PERSPECTIVE

What insights can fMRI offer into the structure and function of mid-tier visual areas?

CHERYL A. OLMAN

Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455

Abstract

Inferring neural responses from functional magnetic resonance imaging (fMRI) data is challenging. Even if we take advantage of high-field systems to acquire data with submillimeter resolution, we are still acquiring data in which a single datum summarizes the responses of tens of thousands of neurons. Excitation and inhibition, spikes and subthreshold membrane potential modulations, local and long-range computations, and tuned and nonselective responses are mixed together in one signal. With *a priori* knowledge of the underlying neural population responses, careful experiment design allows us to manipulate the experiment or task design so that subpopulations are selectively modulated, and our experiments can reveal those tuning functions. However, because we want to be able to use fMRI to discover new kinds of tuning functions and selectivity, we cannot limit ourselves to experiments in which we already know what we are looking for. Broadly speaking, analyses that rely on classification of responses that are distributed across the local neural population [multi-voxel pattern analyses (MVPA)] offer the ability to discover new kinds of information representation and selectivities in neural subpopulations. There is, however, no way to determine how the information discovered with MVPA or other analyses is related to the underlying neuronal tuning functions. Therefore, we must continue to rely on behavioral, computational, and animal models to develop theories of information representation in mid-tier visual cortical areas. Once encoding models exist, fMRI can be powerful for testing these *a priori* models of information representation. As an aide in developing these models, an important contribution that fMRI can make to our understanding of mid-tier visual areas is derived from connectivity analyses and experiments that study information sharing between visual areas. This ability to quantify localized population average responses throughout the brain is the strength we can best leverage to discover new properties of local and long-range neural networks.

Keywords: fMRI, Connectivity, Visual hierarchy

Introduction

Our understanding of the human visual system—from the delineation of regions containing different spatial maps of the visual world to our understanding of just how distributed information processing is—would be far behind current levels without functional magnetic resonance imaging (fMRI). In spite of the fact that fMRI provides only an indirect measure of neural responses (hemodynamic measures like blood flow and oxygenation are proxies for the neural signals we want to study), fMRI experiments have revealed much about the mesoscopic (intra-area organizations such as retinotopy) and macroscopic (delineation of brain regions) structure.

One of the most exciting contributions of fMRI to our understanding of primary visual cortex (V1) is our awareness that responses in V1 are modulated by information encoded outside of V1. That is, to say: V1 fMRI responses are strongly modulated by global scene or behavioral context (visual information) that cannot possibly be explained by the local neuron response properties (i.e., not encoded locally). A caveat about what is meant by "strong" modulation is appropriate here. The signal-to-noise ratio (SNR) of fMRI studies, in general, is not strong—we are generally looking for ~1% signal changes in a signal that has noise processes with standard deviations of 1% or more. So our SNR is about 1:1. A strong modulation, due to attention or scene structure, might be a doubling of the signal. However, given the fundamentally low SNR of fMRI, we still require an average of many trials to discover these strong modulations.

There is some controversy over whether this strong modulation of the fMRI signal actually represents modulation of neural responses. It is, after all, a much larger modulation than is typically seen in the firing rates of neurons (Yoshor et al., 2007). Perhaps, it reflects instead some alteration of local energetics or the glial signal (Iadecola & Nedergaard, 2007), rather than neural firing rates? For an excellent balanced discussion of this, see the discussion of

Address correspondence to: Cheryl A. Olman, Associate Professor, Department of Psychology, University of Minnesota, N218 Elliott Hall, 75 East River Road, Minneapolis, MN 55455. E-mail: caolman@umn.edu

1

Maier et al (2008). However, a reasonable view of these reports of V1 modulation by scene disambiguation (Hegde & Kersten, 2010), object information (Williams et al., 2008), or scene coherence (Mannion et al., 2013), for example, is that interactions with other cortical [or subcortical (Sherman, 2007)] regions are modifying, if not firing rates, then at least some aspect the way in which different V1-encoded visual features are represented in V1. This very integrative view of V1 as a node in a distributed network makes particular sense when one considers that, perhaps, 10% of the inputs to V1 come from "below" V1 (i.e., subcortical sources) (Logothetis, 2008), while the vast majority come from other cortical regions in the brain that rank "higher" in the visual network.

It is amazing that fMRI can tell us so much about what is going on in the human visual system, yet at the same time reveal so little about what we want to know: how does the human visual system encode the visual information in a scene? Visual information is used here as a very broad term—anything an observer can learn from a scene. The term encode, however, is intended to have a narrow interpretation. Information encoding is used here to refer specifically to tuning functions, or response selectivity, of neurons in a localized neural population. In V1, information that meets the definition of being encoded in V1 would be the orientation of a luminance contrast edge that subtends a degree or two of visual angle, or the color contrast of that edge, or perhaps the direction it is drifting.

The assumption for the rest of this article is that our basic research goals are to understand both (1) how a localized neural population encodes information and (2) how that information encoding is modulated by information from other nodes in the visual hierarchy. Increasing sophistication in analysis techniques and experiments makes us excited about accessing large-scale information contained in local neural activity. However, the "Caution" section will identify some concerns about the limitations of quantitative inference of neural response properties from fMRI data. "The particular challenges encountered in hV4" section will take hV4 as an opportunity to consider the strong limitations we have in using fMRI to discover— *de novo*—how mid-tier visual areas represent visual information. Finally, the "So what is fMRI good for?" section will argue that fMRI is particularly well suited for addressing the second of the two problems above—information shared between visual areas.

Caution

Is the term "neural activity" useful?

Not long ago, there was much discussion in the literature about energy budgets and neuro-hemodynamic coupling (e.g., Lennie, 2003). A driving concern was that the fMRI response might be driven by some aspect of physiology or metabolism that was not correlated with the firing rates of the neurons representing the information we want to discover. However, a large body of literature supports the idea that fMRI does indeed reflect neural responses, rather than temporally or spatially dissociated characteristics like hemodynamic auto-regulation or glial energy consumption [although the jury is still out regarding the question of whether glia are part of the computational network (Schummers et al., 2008)]. In fact, the discovery of pericytes (Hall et al., 2014) and the corresponding ability of the vasculature to regulate on a very fine spatial scale, as well as the proof-of-principle visualization of ocular dominance columns (Cheng et al., 2001; Yacoub et al., 2007) and even orientation columns (Yacoub et al., 2008), indicates that the fMRI signal provides a very spatially accurate map of neural activity with accuracy that can be better than 1 mm if we use specialized imaging

techniques. With standard gradient-echo echo-planar imaging (EPI), our resolution is in the 2–5 mm range (Olman & Yacoub, 2011).

The danger of the term "neural activity" is that it can be used to imply homogeneity in a heterogeneous and inadequately sampled neural population. Neuron density is on the order of 10,000 neurons/mm^3 throughout cortex, as high as 40,000 in V1, less in other areas. So even with submillimeter resolution (e.g., 0.8 mm, isotropic, or about 0.5 mm^3), a voxel (fMRI resolution element) contains 5000–20,000 neurons. Blurring is of course present in the data, so a single fMRI datum reflects at best the temporally blurred, spatially aggregated responses of 10,000–50,000 neurons. There is redundancy in the population, of course, but in that population— for V1 at least—a neuroscience graduate student could easily name about 30 different types of neurons (and know that she/he is forgetting another 30 types): several different kinds of inhibitory interneurons in every layer, excitatory interneurons in the input layers, pyramidal cells in deep and superficial layers, and a complete collection of each of these types of cells for every feature-tuned column. In V1, the most obvious feature is orientation; those columns appear to be about 200 microns in diameter in humans, so there are at least four in the smallest voxel, and typically many more. Many different features are mapped across the cortical surface (Swindale, 1992), and for each, there is a small army of excitatory and inhibitory neurons with different connection patterns and roles in shaping the population's selectivity to that feature. When we see a signal increase or decrease in a fMRI voxel, how do we know which neurons have changed their firing rate?

In defense of the term "neural activity" stands the fact that its common interpretation is "average firing rate in the local population," and this kind of measurement does have inherent value. Other techniques for measuring *in vivo* neuronal responses tend to biased toward large neurons: electroencephalography (EEG) and magnetoencephalography (MEG) will detect fields generated by ionic currents driven by large pyramidal neurons at particular orientations (Darvas et al, 2004), and even invasive electrophysiological recordings are most likely to sample larger neurons (Olshausen & Field, 2004). Hemodynamic responses, on the other hand, will reflect activity in a broader cross-section of the neuronal population, which may be why fMRI reflects modulatory signals with surprising strength (Maier et al., 2008).

Further support for the value of population average mapping comes from discoveries made with optical imaging [e.g., (Hubener et al., 1997)]. Optical imaging is also a hemodynamic measurement that pools responses from neurons with many different encoding functions, yet key aspects of cortical organization were discovered with optical imaging. While the fact that neurons with similar response properties cluster in columns that span the cortical depth was discovered with electrophysiology (Mountcastle, 1997), but elucidation of the mesoscopic organization of these columns, into orientation pinwheels, for example (Bonhoeffer & Grinvald, 1993), required a technique that could uniformly sample the population average over a large spatial scale.

Still, it is not an exaggeration to say that the term "neural activity" is about as quantitative as the term "public opinion." Because of the spatial scale and diversity of the neural population, inferring "neural activity" from the response in a fMRI voxel is in many ways like asking the mayor of a small town about a politically charged issue and thinking you have adequately polled the citizens. From one perspective, this is a perfectly valid approach: for a given region of cortex, we cannot possibly sample every response, so we might be happy with an estimate of the average response (or a winner-take-all estimate of the majority vote). However, from the

perspective of understanding information representations in the brain, this is exactly the wrong approach.

A recent article by Brouwer and Heeger (2011) studying cross-orientation suppression showed us one way to circumvent this difficulty. One beauty of cross-orientation suppression is that it calibrates the population response to simultaneously encode total contrast as well as contrast at each of the contributing orientations (McDonald et al., 2012). This, however, means that the population average response (the fMRI signal) will be the same whether a high-contrast vertical grating, a high-contrast horizontal grating, or a high-contrast mixture of vertical and horizontal is presented to the subject. That is, average population response is the same in spite of very different response profiles in the sub-populations of neurons. Brouwer and Heeger took advantage of the fact that—for whatever reason—each fMRI voxel in V1 has a slight orientation bias (Kamitani & Tong, 2005; Sun et al., 2013). Using these biased voxels as indicators of the underlying neural subpopulations, they were able to estimate subpopulation responses that matched behavior and single-unit electrophysiology. His study shows how crucial (and feasible) it is for us to continue to develop tools what will reveal tuning functions in subpopulations of neurons, rather than use fMRI as a measure of locally averaged neural activity. This article also exemplifies how strongly fMRI studies must rely on modeling, behavioral measures, and single-unit recordings; fMRI data cannot be interpreted in isolation.

What does it mean to discover information in a decoding analysis?

While the previous analysis showed that subpopulation responses were indeed accessible behind a bland population response, there is also an important mirror image caveat: the information revealed by a decoding analysis might be reflected or available, but not actually encoded, in a given cortical region. Trivially, we can consider the example of decoding in V1. Many articles have shown that multi-voxel pattern analyses (MVPA) can reveal sensitivity to a broad range of global scene information in the V1 population response. None of these authors, however, is arguing that this high-level information is encoded in V1. Instead, we are fascinated to discover the strength with which it is reflected (presumably *via* long-range connections) in the V1 fMRI signal. So this first caveat about interpreting the results of decoding analyses is that the information reflected in a given cortical region was quite likely encoded in a different cortical area.

It is also easy to envision situations in which information is encoded locally but not along the dimensions detected by the classification analysis. Davis et al. (2014) provides a compelling example. The authors envision an experiment studying animal perception and a neural population that contains two subpopulations: one that responds to the size of an animal and one that responds to the aggressiveness ("predacity") of the animal. Taken together, size and predacity predict scariness. A cow is large but not aggressive; a wolverine is relatively small but worth going out of your way to avoid. A linear classifier seeking to decode this neural population's response to the pictures of these animals (in a subject who has knowledge of their behavior) would discover that this region of cortex encodes scariness. However, that is not actually the information being encoded. Size and predacity are the features encoded locally; scariness is a derivative concept. Discovering the representation of scariness in this neural population does not help us understand the local neural tuning functions or information encoding.

The conclusion of this argument is not that we should stop using MVPA for studying visual information representations in the brain. How else are we going to discover representations that have not yet occurred to us, except by using the technique we have that is most sensitive to information contained in distributed neural populations and robust to between-subject sources of noise (Davis et al., 2014)? The key concern is that we should remain cautions about taking the leap from "information is available here" to "information is computed here." Additionally, as discussed below: given the number of neural responses confounded in very voxel, we cannot claim that fMRI can access the actual neural tuning functions without confirmation by behavior, computational, and where appropriate, animal models.

Encoding models

The need to rely on multiple sources of information to interpret fMRI (or any other neuronal or imaging) data is reflected in a recent growth of interest in encoding models. Encoding models represent an approach to data analysis similar to that previously discussed in Brouwer and Heeger (2011): *a priori* knowledge of neural tuning functions is used to build parameterized models of neural subpopulation responses, and by fitting these models to fMRI data, subpopulation response parameters are estimated in order to explain the behavior of the fMRI data. Recent prime examples have been provided for the encoding of natural images (Naselaris et al., 2009; Nishimoto et al., 2011) and semantic information (Cukur et al., 2013) throughout the visual hierarchy. Population receptive field estimation techniques (Dumoulin & Wandell, 2008) also have their basis in encoding models and are showing good promise for revealing the mesoscopic organization of mid-tier visual areas, where feature selectivity is complex and checkerboards have little utility.

Encoding models have their own set of limitations, of course. A significant one is the number of parameters that need to be used to model the diversity of tuning functions that will contribute to each voxel's or region's fMRI response. The V1 model used in the work (Kay et al., 2008), for example, needed more than 1000 parameters to capture all possible spatial positions, spatial frequencies, and orientations. It takes a lot of data to constrain that many parameters!

Another limitation of encoding models is that, in order to be fit to the fMRI signal, they need to account (or at least allow) for modulation by long-range connections. Even with a very simple encoding model (orientation-dependent surround suppression in V1), my laboratory has struggled for some time now with a finding we published in 2010 (Schumacher & Olman, 2010): the (beautifully) localized fMRI response decreased, rather than increased, as we increased the contrast of Gabor elements when they were flanked by parallel elements. Five experiments later, we are on the verge of publishing "the rest of the story": the V1 fMRI signal is composed of an attention-modulated mixture of first-order and second-order contrast representations. The crux move is realizing that in even very simple tasks, perception and behavioral state shape even very low-level fMRI responses.

At this point, it is worth stopping to consider that this fundamental limitation—the mixture of signals—is not unique to fMRI. All measurements of neural activity in the intact brain of a behaving animal suffer from the same limitation: when a neuron, or a signal reflecting neuronal responses, shows selectivity to a particular bit of information, it cannot be unambiguously determined whether it is local or long-range neural networks that are shaping

that selectivity. In EEG, and to a lesser degree in MEG, the location of the sensors with respect to the neuronal populations keeps this issue at the forefront. In a sense, this makes EEG and MEG the most "honest" techniques: it is always obvious that the measured signal has multiple potential sources. Functional MRI data are so beautifully localized to gray matter, with surprising precision for a noninvasive imaging modality, that it is tempting to believe that we know exactly where that signal is coming from. However, existing literature—some of which has been discussed above—provides ample evidence that this beautifully localized fMRI signal is often modulated by quite remote neuronal populations. So, in this sense, the fMRI signal is not well-localized at all. What is most surprising, perhaps, is that single-unit electrophysiology recordings suffer from a signal-localization problem that is quite similar to the problem suffered by fMRI. It is not possible to sample all neurons in the brain simultaneously, so when a firing rate or spiking threshold is modulated, it is quite difficult to infer the source of that signal modulation from electrophysiological recordings. The "So what is fMRI good for?" section will discuss how the whole-brain coverage of fMRI, particularly when used in the context of connectivity analyses, offers an exciting opportunity to address this "source localization" problem, even though fMRI is frustratingly blind to the specific neurons giving rise to the signal we measure.

The particular challenges encountered in hV4

Thus far, in considering the limitations of fMRI for elucidating the mechanisms of neural information encoding, the key ideas have been (1) lack of selectivity to neuronal subpopulations and (2) the fact that information reflected in any given neuronal population can come from either local computations or modulation by long-range connections to remote neuronal populations. *A priori* encoding models are required to interpret fMRI data. However, if quantitative fits to encoding models are difficult in V1, they are that much more difficult in mid-tier and higher visual areas because of the increased complexity of the feature space encoded by the neurons. A world-view-changing summary of this problem is available in DiCarlo, Zoccolan, and Rust 2012. That article clearly articulates the problem of approaching mid-tier visual areas in the same way that we have approached V1, with normalized linear/nonlinear (NLN) models: it is difficult to conceive of acquiring enough neuroscience data to constrain parameters in the "deep stack" of NLN models required to approach mid-tier regions such as hV4 from the bottom up. This is particularly challenging as we can rely less and less on animal models as we move into higher visual areas.

Not only is the modeling more complicated as we move beyond V1 to mid-tier visual regions, but demands on imaging resolution (not only voxel size but also the accuracy of registration to reference anatomy) increase as we look at smaller visual regions like the subunits of lateral occipital complex (LOC) (Sayres & Grill-Spector, 2008). In V1, it is straightforward to identify representations of separate regions of the visual field and study how these interact (de Wit et al., 2012; Kok & de Lange, 2014); standard 3–5 mm resolution is sufficient. In mid-tier visual areas, however, the entire retinotopic map might span only a few centimeters, and separation of signals from different regions of the visual field requires millimeter-precision.

Imaging is further complicated for mid-tier regions on the ventral surface of the brain because the image quality is degraded by motion artifacts (subject motion, as well as partial-volume effects as the brain moves in response to pulse and respiration) and magnetic field inhomogeneities caused by tissue interfaces. V1, where we have so much experience with high-resolution imaging and model fitting, has a rather privileged location where anatomy is relatively homogenous and image quality is optimal. V1 is further privileged by the fact that it has particularly strong vascularization (Zheng et al., 1991) and potentially stronger hemodynamic responses than other regions in the visual system.

As a good example of the challenge of getting high-quality fMRI data in ventral cortex, we can pick human V4 (hV4). The difficulty of arriving at a consensus on whether cortex adjacent and inferior to dorsal V3 contains a quarterfield representation of the upper visual field (Tootell & Hadjikhani, 2001) or a complete hemifield representation (Witthoft et al., 2014) is a good example of the practical difficulty of studying ventral visual regions. Given the spatial precision of fMRI, it is astonishing that something as big as a retinotopic map in an evidently retinotopic visual area would be difficult to agree on, but the ambiguity has been evident in my laboratory's data, as well. For most subjects, we see a clear hemifield, albeit sometimes in only one hemisphere and not the other. For some subjects, we cannot definitively identify a full hemifield representation in either hemisphere, even though we are convinced it should be there.

It is likely that methodological limitations—specifically distortions and signal loss (Olman et al., 2009) caused by the transverse sinus (Winawer et al., 2010)—play a significant role in creating ambiguity about ventral visual representations. When we do retinotopic mapping with Spin Echo EPI data instead of the standard Gradient Echo (T_2*-weighted) EPI data (Olman et al., 2010), we see more reliable signal from ventral visual regions. Spin Echo EPI is T_2-weighted: insensitive to field perturbations that occur over spatial scales larger than ~100 microns and, therefore, subject to distortion but not signal dropout from the large sinus. T_2-weighted fMRI is not practical at 3 T but produces beautiful data at 7 T, albeit with a contrast-to-noise ratio roughly half that of T_2*-weighted fMRI (Olman et al., 2010). T_2*-weighted EPI data from V4 may not be robust enough—barring the use of high-resolution to combat drop-out, and high parallel imaging reduction factors to combat distortion—to provide a firm enough footing from which to embark on a careful study of neural response properties using fMRI. For studies that want a uniform sampling of ventral visual regions, or the ability to quantify distinct information representations in distinct regions of the visual field, Spin Echo EPI techniques at 7 T are likely worth the trouble.

So what is fMRI good for?

The above arguments are aimed at this central question: can fMRI be used to discover neural information encoding in a given region of cortex? To recap the main arguments thus far: it is prohibitively complicated—or at least it will take a very long time—to work from the bottom up to build an encoding model that represents a first-principled understanding of how neurons extract the information encoded in a mid-tier visual region. Can we take advantage of high-resolution, noninvasive neuroimaging to shortcut the process and discover information representations in mid-tier visual areas? The simplest answer to this question is "no." Even with an encoding model, or at least an *a priori* expectation for what information is encoded locally, our experiences in V1 have taught us that the localized fMRI signal is a mixture of locally encoded information and modulatory signals originating elsewhere in the brain. In the absence of independent measures of local information encoding, any information we discover—whether through a decoding analysis or in the mean signal of a traditional linear regression analysis—cannot be unambiguously assigned as originating locally, nor can it

be clearly mapped to underlying neural tuning functions (as in the scary animal example).

So, if our ideas about neural information encoding in mid-tier visual areas need to come from elsewhere, what is fMRI good for? The remaining discussion focuses on the fact that fMRI is particularly useful for two things: (1) validating or excluding models and (2) identifying information that is held in common throughout multiple visual areas. On these two fronts, fMRI offers unparalleled opportunities.

Model validation

While fMRI might not be able to provide direct measures of local information encoding, fMRI can be very useful for validating models. Recently, Freeman et al. (2013) took this approach in V2. Single-unit electrophysiological recordings indicated that V2 cells in nonhuman primates differentiated between naturalistic images that differed in higher-order pixel luminance statistics (higher than second order). Functional MRI data verified that this was a general property in V2, but not V1, by determining that the signal throughout V2 depended on higher order statistics; the study further confirmed that the V2 sensitivity to higher-order statistics matched observer's behavioral thresholds. The experiment might have been done in reverse order, or the electrophysiological data might have been excluded, since there are other reasons to believe that high-order statistical sensitivity originates in V2. Regardless of the details, the reason this approach was successful in convincing us that V2 encodes a particular type of information is that a direct comparison was made between behavior and a clearly defined type of visual information, which was shown to modulate V2 but not V1.

Connectivity analyses for determining shared information

There are two things we want to know: (1) how information is encoded locally and (2) what information that is shared between visual areas. We cannot use fMRI to address the first point without neurophysiologically plausible (quantitative and falsifiable) encoding models, and these remain challenging to construct. The second problem, however, is perfectly tailored to the strengths of fMRI, since well-localized signals that are simultaneously acquired throughout the brain are readily available.

Visually responsive regions of the brain can profitably be viewed as a network rather than a hierarchy of areas. While some object recognition tasks may be accomplished with a single pass through the system (Epshtein et al., 2008), many visual tasks require iterative computations between different levels in this hierarchy. At about 10 ms per level (Nowak & Bullier, 1997), information representations have been shared between visual areas several times before we recognize an object or saccade to our next target. An alternative to focusing on the local neural networks that encode different kinds of visual information in different visual areas is to focus on discovering what information is shared between visual areas.

Limiting the discussion for now to the retinotopically organized areas that can be revealed with a simple protocol (perhaps V1/2/3/4/3AB/7/IPS1+/pLOC, although there are certainly other candidates), we are considering a network with nodes separated by a few centimeters and distributed across perhaps a 12 cm × 12 cm × 6 cm volume. At the 1–5 mm resolution afforded by standard fMRI techniques, we can resolve representations of different locations in the visual field (retinotopy) within each of these regions, and we can study response dependencies between these regions and subregions. The temporal dynamics of this information sharing are evidently inaccessible except by creative experiment design. However, the spatial attributes are perfectly matched to fMRI.

Results from connectivity analyses are, of course, as open to misinterpretation as results from decoding analyses: all we are doing when we calculate task-dependent correlations between voxels (Baldassano et al., 2012) or regions (McLaren et al., 2012; O'Reilly et al., 2012) is discovering what information is shared between regions. We remain ignorant of how that information is shared or what neural mechanisms are used to calculate it. However, with connectivity analyses, particularly connective field modeling (Haak et al., 2012) and task-dependent connectivity (Friston et al., 1997), combined with intelligent (parameterized) experiment design, we can efficiently explore theories about the kind of information passed between nodes in this hierarchical network. This, in turn, will let us form hypotheses about the inputs and outputs of different nodes in the network and design experiments to quantitatively model selectivity of different regions to information identified through connectivity analyses.

Lately we have invested significant effort in understanding whether depth-resolved measurements can be used to extend inter-regional connectivity analyses by isolating subpopulations at different cortical depths with different input/output relationships (Callaway, 2004 ; Markov et al., 2014). Our preliminary studies are encouraging (Olman et al., 2012), but so far, none can be unambiguously connected to quantitative encoding models. We therefore do not know whether depth-resolved fMRI will provide unique information about how information is shared between visual areas. Even if it does, we will still have dramatically fewer data points than neuron types, so fMRI will never be able to stand on its own as an estimate of the neural information computed in or shared between visual areas. But linked to quantitative computational models informed by single unit, behavioral and EEG/MEG data— as well as imaging modalities we haven't invented yet—fMRI data provide key information about long-range correlations between localized population responses.

Conclusion

The fundamental conclusion is that, while fMRI is very good at discovering information representations in the brain, it cannot be used to unambiguously connect visual information encoding to specific neural populations. Functional fMRI can discover what information is reflected at a given location but not how that information is encoded. Because the simple availability of information does not help us understand the function of the visual system, fMRI cannot stand on its own in studying vision. The connection between information computation (visual feature extraction) and local neural networks needs to be made by encoding models with parameters constrained by every dataset available—behavioral, electrophysiological, theoretical, and computational. Once a model is specified, fMRI experiments are powerful for testing whether a particular local population might support a given model. No other technique allows us to localize signals with the precision of fMRI, so it remains a critical tool for understanding the mechanisms supporting human visual behaviors. But a critical first step is to reduce the dimensionality of the problem with *a priori* knowledge or principled assumptions of the underlying tuning functions.

References

Baldassano, C., Iordan, M.C., Beck, D.M. & Fei-Fei, L. (2012). Voxel-level functional connectivity using spatial regularization. *Neuroimage* **63**, 1099–1106.

Bonhoeffer, T. & Grinvald, A. (1993). The layout of iso-orientation domains in area 18 of cat visual cortex: Optical imaging reveals a pinwheel-like organization. *The Journal of Neuroscience* **13**, 4157–4180.

Brewer, A.A., Liu, J., Wade, A.R. & Wandell, B.A. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nature Neuroscience* **8**, 1102–1110.

Brouwer, G.J. & Heeger, D.J. (2011). Cross-orientation suppression in human visual cortex. *Journal of Neurophysiology* **106**, 2108–2119.

Callaway, E.M. (2004). Feedforward, feedback and inhibitory connections in primate visual cortex. *Neural Networks* **17**, 625–632.

Cheng, K., Wagooner, R.A. & Tanaka, K. (2001). Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron* **32**, 359–374.

Cukur, T., Nishimoto, S., Huth, A.G. & Gallant, J.L. (2013). Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience* **16**. 763–770.

Darvas, F., Pantazis, D., Kucukaltun-Yildirim, E. and Leahy, R.M. (2004). "Mapping human brain function with MEG and EEG: methods and validation." *Neuroimage* **23 Suppl 1**: S289–299.

Davis, T., LaRocque, K.F., Mumford, J.A., Norman, K.A., Wagner, A.D. & Poldrack, R.A. (2014). What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. *Neuroimage* **97**, 271–283.

de Wit, L.H., Kubilius, J., Wagemans, J. & Op de Beeck, H.P. (2012). Bistable gestalts reduce activity in the whole of V1, not just the retinotopically predicted parts. *Journal of Vision* **12**, 1–14.

DiCarlo, J.J., Zoccolan, D. & Rust, N.C. (2012). How does the brain solve visual object recognition. *Neuron* **73**, 415.

Dumoulin, S.O. & Wandell, B.A. (2008). Population receptive field estimates in human visual cortex. *NeuroImage* **39**, 647–660.

Epshtein, B., Lifshitz, I. & Ullman, S. (2008). Image interpretation by a single bottom-up top-down cycle. *Proceedings of the National Academy of Sciences* **105**, 14298–14303.

Freeman, J., Ziemba, C.M., Heeger, D.J., Simoncelli, E.P. & Movshon, J.A. (2013). A functional and perceptual signature of the second visual area in primates. *Nature Neuroscience* **16**, 974–981.

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

Haak, K.V., Winawer, J., Harvey, B.M., Renken, R., Dumoulin, S.O., Wandell, B.A. & Cornelissen, F.W. (2012). Connective field modeling. *Neuroimage* **66C**, 376–384.

Hall, C.N., Reynell, C., Gesslein, B., Hamilton, N.B., Mishra, A., Sutherland, B.A., O'Farrell, F.M., Buchan, A.M., Lauritzen, M. & Attwell, D. (2014). Capillary pericytes regulate cerebral blood flow in health and disease. *Nature* **508**, 55–60.

Hegde, J. & Kersten, D. (2010). A link between visual disambiguation and visual memory. *The Journal of Neuroscience* **30**, 15124–15133.

Hubener, M., Shoham, D., Grinvald, A. & Bonhoeffer, T. (1997). Spatial relationships among three columnar systems in cat area 17. *The Journal of Neuroscience* **17**, 9270–9284.

Iadecola, C. & Nedergaard, M. (2007). Glial regulation of the cerebral microvasculature. *Nature Neuroscience* **10**, 1369–1376.

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

Kay, K., Naselaris, T., Prenger, R.J. & Gallant, J.L. (2008). Identifying natural images from human brain activity. *Nature* **452**, 352–355.

Kok, P. & de Lange, F.P. (2014). Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Current Biology* **24**, 1531–1535.

Lennie, P. (2003). The cost of cortical computation. *Current Biology* **13**, 493–497.

Logothetis, N. (2008). What we can do and what we cannot do with fMRI. *Nature* **453**, 869–878.

Maier, A., Wilke, M., Aura, C., Zhu, C., Ye, F.Q. & Leopold, D.A. (2008). Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. *Nature Neuroscience* **11**, 1193.

Mannion, D.J., Kersten, D.J. & Olman, C.A. (2013). Consequences of polar form coherence for fMRI responses in human visual cortex. *NeuroImage* **78**, 152–158.

Markov, N.T. Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., Barone, P., Dehay, C., Knoblauch, K. & Kennedy, H. (2014). Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *Journal of Comparative Neurology* **522**, 225–259.

McDonald, J.S., Mannion, D.J. & Clifford, C.W. (2012). Gain control in the response of human visual cortex to plaids. *Journal of Neurophysiology* **107**, 2570–2800.

McLaren, D.G., Ries, M.L., Xu, G. & Johnson, S.C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage* **61**, 1277–1286.

Mountcastle, V.B. (1997). The columnar organization of the neocortex. *Brain* **120**, 701–722.

Naselaris, T., Prenger, R.J., Kay, K., Oliver, M. & Gallant, J.L. (2009). Bayesian reconstruction of natural images from human brain activity. *Neuron* **63**, 902–915.

Nishimoto, S., Vu, A.T., Naselaris, T., Benjamini, Y., Yu, B. & Gallant, J.L. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology* **21**, 1641–1646.

Nowak, L.G. & Bullier, J. (1997). The timing of information transfer in the visual system. In *Cerebral Cortex: Extrastriate Cortex in Primate*, ed. Rockland, K.S., Kaas, J.H. & Peters, A., pp. 870. New York: Plenum Publishing Corporation.

O'Reilly, J.X., Woolrich, M.W., Behrens, T.E.J., Smith, S.M. & Johansen-Berg, H. (2012). Tools of the trade: Psychophysiological interactions and functional connectivity. *Social Cognitive and Affective Neuroscience* **7**, 604–609.

Olman, C., Davachi, L. & Inati, S. (2009). Distortion and signal loss in medial temporal lobe. *PLoS ONE* **4**, e8160.

Olman, C.A., Harel, N., Feinberg, D.A., He, S., Zhang, P., Ugurbil, K. & Yacoub, E. (2012). Layer-specific fMRI reflects different neuronal computations at different depths in human V1. *PLoS One* **7**, e32536.

Olman, C.A., Van de Moortele, P-F., Schumacher, J.F., Guy, J., Ugurbil, K. & Yacoub, E. (2010). Retinotopic mapping with Spin Echo BOLD at 7 Tesla. *Magnetic Resonance Imaging* **28**, 1258–1269.

Olman, C.A. & Yacoub, E. (2011). High-field fMRI for human applications: An overview of spatial resolution and signal specificity. *Open Neuroimaging Journal* **5**, 74–89.

Olshausen, B.A. & Field, D.J. (2004). What is the other 85% of V1 doing? In *Problems in Systems Neuroscience*, ed. Sejnowski, T.J. & Van Hemmen, L., Oxford University Press, New York, New York, USA.

Sayres, R. & Grill-Spector, K. (2008). Relating retinotopic and object-selective responses in human lateral occipital cortex. *Journal of Neurophysiology* **10**, 249–267.

Schumacher, J.F. & Olman, C.A. (2010). High-resolution BOLD fMRI measurements of local orientation-dependent contextual modulation show a mismatch between predicted V1 output and local BOLD response. *Vision Research* **50**, 1214–1224.

Schummers, J., Yu, H. & Sur, M. (2008). Tuned responses of astrocytes and their influence on hemodynamic signals in the visual cortex. *Science* **320**, 1638–1643.

Sherman, S.M. (2007). The thalamus is more than just a relay. *Current Opinion in Neurobiology* **17**, 417–422.

Sun, P., Gardner, J.L., Costagli, M., Ueno, K., Waggoner, R.A., Tanaka, K. & Cheng, K. (2013). Demonstration of tuning to stimulus orientation in the human visual cortex: A high-resolution fMRI study with a novel continuous and periodic stimulation paradigm. *Cerebral Cortex* **23**, 1618–1629.

Swindale, N.V. (1992). Elastic nets, travelling salesmen and cortical maps. *Current Biology* **2**, 429–431.

Tootell, R.B.H. & Hadjikhani, N. (2001). Where is 'Dorsal V4' in human visual cortex? Retinotopic, topographic and functional evidence. *Cerebral Cortex* **11**, 298–311.

Williams, M.A., Baker, C.I., Op de Beeck, H.P., Shim, W.M., Dang, S., Triantafyllou, C. & Kanwisher, N. (2008). Feedback of visual object information to foveal retinotopic cortex. *Nature Neuroscience* **11**, 1439–1445.

Winawer, J., Horiguchi, H., Sayres, R., Amano, K. & Wandell, B.A. (2010). Mapping hV4 and ventral occipital cortex: The venous eclipse. *Journal of Vision* **10**, 1.

Witthoft, N., Nguyen, M.L., Golarai, G., LaRocque, K.F., Liberman, A., Smith, M.E. & Grill-Spector, K. (2014). Where is human V4? Predicting the location of hV4 and VO1 from cortical folding. *Cerebral Cortex* **24**, 2401–2408.

Yacoub, E., Harel, N. & Ugurbil, K. (2008). High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 10607–10612.

Yacoub, E., Shmuel, A., Logothetis, N. & Ugurbil, K. (2007). Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. *NeuroImage* **37**, 1161–1177.

Yoshor, D., Ghose, G., Bosking, W.H., Sun, P. & Maunsell, J.H.R. (2007). Spatial attention does not strongly modulate neural responses in early human visual cortex. *Journal of Neuroscience* **27**, 13205–13209.

Zheng, D., LaMantia, A-S. & Purves, D. (1991). Specialized vascularization of the primate visual cortex. *The Journal of Neuroscience* **11**, 2622–262.