point (i.e., its activity pattern) and ii) vector of mean voxel activation values for each *alternate* condition in the training data (i.e., the prototypical activity patterns for the rival conditions).
3. Identify the highest correlation from step 2 (i.e., the highest similarity to an 'incorrect' condition). Fisher-transform to z-score.
4. Multi-voxel Pattern Discriminability = Step 1 – Step 3 (i.e., Relationship to condition's prototypical pattern minus Relationship to the most similar incorrect condition).

The procedure is formalized in the below formulae, where $\overline{x}$ is the normalized 1-by-*m* row vector of *m* voxel activation values at time-point *n,* $\overline{y}$ is the normalized 1-by-*m* row training data vector of mean *m* voxel activation values for the correct (*c*) or incorrect (*i*) conditions relating to time-point *n*. In the analyses conducted here, *m* was 123 (the searchlight volume), and *n* ranged from 1 to 432. The artanh function normalizes the correlation coefficients through Fisher's transform.

$$r_c[n] = \frac{\overline{x}[n].\overline{y}_c'}{m-1}$$

$$r_i[n] = \max\left(\frac{\overline{x}[n].\overline{y}_i'}{m-1}, \forall i \neq c\right)$$

Multi-voxel pattern discriminability = $\operatorname{artanh}\left(r_c[n]\right) - \operatorname{artanh}\left(r_i[n]\right)$

This multi-voxel pattern discriminability metric is calculated for each time-point across the timeseries, giving a dynamic series of values across the fMRI session (see Figure 2). This metric can be intuitively related to the typical binary metric used in classification analyses: The commonly used correlation-based classifier would successfully predict a time-point's condition when its data give a discriminability value above zero. This type of classifier makes a prediction for each time-point based on which class's training pattern is most strongly correlated with the time-point's activity pattern. In our measure, discriminability values are positive when a time-point's multi-voxel pattern is most strongly correlated with the training pattern of the correct class (i.e., the condition that was shown to participants). Positive discriminability values therefore reflect that a time-point's condition can be successfully predicted. A negative value on the other hand, reflects that the training pattern for a non-present (rival) class has the highest correlation with the current time-point, which would lead to an incorrect prediction.

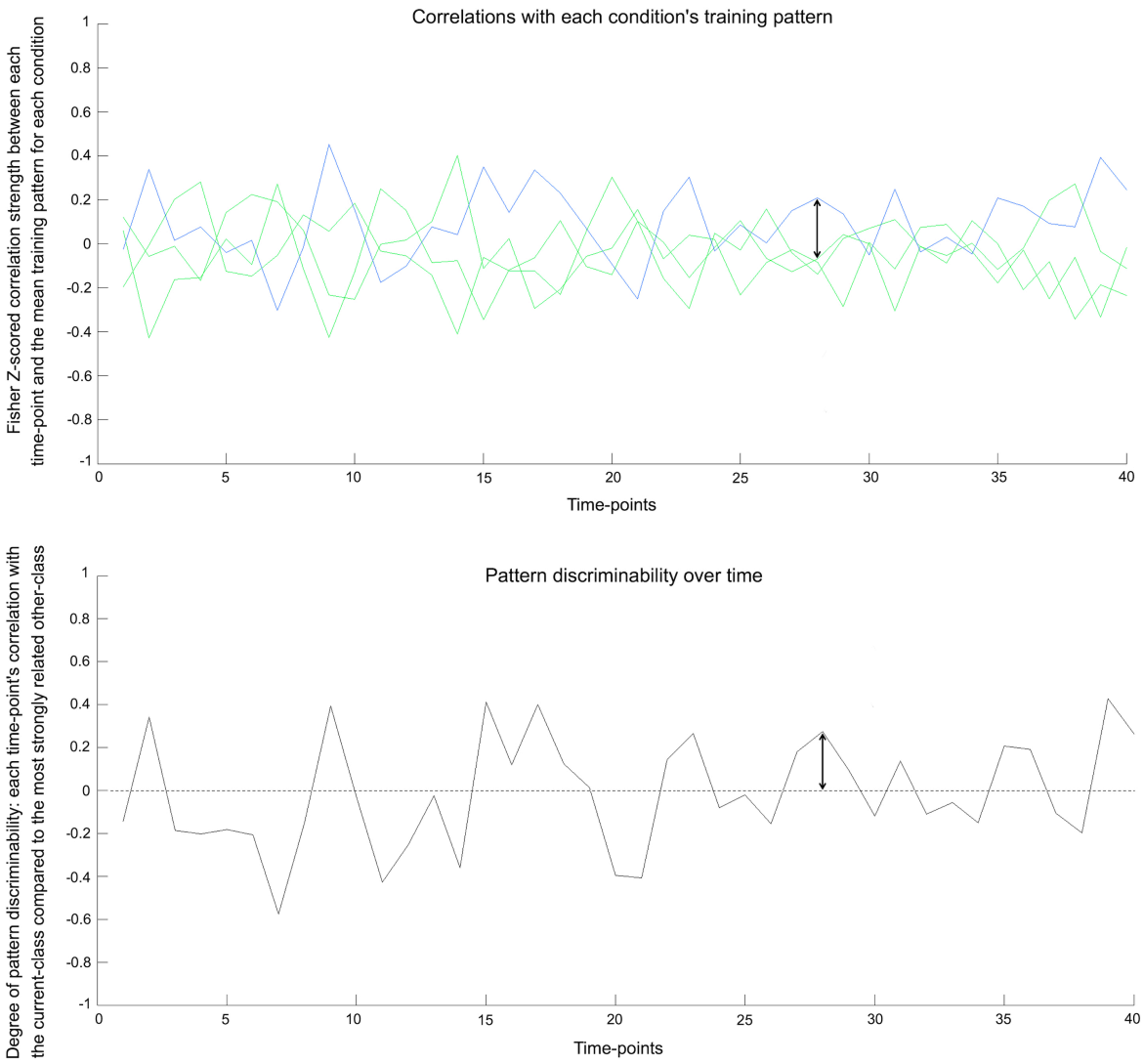


Figure 2: Pattern discriminability over time in real data. Top: The underlying basis for the pattern discriminability metric – shown here for the bottle condition in one seed in one subject. The blue line represents each time-point's Fisher z-scored correlation with the training pattern for the correct class. The green lines show the correlation values with mean training patterns for the three other classes. Bottom: Pattern discriminability is calculated by taking the correlation with the correct class's mean training pattern and subtracting the correlation strength of the strongest incorrect class (see text for details). When a time-point's value surpasses zero, it would reflect a

classifier successfully predicting that time-point's condition. The arrow shows the corresponding values between the plots.

To create an IC map, multi-voxel pattern discriminability is calculated for the timeseries of the seed region, followed by the timeseries of every searchlight sphere identified in the roaming searchlight procedure described above (Kriegeskorte et al., 2006). The timeseries of pattern discriminability from the seed region (i.e., a vector N-trials long) is then correlated with each searchlight's timeseries of discriminability, through a non-parametric Spearman's rank correlation. The resulting r_s-value (representing the strength of the relationship between searchlight and seed) is placed at the voxel that lies at the center of each searchlight (a typical approach to mapping searchlight results; Kriegeskorte et al., 2006). This produces a brain map of values that each reflects how closely the timeseries of multi-voxel pattern discriminability for that (searchlight) area matches the equivalent timeseries of the seed region. The map therefore shows how strongly brain regions are correlated in terms of pattern discriminability (i.e., how 'informationally connected' they are) with the region-of-interest (the seed). Each participant's map is then Fisher-transformed into z-scores, spatially smoothed (8 mm FWHM) and tested for values above zero (i.e., asking which searchlights are significantly correlated with the seed) in a one-way group *t*-test. The tools and scripts for running these analyses are freely available in the Informational Connectivity Toolbox (http://www.informationalconnectivity.org). Statistical significance was tested using the same procedure (described below) for both IC and FC to enable direct comparisons.

Functional connectivity

The IC results were compared to results from a typical FC analysis. We assessed FC for the same TRs analyzed using IC (TRs associated with the four man-made objects). The timeseries of mean activation values for the TRs was extracted for each seed region. This timeseries was then used as a predictor in a whole-brain GLM analysis, with six motion

parameters and mean white matter signal as covariates. Individuals' maps of correlation values, reflecting the correspondence between voxels' and each seed's timeseries, were converted to Fisher-transformed z-scores and spatially smoothed (8 mm FWHM). All subjects' maps were subjected to a one-way group *t*-test for values greater than zero. The method for significance testing is outlined below.

Significance testing

We adopted the same significance testing approach for both IC and FC to enable direct comparisons. For each seed region, the group statistical *t*-maps were first thresholded at $p < 0.001$ (and also at $p < 0.005$ to ensure that the results are not dependent on a particular *t*-threshold) for positive *t*-values in a one-way test to identify regions that were positively correlated with the seed. To correct for multiple-comparisons, we employed permutation testing to determine the minimum cluster size required for corrected significance. The seed's timeseries of values (pattern-discriminability values for IC; univariate activation values for FC) were shuffled by randomly swapping blocks of presentations (i.e., moving the sets of nine contiguous TRs that were separated by rest). One thousand group maps were created (constructed by randomly sampling from a set of one hundred permuted maps for each subject) and submitted to a group test in the same manner as the seed's real (non-permuted) time-course, including thresholding at $p < 0.001$. This gave a null distribution of 1,000 group maps. We used this to determine the minimum cluster size needed in the real (non-permuted) group map for a corrected p-value of < 0.05, by identifying the largest cluster size in each of the 1,000 permuted maps. The 50th largest cluster size from this null distribution is the cluster size that would be expected by chance five times out of 100 (i.e., $p < 0.05$). Any clusters larger than this in the true (non-permuted) group map are significant at $p < 0.05$ corrected. This approach has the advantage of correcting in a manner that accounts for the dataset's own level of smoothing, as each permutation undergoes the same processing as the true order. The minimum cluster sizes were calculated separately for every seed and the two connectivity approaches.

Results

We analyzed a dataset collected while subjects viewed blocks of images of four types of man-made objects, using our novel informational connectivity method to track and compare dynamic change in discriminability of multi-voxel patterns across time. We compared these results to a typical functional connectivity analysis that tracks synchronized changes in univariate activation. We employed six seeds, selected from regions showing univariate variation between conditions, MVPA decoding, or both.

The IC and FC analyses identified different networks of regions, with IC revealing larger networks than FC in this man-made object dataset (Figure 3; Figure 5; Table 1). These FC results were not specific to the p-value selected: Repeating the FC analysis with a more liberal p-value ($p < 0.005$ with a permutation-generated minimum cluster size) generated similar networks of regions. The different seeds varied in how many regions were informationally connected to them: for example, the right postcentral sulcus seed was informationally connected with a large variety of cortical areas, while the left inferior occipital gyrus seed was not (Figure 3).

Region	**Informationally connected clusters**				**Functionally connected clusters**			
	Volume (voxels)	**x**	**y**	**z**	**Volume (voxels)**	**x**	**y**	**z**
Univariate Seed 1: Right Postcentral Sulcus [x=39, y=-42, z=45]								
Left Precuneus	3871*	-11	-43	36				
Left Fusiform Gyrus	3871*	-30	-43	-10				
Left Fusiform Gyrus	3871*	-37	-58	-7	19	-33	-56	-18
Left Middle Temporal Gyrus	3871*	-45	-58	19				
Left Superior Temporal Gyrus	318*	-54	0	-6				
Left Superior Temporal Gyrus	112	-54	-30	12				
Left Parahippocampal Gyrus	83*	-22	-12	-22				
Left Temporal Pole	83*	-19	8	-26				

Left Anterior Cingulate	78	-16	45	1				
Left Inferior Parietal Lobe	3871*	-47	-43	47	31	-37	-38	53
Left Orbital Gyrus	318*	-35	27	-12				
Left Inferior Frontal Gyrus	318*	-51	11	1				
Left Middle Frontal Gyrus	87	-40	19	34				
Left Superior Frontal Gyrus	3871*	-11	0	53				
Left Superior Frontal Gyrus	3871*	-7	17	57				
Left Caudate	170	-5	19	4				
Right Inferior Occipital Gyrus	3871*	37	-67	-7				
Right Fusiform Gyrus	3871*	25	-78	-13				
Right Fusiform Gyrus	3871*	40	-36	-14				
Right Superior Temporal Gyrus	3871*	39	-26	9				
Right Precentral Gyrus	3871*	55	2	22				
Right Supplementary Motor Area	3871*	5	-19	54				
Right Inferior Frontal Gyrus	3871*	46	33	6				
Right Inferior Frontal Gyrus	3871*	34	7	31				
Right Middle Frontal Gyrus	3871*	32	33	19	21	37	26	34
Right Superior Frontal Gyrus	48	19	8	53				
Right Cerebellum	39	44	-41	-44				
Right Cerebellum	3871*	43	-43	-46				
Right Cerebellum	3871*	23	-55	-45				
Right Cerebellum	3871*	24	-43	-27				
Right Thalamus	3871*	8	-15	1				
Univariate Seed 2: Right Supramarginal Gyrus [x=49, y=-24, z=35]								
Left Lingual Gyrus	96	-2	-79	4				
Left Parahippocampal Gyrus	59	-18	-25	-13				

Left Middle Temporal Gyrus	29	-54	-56	19				
Left Cingulate Gyrus	306*	-12	7	30				
Left Cingulate Gyrus	911*	-2	-20	44				
Left Supramarginal Gyrus	911*	-60	-17	33				
Left Precentral Gyrus	911*	-29	-21	61				
Left Inferior Frontal Gyrus	39	-30	23	-14				
Left Cerebellum	89	-51	-56	-26				
Left Cerebellum	66	-12	-68	-37				
Left Thalamus	306*	-4	-10	15				
Right Fusiform Gyrus	37	37	-4	-29				
Right Superior Frontal Gyrus	41	3	44	36				
Right Cerebellum	87	30	-34	-26				
Right Putamen	82	30	-11	-3				
Left Supramarginal Gyrus					12	-58	-26	23
Left Precentral Gyrus					10	-51	4	23
Left Postcentral Gyrus					27	-44	-30	42
Right Postcentral Gyrus					87	47	-8	16
Multi-voxel Seed 1: Left Inferior Occipital Gyrus [x=-30, y=-75, z=-7]								
Left Calcarine Sulcus	225	-14	-96	-5				
Left Fusiform Gyrus	42	-33	-29	-23				
Left Superior Parietal Lobe	28	-18	-63	51	13	-26	-60	42
Left Orbital Gyrus	22	-40	47	-5				
Left Cerebellum	23	-37	-75	-37				
Left Insula	22	-33	-7	16				
Right Inferior Occipital Gyrus	66*	29	-85	-12				
Right Middle Occipital Gyrus	66*	26	-86	12				

Right Cerebellum	28	45	-67	-26				
Left Middle Occipital Gyrus					40	-26	-60	-11
Left Middle Occipital Gyrus					11	-33	-79	27
Multi-voxel Seed 2: Left Superior Temporal Sulcus [x=-51, y=-41, z=8]								
Left Calcarine Gyrus	2401*	-15	-71	12				
Left Fusiform Gyrus	2401*	-45	-40	-22				
Left Inferior Temporal Gyrus	2401*	-39	0	-26				
Left Parahippocampal Gyrus	43	-19	-8	-29				
Left Superior Parietal Lobe	2401*	-30	-64	51				
Left Postcentral Gyrus	2401*	-27	-30	50				
Left Inferior Frontal Gyrus	2401*	-53	14	2				
Right Middle Occipital Gyrus	95*	33	-82	7				
Right Lingual Gyrus	2401*	16	-96	-7				
Right Inferior Temporal Gyrus	95*	47	-59	-2				
Right Angular Gyrus	78	42	-70	38				
Right Supramarginal Gyrus	74	58	-41	38				
Right Precentral Gyrus	45	33	-23	57				
Right Cerebellum	2401*	47	-59	-33				
Right Cerebellum	2401*	12	-55	-15				
Right Insula	146	35	-19	12				
Common Seed 1: Right Inferior Occipital Gyrus [x=45, y=-61, z=-8]								
Left Fusiform Gyrus	76	-44	-56	-14				
Left Middle Temporal Gyrus	68	-54	-39	-5				
Left Supramarginal Gyrus	36	-65	-30	34				
Right Precuneus	31	20	-48	36				
Right Middle Temporal Gyrus	434	49	-72	12				

Right Inferior Parietal Lobe	113*	43	-50	53				
Right Supramarginal Gyrus	113*	58	-41	38				
Right Superior Frontal Gyrus	28	12	15	42				
Right Superior Frontal Gyrus	62	16	53	1				
Right Inferior Occipital Gyrus					97*	30	-84	-8
Right Inferior Temporal Gyrus					97*	56	-53	-8
Common Seed 2: Left Fusiform Gyrus [x=-38, y=-40, z=-16]								
Left Middle Occipital Gyrus	101	-41	-67	7				
Right Middle Occipital Gyrus	101	51	-65	11				
Right Supramarginal Gyrus	49*	64	-40	29				
Right Inferior Parietal Lobe	49*	53	-47	44				
Left Superior Occipital Gyrus					12	-29	-71	26
Right Fusiform Gyrus					24	48	-54	-14

Table 1: Significantly connected regions for IC and FC analysis methods. Significant regions are displayed for IC and FC (at $p < 0.001$ and cluster sizes determined by permutation testing). Similarly located regions are listed in the same row. Clusters significant at the seed's location are not listed to avoid circularity. Coordinates represent the peak of significant voxel-clusters. An asterisk indicates that the cluster contained multiple peaks, each included separately.

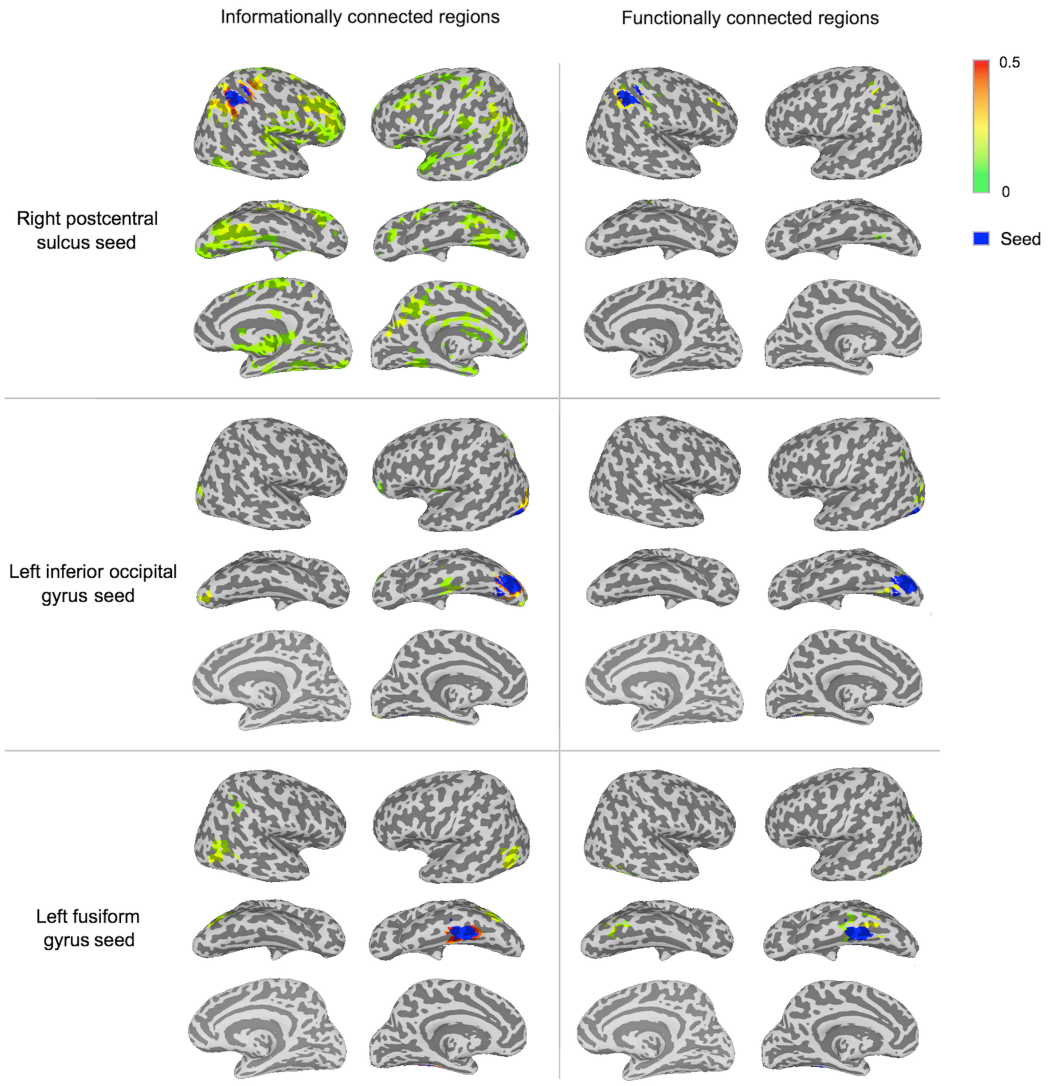


Figure 3: Significantly connected regions in IC and FC analyses for three of the seeds. A group *t*-test ($p < 0.001$ with minimum cluster size from permutation testing) determined significance (described in the Methods). Connectivity strength is displayed between green (lower values) and red (higher values). Each seed region is shown in blue.

To visualize the two methods' results without a minimum spatial extent, Figure 4 shows IC and FC connectivity before applying the cluster-based permutation thresholds. By visualizing the degree of overlap in regions that were significantly informationally and functionally connected

with each seed, we found that the two methods identified either largely distinct or slightly overlapping networks of regions (Figure 5). This is also reflected in the small number of regions that are listed under both methods in Table 1. Many of the areas showing synchronous multi-voxel pattern discriminability include regions that have been implicated in object processing. Evidence underlying this involvement is presented in the discussion.

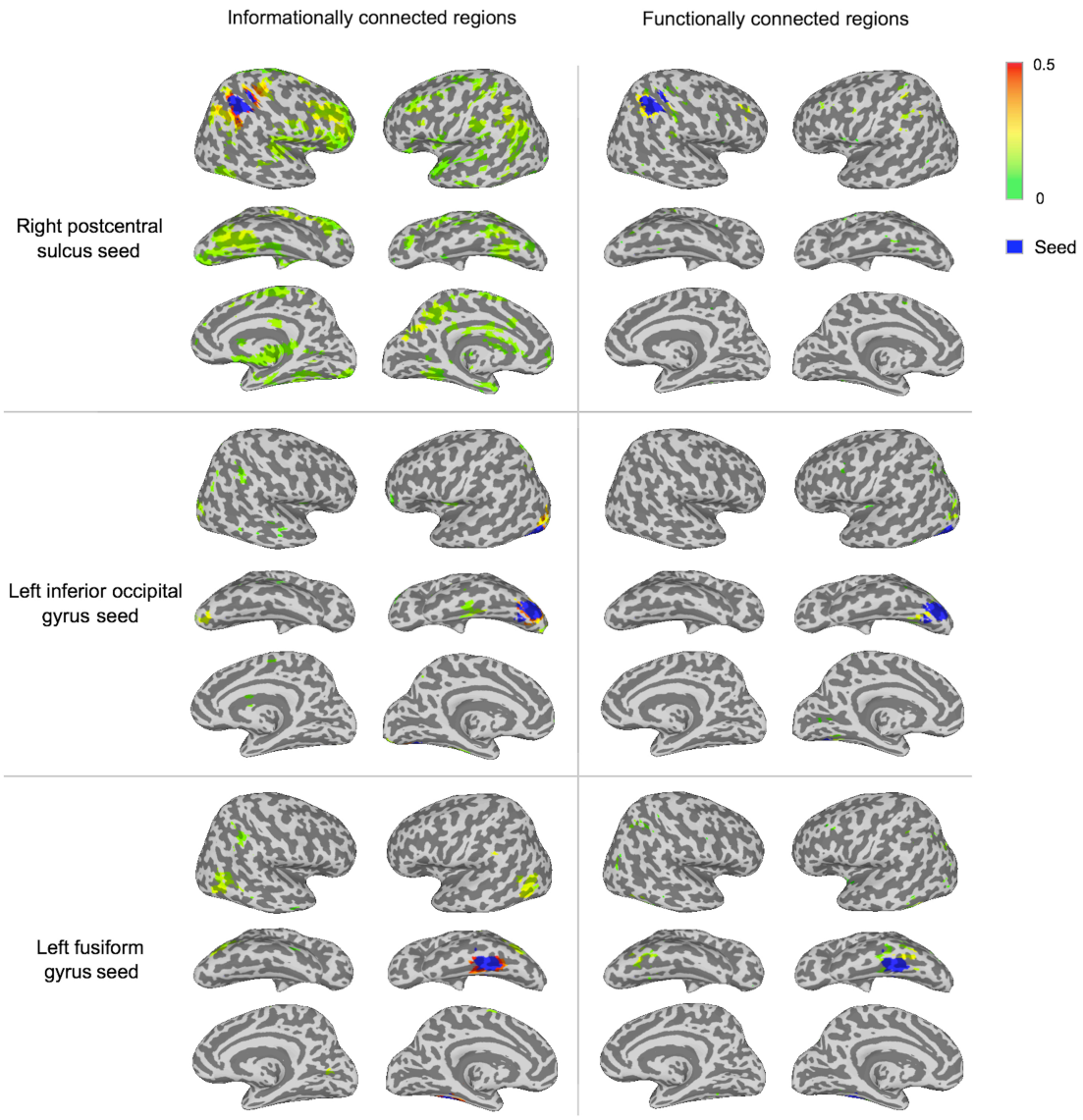


Figure 4: Connectivity strengths before cluster-based thresholding for three of the seeds. The displayed regions have connectivity above zero from the group *t*-test at $p < 0.001$ prior to

thresholding in cluster-based permutation tests, to visualize sub-threshold connectivity for both methods. Connectivity strength is displayed between green (lower values) and red (higher values). Each seed region is shown in blue.

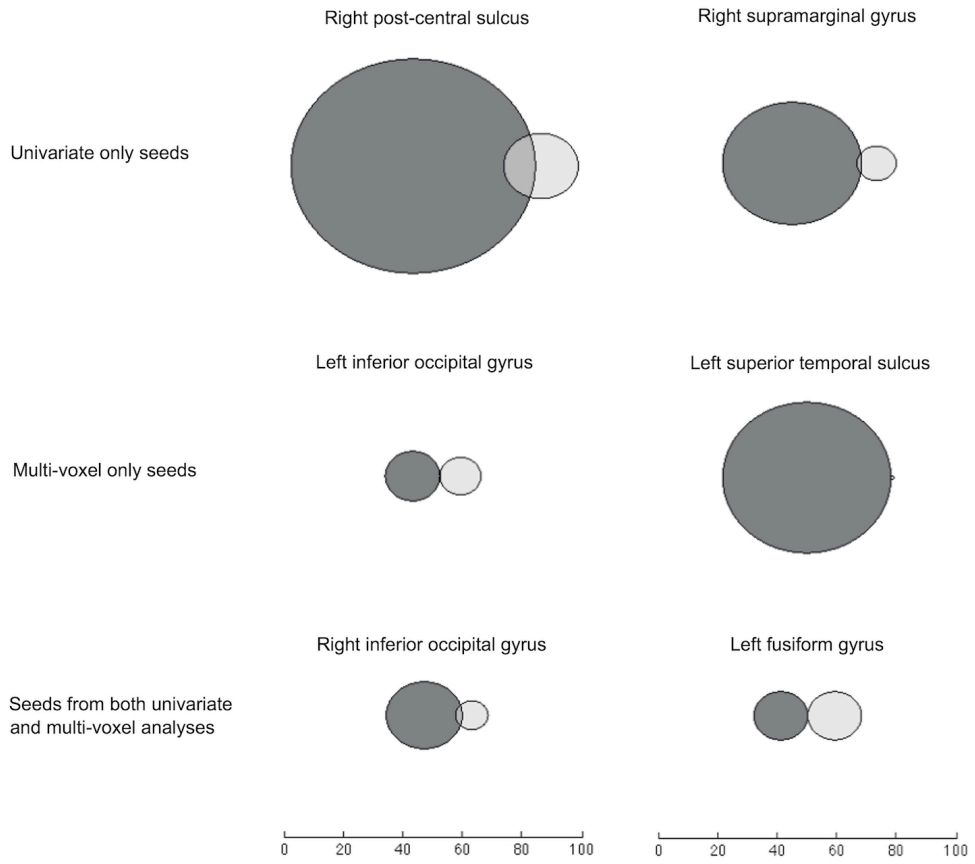


Figure 5: Venn diagrams of voxels significantly connected to each seed through IC (dark gray) and FC (light gray). Searchlights that overlapped with the relevant seed region have been removed. Here, FC results come from an analysis using the timeseries of *searchlights'* (rather than voxels') mean values, to give a suitable comparison with the searchlight-based IC results.

We examined the univariate and multivariate characteristics of searchlights, relative to their strength of IC and FC, and confirmed that the IC approach can highlight regions that would otherwise be ignored by FC. For example, regions with low univariate activation to conditions, yet

decodable multi-voxel information, were ignored by FC, but detected with IC. This can be seen in Figure 6, which shows the group average mean activation, multivariate information and connectivity strength (with the left fusiform gyrus seed) for searchlights across the brain. The empty space visible in the top-left octant (representing searchlights with low response levels despite high decoding accuracy) in the FC, but not IC, graph highlights connectivity that is inaccessible to univariate FC. This pattern was representative of connectivity with other seeds.

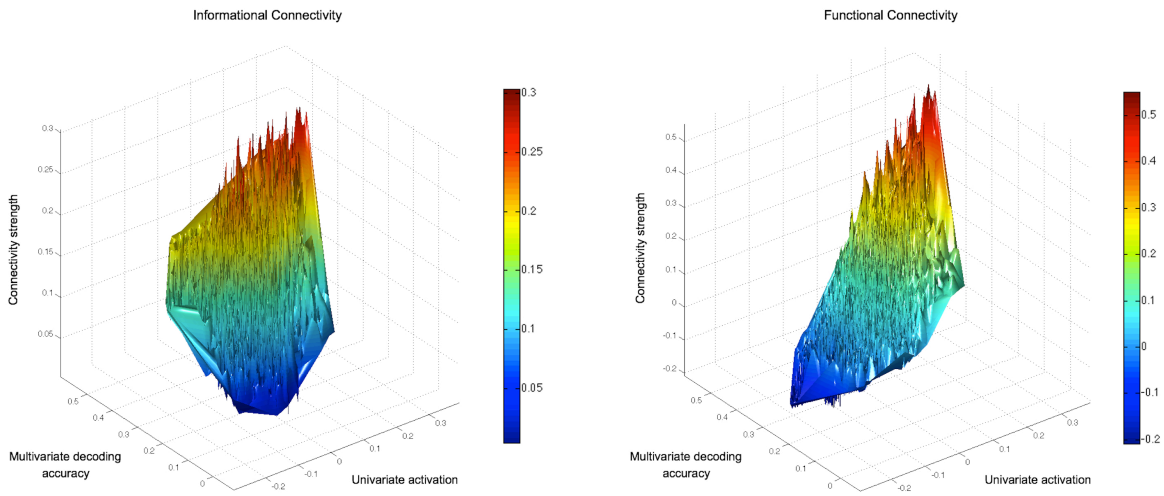
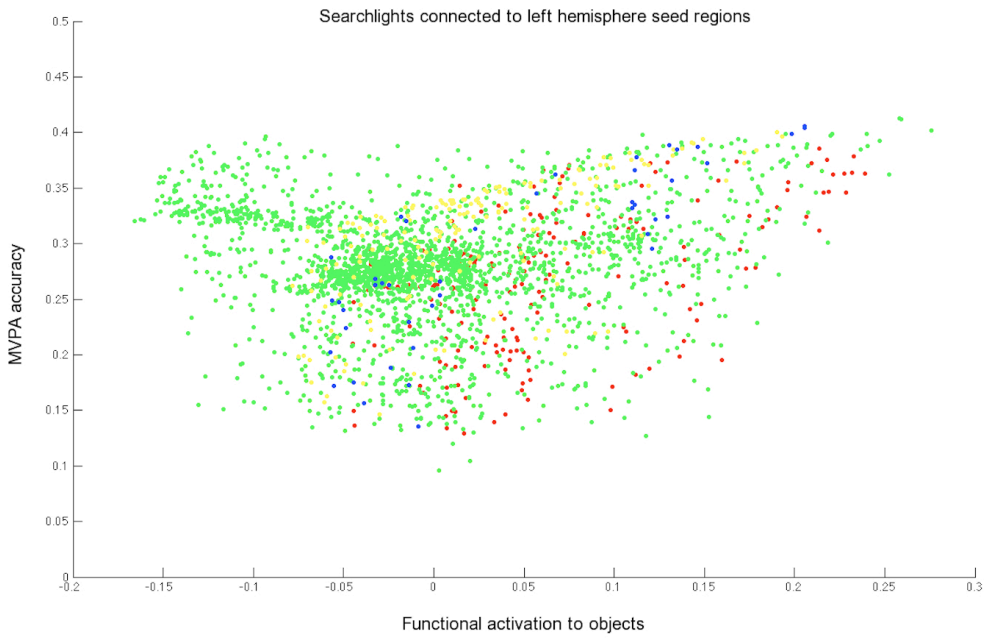


Figure 6: Connectivity strengths of all searchlights with a seed in the left fusiform gyrus (present in both the GLM and MVPA searchlight results). The IC and FC results for every brain searchlight are displayed relative to the searchlight's mean univariate activation to the objects and decoding accuracy in a 4-way classification of object type. Searchlights that overlapped with the seed region have been removed. The FC values reflect the described FC approach, using each searchlight's mean timeseries (rather than each voxel's timeseries) to give a suitable comparison with IC (which reflects information in a searchlight volume). The empty space visible in the top-left octant of the FC graph for searchlights with low response levels (despite high decoding accuracy) highlights connectivity that is inaccessible to univariate FC.

The IC networks detected for each seed were not redundant with each other. A large proportion of searchlights were significantly connected with only one seed (Figure 7) and

although some searchlights were identified in the networks of two seeds (blue in Figure 7), very few were found for three. The distinctiveness apparent for different seed networks also confirms that informational connectivity is not redundant with conducting a typical MVPA searchlight analysis, as it can highlight distinct networks based on the selected seed.



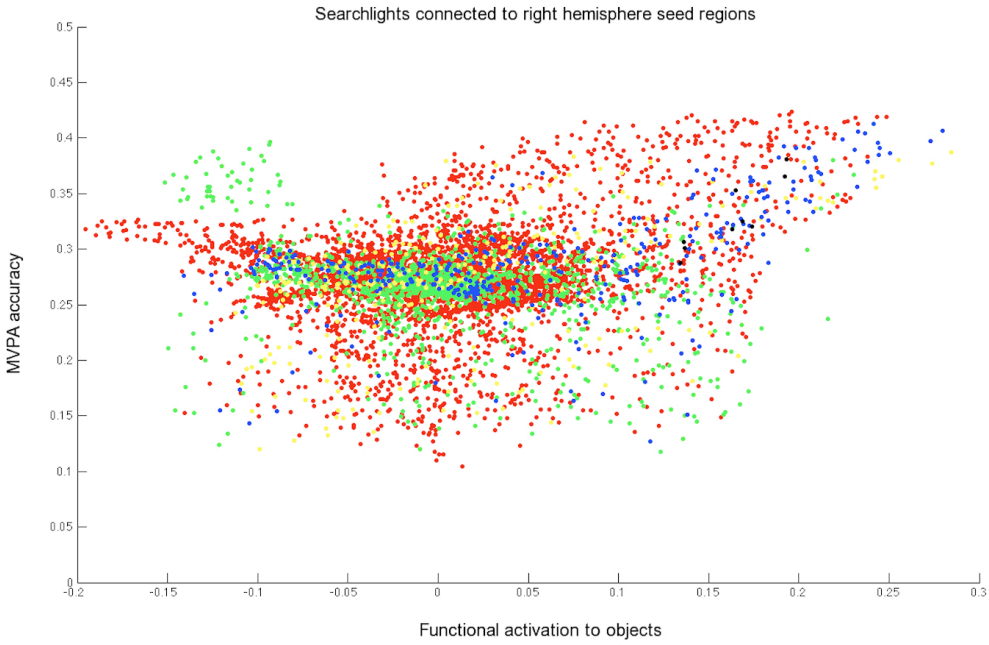


Figure 7: Searchlights with significant informational connectivity to at least one of the three left hemisphere seeds (top) and at least one of the three right hemisphere seeds (bottom), shown against MVPA accuracy and mean functional activation. The green, yellow and red colors each represent searchlights that are connected with just one seed. Blue points show searchlights that are connected to two seeds and black points show searchlights connected to three seeds. Searchlights overlapping with one of the three seed regions were removed from each scatterplot.

Discussion

This paper has presented a new method – informational connectivity – for measuring synchronous discriminability of multi-voxel patterns across the brain. We have described a metric for quantifying multi-voxel pattern discriminability across a timeseries, and conducted an informational connectivity analysis on data collected as subjects viewed four types of man-made objects. The IC method identified networks of synchronized regions that were not identified by functional connectivity. Many of these brain areas are linked to object processing (discussed below), suggesting that multi-voxel pattern discriminability can identify networks involved in processing conditions that are characterized by multi-voxel information (such as perceiving objects).

The limited overlap of regions identified by IC and FC is consistent with a prior report of low commonality between MVPA and univariate measures, with MVPA having greater sensitivity overall (Jimura and Poldrack, 2012). GLM and MVPA approaches have been conceptualized as tapping basic processing (causing changes in univariate activation) versus representations of the content being processed (causing changes in pattern discrimination; Jimura and Poldrack 2012; Mur, Bandettini, & Kriegeskorte, 2009; although MVPA has also been applied to identify cognitive processes e.g., Esterman, Chiu, Tamber-Rosenau, & Yantis, 2009). MVPA investigations into representational content, such as the type of man-made object being processed, have proven effective for advancing our understanding of the visual system (Eger, Ashburner, Haynes, Dolan, & Rees, 2008) and others (e.g., auditory system: Lee, Janata, Frost, Hanke, & Granger, 2011). Analogously, identifying networks characterized by synchronized discriminability of multi-voxel information will be valuable for investigators wishing to study how systems of brain areas are engaged. A related proposed distinction between MVPA and GLM, which frames MVPA as reflecting *sub*-processing that varies during GLM-measured general processing (Jimura and Poldrack, 2012) would suggest that IC's access to multi-voxel patterns would be valuable for mapping sub-processing networks.

Although a comprehensive discussion of the implications of specific findings from this analysis, in terms of our understanding of the visual system, is beyond the scope of this paper, we will make some comments on the types of hypotheses that can be informed by this approach. Firstly, the IC findings are consistent with theories that an object's action representations become automatically activated when its visual or semantic properties are engaged (Chao and Martin, 2000; Johnson-Frey, 2004; Mahon and Caramazza, 2009). A frontal region, the left inferior frontal gyrus, has previously been linked to visual-to-motor transformations (Chao and Martin, 2000) and was informationally connected to several of the seeds here. Equally, the supramarginal gyrus, suggested as a location for representations of object-use skills (Johnson-Frey, 2004), was informationally connected to four of the seeds. Secondly, the distinctions between the IC and FC results for the left fusiform gyrus seed are consistent with a prior fMRI investigation into the organization of object-processing regions (Mahon et al., 2007). Mahon and colleagues (2007) have reported that while the left and right fusiform gyri respond similarly to different object categories in terms of their mean BOLD activation, their underlying neural representations (when measured through repetition suppression) differ. This is supported by the IC and FC differences reported here: the left and right fusiform gyri were *functionally* connected (fitting with Mahon et al.'s mean activation findings) but not *informationally* connected (for the same statistical thresholds), giving support for the left and right fusiform regions differing in their object representations (Mahon et al., 2007). This study is the first to find that object-processing regions are linked together by common fluctuations in multi-voxel patterns for different types of man-made objects.

As a primary analysis method, a key advantage of IC is its ability to examine synchrony within condition-related information that is not accessible from univariate response levels, such as object identity. Dynamically changing cognitive states (such as attention to objects or visual properties) will also differentially affect systems during the time-course of an experiment. For example, time-points marked by greater or reduced attention will likely show increased or decreased pattern discriminability. Regions that process stimuli as part of an interconnected

system will often share these effects. As well as acting as a primary analysis method, IC can be used as a further analysis after an MVPA searchlight procedure, which is often used to identify regions that have condition-relevant information or a relationship to individual differences (e.g., Coutanche, Thompson-Schill, & Schultz, 2011). The brain regions identified in a searchlight analysis will likely decode conditions using a variety of separation principles and forms of relevant information (i.e., the analysis is "opportunistic"; p. 550, Jimura and Poldrack, 2012). For example, man-made objects could be separated by visual appearance in early visual areas, viewpoint-independent identity in later visual areas, associated motor movements in motor areas, and so on. A region's basis for its distributed information will strongly influence which stimuli and time-points are particularly discriminable. The IC approach can help identify different networks of regions, moving beyond one overall MVPA searchlight map. The ability to separate regions based on decoding principles is visible in the IC results for a left inferior occipital gyrus seed in this work. This posterior occipital region showed strong informational connectivity with occipital regions in the opposite hemisphere, but little other contralateral cortex. In contrast, more anterior seeds had more extensive IC. This result was expected, given the basic visual properties that are processed in these early visual areas (Kamitani and Tong, 2005). Once the visual processing stream moves to more anterior brain areas, the processing target moves away from basic visual properties to whole objects, which are processed across different brain regions.

Among other applications, IC can also be used to compare groups by directly contrasting subjects' informational connectivity values, or to examine differences in informational connectivity strengths between tasks. For example, certain networks may show connectivity increases if participants make action-related, compared to visual, judgments of objects. The IC method's general framework can be extended to use classifiers other than the correlation-based approach employed here. Many classifiers, including support vector machines, assign continuous values to the potential conditions for each time-point. These condition-weights determine a classifier's predictions, and incorporate how well the conditions' multi-voxel patterns can be distinguished from each other. By extracting and treating these values in the manner outlined here for

correlations (i.e., correlating a timeseries of classifier condition-weights instead of z-scored correlation coefficients), investigators can draw on the advantages of a range of classification methods.

Although we employed IC using spherical volumes for seeds and searchlights, the method is compatible with seeds and targets that are defined in other ways, such as through anatomical masks or a separate functional localizer. In some cases, it might be desirable to select a seed with a theoretically driven size. For example, an investigator may wish to ensure that the entire visual field of retinotopic V1 is selected as a seed so that the pattern discriminability metric reflects the information available from this entire region. Future investigations that employ both FC and IC can examine seeds that are defined according to a variety of criteria. Here, we selected the univariate-based seeds based on GLM contrasts that were directly comparable to the selection criteria of the multivariate seeds (condition discriminability), but a connectivity seed can be defined in a number of different ways, such as selecting regions with high within-condition variance. The seed and data used in an FC or IC analysis may be influenced by the particular question under investigation. Whereas studies of the object-processing system, for example, may examine a timeseries that fluctuates with different conditions, other targets, such as the influence of attention, may be accessible from seeds that show fluctuating responses within a condition.

One methodological question concerns the length of a timeseries required for robust IC results. The specific data requirements for a given experiment will depend strongly on a number of factors, including the conditions that trigger the data. For IC, experimental paradigms that extensively sample a stimulus space, or that challenge a neural system to varying degrees, will likely produce strongly fluctuating multi-voxel discriminability, potentially increasing the opportunity to sensitively detect relationships between regions. Similarly, an engaging task will likely reduce participant fatigue, and more reliably engage neural representations, thereby producing a more robust measure of discriminability at each time-point. In addition to influencing the quantity of time-points in the IC timeseries, the amount of collected data will influence the

robustness of the training model. This factor is well known to MVPA investigators, and readers are referred to relevant discussions (e.g., Mur et al., 2009; O'Toole et al., 2007) or approaches to improving training data (e.g., Coutanche & Thompson-Schill, 2012) for further information. We note that for the data analyzed here, we observed (from re-running analyses with randomly selected subsets of runs) that the reported informationally connected regions reached significance (as measured with a group mean *t*-value) when the subjects' IC values were calculated from a minimum of between seven and eleven runs (depending on the seed). For this particular set of stimuli and participants, approximately seven to eleven blocks of each condition were therefore sufficient for identifying the brain networks reported above.

Although we found that a prototypical FC analysis was unable to identify the networks found using IC, we acknowledge that a variety of FC analysis measures are available, and others may be more effective. Future work may wish to compare IC results to other FC analysis approaches. Equally, there may be circumstances where investigators wish to track variations in a general process, without influence from sub-process or representational nuances. Analyzing data with FC or IC does not preclude using the other method: in many circumstances, they could be used together and their results compared, as discussed above for the left and right fusiform gyri. A joint approach may lead to a more nuanced understanding of relevant networks. We have created and made available an Informational Connectivity Toolbox online (http://www.informationalconnectivity.org) to aid investigators in applying this technique to their own data.

CHAPTER 4: GENERAL DISCUSSION

"If the instruments played the identical parts, but in a different order, the nature of the symphony changes."

– McIntosh (2000, p. 863)

How are distinct multi-voxel codes integrated across object-processing regions of the human brain? Chapter 2 examined the activation of memory representations for objects' specific features and their convergence to a neural representation of identity. Brain activity was recorded as participants attempted to detect one of four types of fruits and vegetables that differed orthogonally in shape and color. Top-down object knowledge was examined by analyzing data from time periods in which the search target had not yet appeared. Activity patterns of the left anterior temporal lobe (ATL) contained information about a retrieved object's identity, but not its features. In contrast, the object's specific shape and color, but not its identity, were decoded from distinct regions of early visual cortex. In support of theories that hypothesize the existence of a convergence zone, the presence of identity information within the left ATL predicted concurrent decoding of the retrieved object's color in right V4, and shape in bilateral lateral occipital cortex. Chapter 3 outlined a novel analysis method – informational connectivity (IC) – that identifies networks of regions with synchronized fluctuations in multi-voxel information. The IC method can detect connectivity for conditions with multi-voxel pattern signatures, and can evaluate whether particular regions' connectivity is sensitive to the fine-grained perceptual and cognitive distinctions encoded in distributed activity patterns. This new analytical method detected networks of regions that are informationally connected during object processing. Many of the identified inter-regional relationships were not detectable from fluctuating univariate responses, suggesting they are specific to information within multi-voxel patterns, and the high degree of perceptual and cognitive specificity that multi-voxel patterns reflect (Brants, Baeck, Wagemans, & Op de Beeck, 2011; Eger, Ashburner, Haynes, Dolan, & Rees, 2008). These results highlight

that, in addition to multi-voxel patterns playing an important role within individual regions, networks of object-processing areas share fluctuations in multi-voxel information.

Collectively, the results presented here show that networks of regions can be characterized by synchronous pattern-emergence across time. Furthermore, because multi-voxel patterns can reflect cognitive and perceptual targets with a high-degree of specificity, such as differences in shape or object-type, these findings in turn reveal that brain networks are modulated by fine-grained distinctions in cognitions and perceptions. This did not have to be true. If inter-regional connectivity depended solely on *engagement* of a process (e.g., whether shape processing is occurring), rather than the *specific* processing target (e.g., whether a sphere or cylinder is retrieved), multi-voxel information synchrony would not have: a) revealed a link between specific features and object identity in the left ATL, shown in Chapter 2; and b) detected large object networks that are not present in univariate fluctuations, shown in Chapter 3. By examining synchronous decoding, the investigations in this dissertation have answered questions that are not addressable using existing approaches, such as MVPA (which does not typically relate regions across time) and functional connectivity (which is insensitive to the fine-grained cognitive distinctions encoded within the brain's multi-voxel patterns). The novel approaches employed here have allowed the testing of a convergence zone theory (Chapter 2) and the existence of distributed semantic networks that would otherwise have remained inaccessible (Chapter 3).

The results presented here speak to an existing debate regarding how object concepts are represented in the human brain (Binder & Desai, 2011; Martin, 2007; Simmons & Barsalou, 2003). The results of Chapter 2 are consistent with the existence of a convergence zone, particularly findings of: feature-specific decoding in specialized areas of visual cortex (i.e., the substrates of convergence), object-identity decoding in the potential convergence zone (i.e., the results of convergence), and a direct relationship between the zone's retrieval of object identity and synchronous activation of the relevant feature fragments (i.e., the convergence of features to identity). As discussed in Chapter 2, the identification of this color-and-shape convergence zone

offers a new explanation of recent reports of severely affected semantic dementia patients, for whom knowledge of items that are rich in both color and shape was particularly vulnerable to disruption (Hoffman, Jones, & Lambon Ralph, 2012). If temporal lobe atrophy were to reach the color-and-shape convergence zone identified in Chapter 2, we would expect this result. Hoffman and colleagues suggested that atrophy may have extended into feature regions in these patients. Although atrophy of feature regions is possible, disruption of a convergence zone is a more parsimonious explanation for the patients' color and shape deficits, than damage to distinct color and shape systems.

Additionally, the studies' findings emphasize an important role for distributed sensorimotor cortex. First, the discovered link between a retrieved object's identity code and simultaneous feature codes is more consistent with a reciprocal active connection between the integration region and sensorimotor cortices than it is with a single amodal hub. Additionally, the IC analysis of object processing in Chapter 3 revealed large informationally connected networks within frontoparietal action cortex, during a basic (1-back) task that did not require explicit retrieval of action knowledge. These results are consistent with prior findings that action-related processing is engaged when manipulable objects are processed (Chao & Martin, 2000). The IC findings are therefore consistent with sensorimotor cortices playing a central role in the perceptual and conceptual processing of object-concepts, at a high degree of specificity.

Taken together, the results support the existence of one or more integration zones, with an important role for early sensorimotor cortex. One model that is consistent with these findings is the "hub-and-spoke" model of semantic memory (Lambon Ralph, Sage, Jones, & Mayberry, 2010). In the context of the studies presented here, the left ATL region from Chapter 2 has the role of an integrative hub, while lateral occipital cortex, right V4 (Chapter 2), and the dorsal action network (Chapter 3) are spokes.

Theories proposing integration regions differ in a number of respects, such as the number of proposed regions (e.g., Damasio, 1989; Patterson, Nestor, & Rogers, 2007) and whether they extend beyond the temporal lobe (Binder & Desai, 2011). Future research will be

needed to distinguish between these alternatives. The new interpretation of patient deficits presented above is more consistent with multiple convergence zones, as items rich in sound and motion or tactile and action properties (alternative convergences) were unaffected by disruption to the potential convergence zone (Hoffman et al., 2012). Some of the regions identified in the informationally connected networks of Chapter 3 may act as convergence zones. For example, multi-voxel patterns in frontoparietal action cortex may have dependences on synchronous properties, such as a particular grip and type of manipulation.

The work here suggests several avenues of future research. The findings and analysis methods of Chapter 2 can be applied to other visual features; for example motion, to ask whether other feature combinations (e.g., shape and motion) would converge on the same or a different cortical location as shape and color. Relatedly, does the addition of a third feature (e.g., shape, color and motion) prompt convergence in the same region, or at a more anterior position in the temporal lobe? Such a result might be predicted from accounts hypothesizing a hierarchy of convergence zones (Meyer & Damasio, 2009).

Future work may also further examine the implications of two regions having greater multi-voxel pattern synchrony than univariate synchrony, and *vice versa*. The direction of this difference might indicate the specificity at which a brain network is processing a perceptual or cognitive target, in the same manner that multi-voxel and univariate analyses reflect this (Brants et al., 2011; Coutanche, 2013). Systematically manipulating the specificity of object distinctions, and measuring changes in IC between object regions, could help illuminate this further.

Our knowledge of an object includes its features, and how they are bound together into a coherent identity. The research presented here has found that perceptual and conceptual object processing is associated with the synchronous emergence of multi-voxel patterns across a network of regions. A region in the left ATL was found to have a number of properties that are predicted by a theory of a neural convergence zone, including a close link between information about object identity and the synchronous activation of an object's specific features in specialized visual areas of cortex. Further, the findings reported here suggested that when we perceive

objects, large networks of cortex show shared fluctuations of information within distributed activity patterns. This work reveals that connectivity between object-processing regions is affected by variations across features and objects at a high degree of specificity.

BIBLIOGRAPHY

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*(11), 527–536.

Biswal, B., Zerrin Yetkin, F., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, *34*(4), 537–541.

Bouvier, S. E., & Engel, S. A. (2006). Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral Cortex*, *16*(2), 183–191.

Bramão, I., Faísca, L., Forkstam, C., Reis, A., & Petersson, K. M. (2010). Cortical brain regions associated with color processing: an FMRI study. *The open neuroimaging journal*, *4*, 164–173.

Brants, M., Baeck, A., Wagemans, J., & Op de Beeck, H. P. (2011). Multiple scales of organization for object selectivity in ventral visual cortex. *NeuroImage*, *56*(3), 1372–1381.

Carlson, T., Hogendoorn, H., Hubert Fonteijn, & Verstraten, F. A. J. (2011). Spatial coding and invariance in object-selective cortex. *Cortex*, *47*(1), 14–22.

Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., & Milner, A. D. (2010). Separate processing of texture and form in the ventral stream: evidence from FMRI and visual agnosia. *Cerebral Cortex, 20*(2), 433–446.

Chai, B., Walther, D. B., Beck, D. M., and Fei-Fei, L. (2009). Exploring functional connectivity of the human brain using multivariate information analysis. *Advances in neural information processing systems*, *22*, 270–278.

Chao, L. L., and Martin, A. (2000). Representation of Manipulable Man-Made Objects in the Dorsal Stream. *NeuroImage*, *12*(4), 478–484.

Chiu, Y.-C., Esterman, M. S., Gmeindl, L., and Yantis, S. (2012). Tracking cognitive fluctuations with multivoxel pattern time course (MVPTC) analysis. *Neuropsychologia*, *50*(4), 479–486.

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science*, *248*(4962), 1556–1559.

Coutanche, M. N. (2013). Distinguishing multi-voxel patterns and mean activation: Why, how, and what does it tell us? *Cognitive, affective & behavioral neuroscience*.

Coutanche, M.N., and Thompson-Schill, S.L. (2011, April). Informational Connectivity: a novel fMRI analysis method for identifying brain areas that share distributed encoding principles. Poster presented at the Cognitive Neuroscience Society Annual Meeting 2011, San Francisco, CA.

Coutanche, M. N., and Thompson-Schill, S. L. (2012). The advantage of brief fMRI acquisition runs for multi-voxel pattern detection across runs. *NeuroImage*, *61*(4), 1113–1119.

Coutanche, M. N., Thompson-Schill, S. L., and Schultz, R. T. (2011). Multi-voxel pattern analysis of fMRI data predicts clinical symptom severity. *NeuroImage*, *57*(1), 113–123.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173.

Damasio, A. R. (1989). The Brain Binds Entities and Events by Multiregional Activation from Convergence Zones. *Neural Computation*, *1*(1), 123–132.

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on “sensory-specific” brain regions, neural responses, and judgments. *Neuron*, *57*(1), 11–23.

Eger, E., Ashburner, J., Haynes, J.-D., Dolan, R. J., & Rees, G. (2008). fMRI activity patterns in human LOC carry information about object exemplars within category. *Journal of Cognitive Neuroscience*, *20*(2), 356–370.

Esterman, M., Chiu, Y.-C., Tamber-Rosenau, B. J., and Yantis, S. (2009). Decoding cognitive control in human parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(42), 17974–17979.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., … Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*(3), 341–355.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(27), 9673–9678.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., and Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*, *6*(3), 218–229.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., and Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping*, *2*(4), 189–210.

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*(1), 187–203.

Hanson, S. J., Matsuka, T., and Haxby, J. V. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a “face” area? *NeuroImage*, *23*(1), 156–166.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425–2430.

Haynes, J.-D., and Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews. Neuroscience*, *7*(7), 523–534.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic Dementia Progressive Fluent Aphasia with Temporal Lobe Atrophy. *Brain*, *115*(6), 1783–1806.

Hoffman, P., Jones, R. W., & Lambon Ralph, M. A. (2012). The degraded concept representation system in semantic dementia: damage to pan-modal hub, then visual spoke. *Brain*, *135*(12), 3770–3780.

Jimura, K., and Poldrack, R. A. (2012). Analyses of regional-average activation and multivoxel pattern information tell complementary stories. *Neuropsychologia*, *50*(4), 544–552.

Johnson-Frey, S. H. (2004). The neural bases of complex tool use in humans. *Trends in Cognitive Sciences*, *8*(2), 71–78.

Kamitani, Y., and Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*(5), 679–685.

Kamitani, Y., and Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Current biology: CB*, *16*(11), 1096–1102.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, *48*(7), 805–825.

Kosslyn, S. M., Thompson, W. L., Costantini-Ferrando, M. F., Alpert, N. M., & Spiegel, D. (2000). Hypnotic visual illusion alters color processing in the brain. *The American journal of psychiatry*, *157*(8), 1279–1284.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3863–3868.

Lambon Ralph, M. A., & Patterson, K. (2008). Generalization and Differentiation in Semantic Memory. *Annals of the New York Academy of Sciences*, *1124*(1), 61–76.

Lambon Ralph, M. A., Sage, K., Jones, R. W., & Mayberry, E. J. (2010). Coherent concepts are computed in the anterior temporal lobes. *Proceedings of the National Academy of Sciences*, *107*(6), 2717–2722.

Lee, Y.-S., Janata, P., Frost, C., Hanke, M., and Granger, R. (2011). Investigation of melodic contour processing in the brain using multivariate pattern-based fMRI. *NeuroImage*, *57*(1), 293–300.

Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2012). Disentangling visual imagery and perception of real-world objects. *NeuroImage*, *59*(4), 4064–4073.

Lizier, J. T., Heinzle, J., Horstmann, A., Haynes, J.-D., and Prokopenko, M. (2011). Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fMRI connectivity. *Journal of computational neuroscience*, *30*(1), 85–107.

Mahon, B. Z., and Caramazza, A. (2009). Concepts and categories: a cognitive neuropsychological perspective. *Annual review of psychology*, *60*, 27–51.

Mahon, B. Z., Milleville, S. C., Negri, G. A. L., Rumiati, R. I., Caramazza, A., and Martin, A. (2007). Action-related properties shape object representations in the ventral stream. *Neuron*, *55*(3), 507–520.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25–45.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Current Opinion in Neurobiology*, *11*(2), 194–201.

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in neurosciences*, *29*(6), 317–322.

McIntosh, A. R. (2000). Towards a network theory of cognition. *Neural networks*, *13*(8-9), 861–870.

McKeefry, D. J., & Zeki, S. (1997). The position and topography of the human colour centre as revealed by functional magnetic resonance imaging. *Brain*, *120 (12)*, 2229–2242.

McNorgan, C. (2012). A meta-analytic review of multisensory imagery identifies the neural correlates of modality-specific and modality-general imagery. *Frontiers in Human Neuroscience*, *6*, 285.

Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in neurosciences*, *32*(7), 376–382.

Mitchell, T., Hutchinson, R., Niculescu, R., Pereira, F., Wang, X., Just, M., & Newman, S. (2004). Learning to Decode Cognitive States from Brain Images. *Machine Learning*, *57*(1), 145–175.

Morita, T., Kochiyama, T., Okada, T., Yonekura, Y., Matsumura, M., & Sadato, N. (2004). The neural substrates of conscious color perception demonstrated using fMRI. *NeuroImage*, *21*(4), 1665–1673.

Mur, M., Bandettini, P. A., and Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI—an introductory guide. *Social Cognitive and Affective Neuroscience*, *4*(1), 101 –109.

Ng, A. Y., & Jordan, M. I. (2002). On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in neural information processing systems*, *2*(14), 841–848.

Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430.

O'Toole, A. J., Jiang, F., Abdi, H., and Haxby, J. V. (2005). Partially distributed representations of objects and faces in ventral temporal cortex. *Journal of Cognitive Neuroscience*, *17*(4), 580–590.

O'Toole, A. J., Jiang, F., Abdi, H., Pénard, N., Dunlop, J. P., and Parent, M. A. (2007). Theoretical, Statistical, and Practical Perspectives on Pattern-based Classification Approaches to the Analysis of Functional Neuroimaging Data. *Journal of Cognitive Neuroscience*, *19*(11), 1735–1752.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976–987.

Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(29), 12125–12130.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *NeuroImage*, *45*(1 Suppl), S199–209.

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010). Category-specific versus category-general semantic impairment induced by transcranial magnetic stimulation. *Current biology*, *20*(10), 964–968.

Quamme, J. R., Weiss, D. J., and Norman, K. A. (2010). Listening for recollection: a multi-voxel pattern analysis of recognition memory retrieval strategies. *Frontiers in Human Neuroscience*, *4*.

Raizada, R. D. S., and Connolly, A. C. (2012). What makes different people's representations alike: neural similarity space solves the problem of across-subject fMRI decoding. *Journal of cognitive neuroscience*, *24*(4), 868–877.

Reddy, L., Tsuchiya, N., & Serre, T. (2010). Reading the mind's eye: decoding category information during mental imagery. *NeuroImage*, *50*(2), 818–825.

Roe, A. W., Chelazzi, L., Connor, C. E., Conway, B. R., Fujita, I., Gallant, J. L., … Vanduffel, W. (2012). Toward a unified theory of visual area V4. *Neuron*, *74*(1), 12–29.

Rogers, T. T., Patterson, K., & Graham, K. (2007). Colour knowledge in semantic dementia: It is not all black and white. *Neuropsychologia*, *45*(14), 3285–3298.

Seghier, M. L. (2013). The Angular Gyrus Multiple Functions and Multiple Subdivisions. *The Neuroscientist*, *19*(1), 43–61.

Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, *20*(3), 451–486.

Singh, V., Miyapuram, K. P., & Bapi, R. S. (2005). Detection of Cognitive States from fMRI data using Machine Learning Techniques. *In: Proceedings of Twentieth International Conference on Artificial Intelligence. (2007) 587–592*.

Spiridon, M., & Kanwisher, N. (2002). How distributed is visual category information in human occipito-temporal cortex? An fMRI study. *Neuron*, *35*(6), 1157–1165.

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *The Journal of Neuroscience, 29*(5), 1565–1572.

Stokes, M., Thompson, R., Nobre, A. C., & Duncan, J. (2009). Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proceedings of the National Academy of Sciences*, *106*(46), 19569–19574.

Thirion, B., Pinel, P., Mériaux, S., Roche, A., Dehaene, S., and Poline, J.-B. (2007). Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. *NeuroImage*, *35*(1), 105–120.

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring “how” from “where.” *Neuropsychologia*, *41*(3), 280–292.

Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual review of psychology*, *63*, 483–509.

Walther, D. B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2009). Natural scene categories revealed in distributed patterns of activity in the human brain. *The Journal of neuroscience*, *29*(34), 10573–10581.

Welchew, D. E., Ashwin, C., Berkouk, K., Salvador, R., Suckling, J., Baron-Cohen, S., and Bullmore, E. (2005). Functional disconnectivity of the medial temporal lobe in Asperger’s syndrome. *Biological psychiatry*, *57*(9), 991–998.

Zhao, M., & Lyengar, S. (2010). Nonconvergence in Logistic and Poisson Models for Neural Spiking. *Neural Computation*, *22*(5), 1231–1244.