



OPEN ACCESS

EDITED BY

Jodi M. Gilman,
Massachusetts General Hospital and Harvard
Medical School, United States

REVIEWED BY

Jaroslav Tintera,
Institute for Clinical and Experimental Medicine
(IKEM), Czechia
Stephanie Samantha Buss,
Beth Israel Deaconess Medical Center
and Harvard Medical School, United States

*CORRESPONDENCE

Cheryl A. Olman
✉ caolman@umn.edu

SPECIALTY SECTION

This article was submitted to
Cognitive Neuroscience,
a section of the journal
Frontiers in Human Neuroscience

RECEIVED 30 December 2022

ACCEPTED 20 March 2023

PUBLISHED 06 April 2023

CITATION

Olman CA (2023) What multiplexing means
for the interpretation of functional MRI data.
Front. Hum. Neurosci. 17:1134811.
doi: 10.3389/fnhum.2023.1134811

COPYRIGHT

© 2023 Olman. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

What multiplexing means for the interpretation of functional MRI data

Cheryl A. Olman*

Department of Psychology, University of Minnesota, Minneapolis, MN, United States

Despite technology advances that have enabled routine acquisition of functional MRI data with sub-millimeter resolution, the inferences that cognitive neuroscientists must make to link fMRI data to behavior are complicated. Thus, a single dataset subjected to different analyses can be interpreted in different ways. This article presents two optical analogies that can be useful for framing fMRI analyses in a way that allows for multiple interpretations of fMRI data to be valid simultaneously without undermining each other. The first is reflection: when an object is reflected in a mirrored surface, it appears as if the reflected object is sharing space with the mirrored object, but of course it is not. This analogy can be a good guide for interpreting the fMRI signal, since even at sub-millimeter resolutions the signal is determined by a mixture of local and long-range neural computations. The second is refraction. If we view an object through a multi-faceted prism or gemstone, our view will change—sometimes dramatically—depending on our viewing angle. In the same way, interpretation of fMRI data (inference of underlying neuronal activity) can and should be different depending on the analysis approach. Rather than representing a weakness of the methodology, or the superiority of one approach over the other (for example, simple regression analysis versus multi-voxel pattern analysis), this is an expected consequence of how information is multiplexed in the neural networks of the brain: multiple streams of information are simultaneously present in each location. The fact that any one analysis typically shows only one view of the data also puts some parentheses around fMRI practitioners' constant search for ground truth against which to compare their data. By holding our interpretations lightly and understanding that many interpretations of the data can all be true at the same time, we do a better job of preparing ourselves to appreciate, and eventually understand, the complexity of the brain and the behavior it produces.

KEYWORDS

fMRI, primary visual area (V1), predictive coding, feedback–FB, neural circuits

1. Introduction and outline

This article has five main sections. First, a background section contains three very brief overviews that summarize some key background for this perspective and point to more in-depth review articles for readers unfamiliar with the topics. Next, a short section defines and summarizes the problem of multiplexing. Then, two sections present two different analogies—reflection and refraction—for thinking about what multiplexing means for the

interpretation of fMRI data. Finally, a summary section highlights three main consequences of multiplexing: the fact that there are always multiple possible interpretations for each fMRI result does not invalidate the data; we can learn something from fMRI experiments even when they cannot be validated against ground-truth datasets and experiments; feedforward and feedback processes might be impossible to separate in our data.

1.1. Neural architecture

First, we will consider the neural architecture of visual pathways since examples from vision science will support key elements of our argument. Although feedforward and feedback pathways are well-defined (Felleman and van Essen, 1991; Callaway, 2004; Barone et al., 2020), it is rare to observe unidirectional information flow in the brain. Visual representations of scenes and objects do make their way up through the feedforward pathways of the early visual hierarchy to eventually reach higher visual areas and non-visual areas to guide behavior. However, at each stage, deep-layer pyramidal neurons begin to integrate/summarize the state of the local network and feed that information back to earlier stages of information processing, morphing the signal as it rises through the hierarchy. Thus, the firing rates of neurons in V1 are only predicted by retinal inputs for the first 50–80 ms (Gieselmann and Thiele, 2022). After that, we need knowledge of activity throughout the visual and non-visual networks of the brain in order to predict neural responses in even low-level sensory regions.

1.2. High-resolution fMRI

Next, we will consider the spatial scale of fMRI data compared to the neural networks we want to study. Over the past decade, there has been a notable increase in the availability of ultra high-field scanners and sub-millimeter data acquisition techniques. Advances in both acquisition (Adriany et al., 2012; Huber et al., 2017) and denoising (Vizioli et al., 2021) approaches are facilitating the widespread use of depth-resolved fMRI (Koopmans et al., 2010; Olman et al., 2012; Kok et al., 2016; Jia et al., 2020; de Hollander et al., 2021), which seeks to distinguish fMRI responses that are driven primarily by feedforward, intra-regional computations, or feedback based on the location of the response in the depth of cortex. While depth-resolved fMRI signals do probe neural populations that might be biased toward feedforward or feedback responses, fMRI cannot separate these pathways, which are mixed at every depth. This is not a unique limitation of the methodology—even electrophysiological or optical measures that can isolate responses of individual neurons must also wrestle with the fact that the membrane potential, spiking rate, and spiking pattern of each unit is determined by a different mixture of feedforward, lateral, and feedback signals (Figure 1).

1.3. Theoretical neuroscience

Finally, we want to bring into the conversation some perspectives from computational neuroscience, which considers

questions about what information is represented in these neural networks, and why. Most leading theories include the idea that the computation is iterative: while different regions of the brain do produce responses that are selective for different aspects of a given task, none works in isolation and feedforward/feedback loops are pervasive (Kawato et al., 1993; Rao and Ballard, 1999; Friston, 2003; Aitchison and Lengyel, 2017). Not only does it seem as though our brains, in everyday life, use sensory data to update an existing understanding of the world (as opposed to building that understanding from scratch in each moment) but it also appears as though this updating is accomplished by feedback from higher to lower visual areas (Bastos et al., 2012). Thus, there are at least two kinds of signals simultaneously present in a given area of the brain: signals from the feedforward pathway and signals from the feedback pathway. Given the existing electrophysiological evidence for the heterogeneity of contextual modulation of neurons in primary visual cortex at each depth in cortex, it is likely that competing hypotheses are represented in spatially mixed neural populations, some receiving stronger feedforward inputs and others receiving stronger feedback inputs (Figure 1C). Thus, it is very difficult to predict—either for a single neuron or for a well-localized, 0.5 mm voxel—which computation is more strongly reflected in the mixture of field potentials and action potentials that determines the local cortical response.

2. Multiplexing is not a surprise

At this point, we've laid out the basics for an argument that there is no such thing as “the fMRI response in ___,” where you can fill in the blank with whatever brain region you are interested in understanding. Of course, there is a single response, in the sense that each voxel has a single measured intensity at a single point in time. But it can be problematic when “the fMRI response” is used to imply a singular or uniform neural response, because there is not just one neural population response in a region or a voxel. The diversity within the population is striking, and the sources of the signals that influence the local responses are far flung, such that interleaved sub-populations are simultaneously representing different sensory and behavioral states and goals and reflecting computations throughout the brain. This is the problem of multiplexing, and it will remain with us even as technology continues to improve and our resolution becomes more precise.

This problem of multiplexing is not unique to fMRI and has been considered in single-unit electrophysiology for a long time. A full understanding of a single neuron's response requires not just measurement of the firing rate, but also the timing of specific spikes which can, for example, multiplex representation of color and pattern (McClurkin and Optican, 1996). In another example, the variance of a single neuron's firing rate might signal uncertainty (Lengyel and Dayan, 2006). From these and other examples, it has been well established that, because multiple signals are multiplexed in even a single neuron's spiking activity, the answer to “What information does this neuron represent?” depends on the techniques we use to ask the question.

Functional MRI researchers have long been cognizant of this problem, as well, and a large array of analysis techniques are

available to address the issue on a scale appropriate for fMRI. Straight-forward univariate analyses estimate response amplitudes by regressing voxel time series against predictors derived from the experiment design and behavioral responses. Univariate analyses are still common and a valuable tool for understanding local neuronal population activity, but they are rarely considered best practice if applied in isolation. Multi-voxel pattern analyses, encoding models, and connectivity analyses are now part of the standard toolkit for fMRI analysis, and the remaining sections of this paper argues that their power and popularity come from the fact that they address the fact that neural population responses are reflected and refracted in the fMRI signal.

3. Reflection: Treating each fMRI response as just one node in a dynamic network

When we see a reflection in a mirror, we generally know that the object or scene that we are seeing is not actually in the mirror but is standing or happening outside the mirror. The reflection we see is made possible by the nature of the mirror—due to metallic polish or layering of dielectrics, light is turned. So we do learn something about the makeup of an object when we see a reflection in it; rarely, however, do we confuse the mirrored surface for the reflected object we are perceiving.

This section picks three lines of research addressing visual representations of information in primary visual cortex—imagery studies, contextual modulation studies, and estimation of receptive field location—to demonstrate the degree to which visual responses originating in cortical regions outside V1 are reflected in the V1 fMRI signal. In each case, the nature of the information modulating the V1 signal is such that it must originate in cortical regions outside V1, i.e., through feedback pathways. These examples are selected to illustrate instances in which cortico-cortical feedback is just another type of input to V1 and highlight the importance of analysis techniques, such as functional connectivity, that let us characterize these crucial signals driving V1 responses.

Perhaps the most straightforward situation in which the V1 fMRI response reflects neural computations that originate elsewhere in the brain is in studies of visual imagery or expectation. In these studies, a retinotopically-localized response in V1 is measured even though there is no stimulus in the corresponding region of the visual field (i.e., no retinal inputs). If scene context elsewhere in the visual field suggests structure that should be present in a blank region of the visual field, the localized V1 responses reflects the implied structure (Muckli et al., 2015; Kok et al., 2016). If attention is directed to a peripheral object and participants are asked to perform a difficult feature-discrimination task on that object, V1 responses in the fovea, where there is no stimulus, indicate the attended features in the periphery (Williams et al., 2008). Finally, when participants are trained to recall specific images, on cue, while viewing a blank screen, the V1 responses show selectivity for the location and structure of the imagined images (Breedlove et al., 2020). In each case, the identity of the stimulus that is not presented, but rather imagined or inferred, can be decoded (or recovered by the solution of an inverse encoding model, as discussed below) from the localized V1 response. The

selectivity of these signals is arguably due to a combination of (1) the response properties of the local neurons (why else would feedback be targeting these specific retinotopic regions?), and (2) the visual information representations created by remote neural populations. However, in these cases, the emergent V1 fMRI response is fully determined by the observer's behavioral goals and not by any feedforward inputs.

An apparently simple but deceptively complex stimulus paradigm in vision science is orientation-tuned surround suppression. The apparent simplicity derives from its predictability: the overall response in a well-localized population of V1 neurons will be reliably suppressed when a uniform stimulus extends beyond the (shared) classical receptive field of that population (Sillito et al., 1995; Shushruth et al., 2013). If a feature contrast, such as a change in orientation, is introduced between the stimulus in the central (classical RF) region and the stimulus in the extra-classical surround region, the suppression will be reduced or eliminated. Individual neurons might actually experience response facilitation due to the feature contrast, and this is where the complexity begins to enter the picture. Several different V1-intrinsic and -extrinsic factors contribute to the suppression and release from suppression (Zipser et al., 1996; Coen-Cagli et al., 2015), individual neurons experience these effects to different degrees (Cavanaugh et al., 2002), and the magnitude of the effects depends on how the observer's attention is directed (Roberts et al., 2007; Reynolds and Heeger, 2009). Functional MRI studies of these contrast surround suppression phenomena have produced mixed results, ranging from results that are oppose predictions from electrophysiology experiments (Schumacher and Olman, 2010), to results that show little sensitivity of the univariate fMRI response to orientation (Schallmo et al., 2016), to results that nicely match predictions from electrophysiology and behavioral measures of perceived contrast (Zenger-Landolt and Heeger, 2003). These disparate results can be reconciled if a wider range of behavior-dependent, cortico-cortical feedback signals to V1, e.g., 2nd-order contrast (Flevaris and Murray, 2015) are included among the explanatory variables that predict the V1 fMRI response magnitude. Thus, the V1 fMRI response cannot be predicted solely from firing rates measured in similar electrophysiology experiments. Accurate prediction of the V1 fMRI response requires us to include a component of the response that reflects feedback from other regions in the brain (Figure 2A).

Finally, If there is one neuroscientific claim that is broadly accepted since its first demonstration (Hubel and Wiesel, 1959), it is that neurons in primary visual cortex (V1) have spatially restricted receptive fields: a response from a given V1 neuron reliably signals the presence of an object (real or imagined) in that location. However, the vision research literature is rich with examples of the malleability of receptive fields (e.g., David et al., 2004) and we now understand even the V1 receptive field location to be an emergent property of the neural network—something that cannot be completely defined from anatomical measurements of the inputs and outputs for a neuron, but must be measured as a result of aggregated excitatory and inhibitory inputs to a neuron while the stimulus is present. The term “feedback receptive field” has recently been coined (Keller et al., 2020; Kirchberger et al., 2022) to describe the reliable way in which V1 neurons' responses depend on cortico-cortical connections in addition to thalamocortical connections. What is important for this argument is that it is an emergent

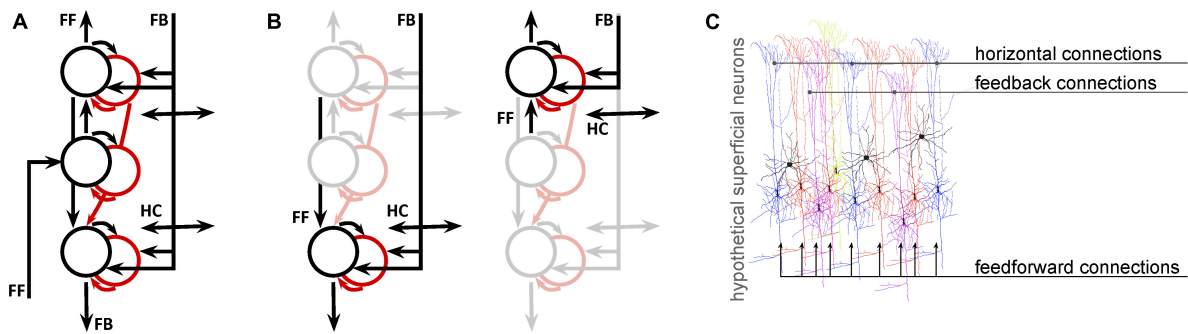
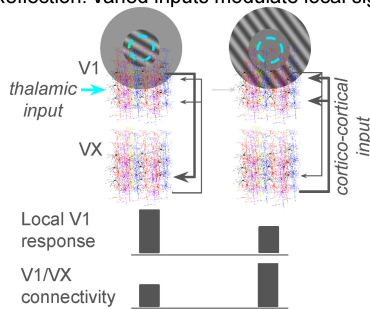


FIGURE 1

Even for fMRI data with sub-millimeter resolution, analysis and interpretation need to address the fact that multiple signals coexist at every location. (A) A simplified network model of deep (toward the bottom), middle, and superficial neural populations (circles: black = excitatory; red = inhibitory) and connections (arrows) in primary visual cortex. Proportions are not represented. (B) Copies of the first panel highlight the fact that, at the mesoscopic level (0.1–0.5 mm), neural populations receive both feedforward and feedback connections. (C) At the microscopic (single-neuron level) it is likely that different neurons receive different projections in different proportions. This is an entirely hypothetical illustration of superficial pyramidal neurons receiving both feedforward, horizontal (intermediate-range), and feedback (long-range) inputs, color coded according to the degree to which each neuron is modulated by each type of input. This is not intended to have specific anatomical validity, but to illustrate the likelihood that even the best-localized population has a heterogeneous feedforward/feedback balance, allowing simultaneous representation of different potential behavioral or perceptual states.

A Reflection: varied inputs modulate local signal



B Refraction: different perspectives show different patterns

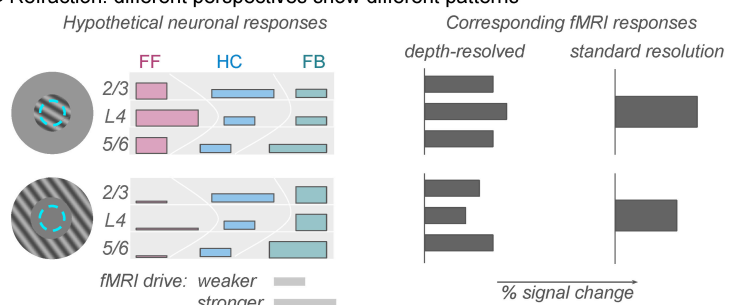


FIGURE 2

The analogies of reflection and refraction are helpful for framing population responses in V1 as a function of spatial context. (A) Reflection: feed-forward receptive fields in V1 are determined by thalamic inputs to the brain (cyan circle: population receptive field location), but the local V1 fMRI response can also reflect strong drive from cortical sources (gray arrows from extrastriate regions, VX). The reflection analogy refers to the fact that this non-zero fMRI response is reflecting figure/ground segmentation computations from other regions in the brain. This allows the local V1 population to perform an important comparison between new sensory information and expected scene configuration. Task-dependent connectivity analyses will discover that correlations between unmodeled fluctuations in the V1 signal and extrastriate (VX) signals are strong during the second condition. (B) Refraction: different analyses will show different patterns of response, but one does not invalidate the other. On the left, colored bars indicate rough predictions of responses in a population receptive field (cyan dashed circle) to stimuli that invoke different levels of feedforward and feedback inputs to the local population (approximately the size of a $3 \times 3 \times 3$ mm fMRI voxels). Top row: stimulation of just the classical receptive field would produce dominantly feedforward responses (pinkish colors), which are strongest in the middle of the gray matter thickness. Bar height represents relative fMRI drive of neurons participating in that pathway; bar width represents estimates from recent literature of the relative fMRI response elicited by feedforward, horizontal, or feedback pathways. Bottom row: with no retinal drive in the classical receptive fields, horizontal and feedback connections will still provide input to the local population whose receptive field is stimulated. Relatively high green bars indicate feedback inputs due to strong figure/ground segmentation cues. Just as the heights and widths of the colored bars on the left are only rough summaries of the anatomical literature, there are also not yet enough data to constrain precise predictions of fMRI response from the diverse neuronal population responses. However, for the sake of argument, illustrative predictions are shown on the right. A depth-resolved analysis with sub-millimeter resolution should show reduced responses in middle and superficial depths in the absence of thalamic input, but deep layers should show relatively large responses. A standard fMRI experiment, on the other hand, should simply show reduced (but not zero) response to the surround-only stimulus.

property not just of the local neural network, but of the extended (extrastriate) network. One experiment that illustrates this effect in human fMRI is the demonstration of a small but reliable spatial shift in the location of V1 responses when an object is perceived to be larger in one context compared to another (Murray et al., 2006).

Each of these examples points to the fact that V1 fMRI responses cannot be understood based solely on feedforward information and intrinsic computations mediated by lateral

interactions; correct analysis of the V1 fMRI response requires additional information about inputs from other cortical regions (feedback).

This challenge is not unique to fMRI. Even electrophysiological studies that sample single-unit responses, and thus are able to characterize the temporal evolution of responses, are rarely able to determine the sources of the modulatory signals that incorporate information about context into the V1 field potentials and firing

rates after the first hundred milliseconds. In sum, these raise the question of whether feedback inputs should just be considered as much a part of the V1-intrinsic response as feedforward inputs.

The challenge, of course, is that feedforward signals are much more amenable to experimental control; they can be largely determined by experimental stimuli and task instructions. Feedback signals, on the other hand, are generated by disparate brain networks that we do not fully understand, which can depend on each person's interpretation of the scene and context that we cannot control. A powerful solution, for fMRI experiments, is offered by task-dependent connectivity analysis.

Connectivity analyses use correlations between fMRI signal fluctuations in different locations to infer neural communication or connectivity between these areas (Friston et al., 1997; Gitelman et al., 2003; Haak et al., 2012; Cisler et al., 2014). Care must be taken to exclude uninteresting sources of common fluctuation, like brain motion or imaging artifacts (Smith et al., 2013) or stimulus-evoked responses (Al-Aidroos et al., 2012). But smart analyses can rule out these confounds and use the fMRI signal to discover something about which we had no *a priori* knowledge: the degree to which signal modulation in one brain region/voxel/neural population/node in a neural network can explain or predict signal modulation in another region. Thus, connectivity analyses offer an important tool for teasing apart the intrinsic and extrinsic factors that drive a well-localized fMRI response (Schindler and Bartels, 2017).

In each of the experiments described above, performing a task-dependent connectivity analysis with the V1 response as a seed should/could/does reveal one or more other regions in the brain where residual signal (the portion of the signal that is left over after stimulus-evoked responses and physiological artifacts are removed by regression) shows task-dependent correlations with the V1 signal. Our lab is one of many that have begun including connectivity analyses in studies of V1 responses; in our case, functional connectivity has been useful for describing responses to contours embedded in complex scenes (Qiu et al., 2016; Pokorný et al., 2021). We find it difficult to draw conclusions about what these long-range correlation signals mean, but it is important to start identifying potential cortical sources of input to V1 alongside known thalamic drive. Directionality of this coupling is difficult (perhaps unwise) to determine with these correlation-based approaches. However, this ability to discover communication partners for V1 is crucial for achieving a full accounting of the inputs to V1 that determine the response, since the V1 fMRI response commonly reflects computations elsewhere in the brain that are multiplexed in the V1 signal as the visual system attempts to make sense of the visual world.

4. Refraction: Accepting that the fMRI signal changes appearance based on perspective

When fMRI was a relatively new technique, vision scientists with a strong background in engineering executed careful experiments designed to validate the fMRI response against neural response properties that were well-characterized by invasive electrophysiology experiments. Retinotopic mapping experiments,

and then ocular dominance and orientation column mapping, validated the spatial precision of fMRI. Drawing on a strong psychophysical tradition in which behavioral sensitivity to simple image features like contrast and orientation can be correlated with the firing rates of superficial pyramidal neurons in V1 (outputs), these early validation experiments succeeded in finding strong and positive relationships between fMRI response magnitude and presumed V1 population firing rates, corroborated by perceptual reports of contrast discrimination (Boynton et al., 1999; Zenger-Landolt and Heeger, 2003).

Another series of careful experiments then started the debate about whether fMRI represents the output of a given brain region (i.e., Layer 2/3 firing rates, as in the early validation experiments) or the local computation—the mess of excitatory and inhibitory synaptic activity that integrates input, local, and feedback connections to shape the selectivity of the output responses. Comparisons between invasive electrophysiology measurements, in both animals and humans, has led to the general conclusion that local field potentials (LFPs), which measure background in flux rather than spiking of large pyramidal neurons, are the best predictors of the blood oxygenation level-dependent (BOLD) hemodynamic response (Logothetis et al., 2001; Maier et al., 2008; Hermes et al., 2019).

There is no inherent conflict between these research perspectives. Both output spiking activity and local field potentials contribute to the fMRI signal. It is likely that they do so to different degrees at different locations depending on the task and experiment design, and our current challenge is in determining the rules for those mixtures. As early pioneers, vision scientists definitely led many quantitative advancements in fMRI (Himmelberg et al., 2022). But one could also argue that vision science, with well-defined sets of external stimuli and a historical focus on recording action potentials from pyramidal cells (Olshausen and Field, 2004), has also encouraged an over-reliance on firing rates and an over-emphasis of external inputs in predicting fMRI responses. Since local field potentials reflect afferent neural activity more strongly than local firing rates (Herreras, 2016), the fact that fMRI is sensitive to both spikes and LFPs means that fMRI is an excellent tool for detecting the diverse information sources that are multiplexed in local neural networks.

Multi-voxel pattern analysis (MVPA) was first introduced as a simple linear classifier applied to visual object recognition (Haxby et al., 2001) and has grown to encompass a wide range of model-free approaches that determine whether the pattern of fMRI responses in a given region can be used to distinguish between different behavioral states or experimental manipulations. As a very open-handed approach (the classifier really is free to use any aspect of signal modulation in the included voxel population to distinguish between experimental conditions) MVPA analyses are well-suited to situations in which we cannot build an *a priori* model for exactly how the underlying neuronal population or the multiplexed signal represents a stimulus or behavioral state. Researchers who seek a principled understanding of causal relationships between task, fMRI response, and neural population activity (a “forward model”) can be frustrated by the difficulty of interpreting a positive result in MVPA analysis, but MVPA techniques are powerful tools for reliably determining whether a particular region of the brain is impacted, in some way, by an experimental manipulation (Kriegeskorte et al., 2008; Williams et al., 2008).

In contrast to MVPA, approaches that use encoding models rely more heavily on *a priori* understanding of the underlying neural responses. An encoding model is a computer simulation of what the fMRI response might be expected to look like, given our expectations for the neuronal activity in a given brain region during an experiment. Once again, vision scientists have been early adopters of this approach (Naselaris et al., 2011), but applications are certainly not limited to vision science [for examples in language and music, (Hoefle et al., 2018; Jain and Huth, 2018)]. The beauty of fMRI analyses using inverted encoding models (Brouwer and Heeger, 2011) is that one can build all sorts of neural response complexity into predictors for individual voxel responses, so the model can capture subtle differences in responses to different experimental conditions. The fact that each brain region has hundreds or thousands of voxels that can be modeled helps address the problem that the models tend to have many more degrees of freedom than the fMRI response. But the challenge, of course, is that the model can only characterize responses anticipated by the experimenter, and fMRI signal modulated along dimensions not captured by the model will go undetected (categorized as unexplained variance).

Importantly, both MVPA and encoding model approaches will produce analysis results that depend on decisions made while setting up the analysis. The results of an MVPA analysis can depend on how you construct and train a classifier, and the results of an encoding model analysis will depend on assumptions made about either the structure and function of the underlying neural responses and/or the neurohemodynamic coupling model used to translate simulated neural responses to simulated fMRI responses (e.g., the degree to which local inhibition and/or LFPs vs. spikes contribute to positive fMRI BOLD responses). As simple examples, either color or orientation (or the conjunction) can be decoded in V1 responses to stimuli that vary along both dimensions (Seymour et al., 2010), and both attended and presented direction can be decoded from population responses to moving stimuli (Kamitani and Tong, 2006). However, the potential variability in results from these analyses represents a limitation that we all need to be aware of, not an inherent flaw that negates the results of the analysis. This is where the analogy of a view through a multi-faceted prism is useful—each view provides useful information about an object that has not yet been fully described (Figure 2B). No view is sufficient to describe the object completely, but each view provides additional information about the neural responses we are trying to understand.

5. Discussion

The ever-increasing spatial precision of fMRI presents us with two challenges. First, the ability to pinpoint responses with sub-millimeter accuracy leads us to make claims that we have identified the location of a computation. However, if each location in the cortex is viewed as a node in an interconnected network, and if the fMRI signal reflects both local spiking responses and the sub-threshold modulatory signals measured in LFPs, then it becomes clear that localized responses reflect remote computations just as meaningfully as they reflect local computations. Second, in even the smallest fMRI voxels there are thousands of neurons,

the responses of which are regulated by different mixtures of feedforward, horizontal, and feedback terms. Thus, no matter how high our resolution, multiple signals are multiplexed in each voxel such that there is not a single computation to uncover.

In the field of depth-resolved (layer) fMRI, it has been proposed that fMRI data can only be correctly interpreted in the context of a forward encoding model that relies on invasive electrophysiological measurements to provide complete characterization of the expected neural responses (Merriam and Kay, 2022). While I am entirely sympathetic with this argument and with the desire to reference experiments against solid ground truth, I also believe that this is currently an unattainable ideal. A ground truth-referenced approach requires either knowing the result of an experiment before conducting it (i.e., merely doing experiments that test whether inference of neural activity provided by non-invasive fMRI in humans produces the right answer) or possessing a validated computational model that can generate accurate predictions for laminar neural activity for novel paradigms. The latter is indeed a primary goal of computational and vision neuroscience; we seek a complete enough understanding of neural code that we can simulate realistic feedforward, lateral, and feedback neural responses. However, the field is not there yet. In the meantime, my argument is that we can continue to do fruitful, informative experiments that do update our understanding of neural responses in human brain, as long as we frame our results as pieces of the puzzle and not the whole picture.

One fundamental challenge that experimental cognitive neuroscientists will continue to wrestle with, in our continued quest to build up a more complete understanding of the brain's function that can generalize from established paradigms to novel paradigms, is our need to study (habit of studying?) the brain out of context. True to our scientific training, we try to control experiments. Experiment control in cognitive neuroscience generally looks like asking a participant to do a specific, narrowly defined task, on cue, for a relatively short amount of time (because we need to repeat each trial several times to get reliable data). The result of careful experiment design is often that, when we isolate the elements of a task, we create disjoint and unpredictable experiences. A predictive coding framework would argue that many of the task fMRI datasets we analyze are biased toward “error” signals because of the abrupt presentation of new information. So while our careful experiment design might do an excellent job of isolating specific neural computations, which we need to do in order to understand the types of things that the brain does, it can lead us assign them the wrong priority, which makes it difficult to predict the outcomes of a new experiment with different context and task demands.

While anatomists and theorists have long known that cytoarchitecturally distinct regions of the brain are richly connected by forward and backward pathways, which are crucial to our ability to infer and learn (Friston, 2003), one of the biggest impacts fMRI has had on cognitive neuroscience has been the provision of tangible views of the interconnectivity and interdependence of brain regions, even at rest (Smith et al., 2013). Different locations in the brain are most certainly specialized for different tasks, with unique connectivity patterns that support different computations or roles in behavior [although see Westlin et al. (2023) for an argument against the “localization assumption”]. However, computations that can be detected at particular locations

quite often originate elsewhere, or would not exist without both the intrinsic and extrinsic connections in a given region [e.g., V1 neurons can be described as having both feedforward and feedback receptive fields (Keller et al., 2020)]. When, for example, the widening or narrowing of a V2 population receptive field coincides with the deployment of spatial attention or an improvement in performance on a crowding task (He et al., 2019), what does it mean that the population receptive field is measured in V2? The size of the receptive field is determined both by the properties of V2 neurons and by V2-extrinsic factors, such that the size of the receptive field is both a property of the responses in V2 and of responses in other regions of the brain yet to be determined. In the end, the multiplexing of the feedforward and feedback signals requires us to draw meaning from fractured and partial views, holding multiple perspectives on neuronal function to be simultaneously true while we figure out how they fit together.

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

References

- Adriany, G., Waks, M., Tramm, B., Schillak, S., Yacoub, E., DeMartino, F., et al. (2012). "An open faced 4 ch. loop transmit / 16 ch. receive array coil for HiRes fMRI at 7 Tesla," in *Proceedings of the 20th annual meeting of the international society for magnetic resonance in medicine*, Melbourne, VIC.
- Aitchison, L., and Lengyel, M. (2017). With or without you: Predictive coding and Bayesian inference in the brain. *Curr. Opin. Neurobiol.* 46, 219–227.
- Al-Aidroos, N., Said, C. P., and Turk-Browne, N. B. (2012). Top-down attention switches coupling between low-level and high-level areas of human visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14675–14680. doi: 10.1073/pnas.1202095109
- Barone, P., Batardiere, A., Knoblauch, K., and Kennedy, H. (2020). Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *J. Neurosci.* 20, 3263–3281.
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron* 76:695.
- Boynton, G. M., Demb, J. B., Glover, G. H., and Heeger, D. J. (1999). Neuronal basis of contrast discrimination. *Vis. Res.* 39, 257–269.
- Breedlove, J. L., St-Yves, G., Olman, C. A., and Naselaris, T. (2020). Generative feedback explains distinct brain activity codes for seen and mental images. *Curr. Biol.* 30, 2211–2224.e2216. doi: 10.1016/j.cub.2020.04.014
- Brouwer, G. J., and Heeger, D. J. (2011). Cross-orientation suppression in human visual cortex. *J. Neurophysiol.* 106, 2108–2119.
- Callaway, E. M. (2004). Feedforward, feedback and inhibitory connections in primate visual cortex. *Neural Netw.* 17, 625–632.
- Cavanaugh, J. R., Bair, W., and Movshon, J. A. (2002). Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *J. Neurophysiol.* 88, 2547–2556. doi: 10.1152/jn.00693.2001
- Cisler, J. M., Bush, K., and Steele, J. S. (2014). A comparison of statistical methods for detecting context-modulated functional connectivity in fMRI. *Neuroimage* 84, 1042–1052. doi: 10.1016/j.neuroimage.2013.09.018
- Coen-Cagli, R., Kohn, A., and Schwartz, O. (2015). Flexible gating of contextual influences in natural vision. *Nat. Neurosci.* 18, 1648–1655. doi: 10.1038/nn.4128
- David, S. V., Vinje, W. E., and Gallant, J. L. (2004). Natural stimulus statistics alter the receptive field structure of v1 neurons. *J. Neurosci.* 24, 6991–7006. doi: 10.1523/JNEUROSCI.1422-04.2004
- de Hollander, G., van der Zwaag, W., Qian, C., Zhang, P., and Knapen, T. (2021). Ultra-high field fMRI reveals origins of feedforward and feedback activity within laminae of human ocular dominance columns. *Neuroimage* 228:117683. doi: 10.1016/j.neuroimage.2020.117683
- Felleman, D. J., and van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47.
- Flevaris, A. V., and Murray, S. O. (2015). Attention determines contextual enhancement versus suppression in human primary visual cortex. *J. Neurosci.* 35, 12273–12280. doi: 10.1523/JNEUROSCI.1409-15.2015
- Friston, K. (2003). Learning and inference in the brain. *Neural Netw.* 16, 1325–1352.
- 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, 218–229.
- Gieselmann, M. A., and Thiele, A. (2022). Stimulus dependence of directed information exchange between cortical layers in macaque V1. *Elife* 11:e2949. doi: 10.7554/eLife.62949
- Gitelman, D. R., Penny, W. D., Ashburner, J., and Friston, K. J. (2003). Modeling regional and psychophysiological interactions in fMRI: The importance of hemodynamic convolution. *NeuroImage* 19, 200–207. doi: 10.1016/s1053-8119(03)00058-2
- Haak, K. V., Winawer, J., Harvey, B. M., Renken, R., Dumoulin, S. O., Wandell, B. A., et al. (2012). Connective field modeling. *Neuroimage* 66, 376–384.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.

Funding

Funding was provided by NIH R01 NS123482 to CO.

Acknowledgments

The author thanks the reviewers as well as Lynda Olman, Steve Engel, and Karen Navarro for feedback on earlier versions of this manuscript.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- He, D., Wang, Y., and Fang, F. (2019). The critical role of V2 population receptive fields in visual orientation crowding. *Curr. Biol.* 29, 2229–2236.e2223. doi: 10.1016/j.cub.2019.05.068
- Hermes, D., Petridou, N., Kay, K. N., and Winawer, J. (2019). An image-computable model for the stimulus selectivity of gamma oscillations. *Elife* 8:e47035. doi: 10.7554/eLife.47035
- Herreras, O. (2016). Local field potentials: Myths and misunderstandings. *Front. Neural Circuits* 10:101. doi: 10.3389/fncir.2016.00101
- Himmelberg, M. M., Gardner, J. L., and Winawer, J. (2022). What has vision science taught us about functional MRI? *Neuroimage* 261:119536. doi: 10.1016/j.neuroimage.2022.119536
- Hoefle, S., Engel, A., Basilio, R., Alluri, V., Toivainen, P., Cagy, M., et al. (2018). Identifying musical pieces from fMRI data using encoding and decoding models. *Sci. Rep.* 8:2266. doi: 10.1038/s41598-018-20732-3
- Hubel, D. H., and Wiesel, T. N. (1959). Receptive fields of single neurons in the cat's striate cortex. *J. Physiol.* 148, 574–591.
- Huber, L., Handwerker, D., Jangraw, D. C., Goense, J. B. M., Poser, B. A., and Bandettini, P. A. (2017). High-resolution CBV-fMRI allows mapping of laminar activity and connectivity of cortical input and output in human M1. *Neuron* 96, 1253–1263. doi: 10.1016/j.neuron.2017.11.005
- Jain, S., and Huth, A. (2018). "Incorporating context into language encoding models for fMRI," in *Proceedings of the 32nd conference on neural information processing systems (NeurIPS 2018)*, Montreal, QC.
- Jia, K., Zamboni, E., Kemper, V., Rua, C., Goncalves, N. R., Ng, A. K. T., et al. (2020). Recurrent processing drives perceptual plasticity. *Curr. Biol.* 30, 4177–4187.e4174. doi: 10.1016/j.cub.2020.08.016
- Kamitani, Y., and Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Curr. Biol.* 16, 1096–1102.
- Kawato, M., Hayakawa, H., and Inui, T. (1993). A forward-inverse optics model of reciprocal connections between visual cortical areas. *Netw. Comput. Neural Syst.* 4, 415–422.
- Keller, A. J., Roth, M. M., and Scanziani, M. (2020). Feedback generates a second receptive field in neurons of the visual cortex. *Nature* 582, 545–549.
- Kirchberger, L., Mukherjee, S., Self, M. W., and Roelfsema, P. R. (2022). Contextual drive of neuronal responses in mouse V1 in the absence of feedforward input. *Sci. Adv.* 9:eadd2498. doi: 10.1126/sciadv.add2498
- Kok, P., Bains, L. J., van Mourik, T., Norris, D. G., and de Lange, F. P. (2016). Selective activation of the deep layers of the human primary visual cortex by top-down feedback. *Curr. Biol.* 26, 371–376.
- Koopmans, P. J., Barth, M., and Norris, D. G. (2010). Layer-specific BOLD activation in human V1. *Hum. Brain Mapp.* 31, 1297–1304.
- Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., et al. (2008). Matching categorical object representations in man and monkey. *Neuron* 60, 1126–1141.
- Lengyel, M., and Dayan, P. (2006). Uncertainty, phase and oscillatory hippocampal recall. *Adv. Neural Inf. Process. Syst.* 19, 833–840. doi: 10.1186/s12868-016-0283-6
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157.
- Maier, A., Wilke, M., Aura, C., Zhu, C., Ye, F. Q., and Leopold, D. A. (2008). Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. *Nat. Neurosci.* 11:1193. doi: 10.1038/nn.2173
- McClurkin, J. W., and Optican, L. M. (1996). Primate striate and prestriate cortical neurons during discrimination I. Simultaneous temporal encoding of information about color and pattern. *J. Neurophysiol.* 71:481.
- Merriam, E. P., and Kay, K. (2022). *The need for validation in layer-specific fMRI*. Available online at: <https://osf.io/f9vqc/> (accessed December 15, 2022).
- Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., et al. (2015). Contextual Feedback to superficial layers of V1. *Curr. Biol.* 25, 2690–2695.
- Murray, S. O., Boyaci, H., and Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nat. Neurosci.* 9, 429–434. doi: 10.1038/nn1641
- Naselaris, T., Kay, K. N., Nishimoto, S., and Gallant, J. L. (2011). Encoding and decoding in fMRI. *Neuroimage* 56, 400–410.
- Olman, C. A., Harel, N., Feinberg, D. A., He, S., Zhang, P., Ugurbil, K., et al. (2012). Layer-specific fMRI reflects different neuronal computations at different depths in human v1. *PLoS One* 7:e32536. doi: 10.1371/journal.pone.0032536
- Olshausen, B. A., and Field, D. J. (2004). "What is the other 85% of V1 doing?," in *Problems in systems neuroscience*, eds T. J. Sejnowski and L. van Hemmen (New York, NY: Oxford University Press).
- Pokorny, V. J., Espensen-Sturges, T. D., Burton, P. C., Sponheim, S. R., and Olman, C. A. (2021). Aberrant cortical connectivity during ambiguous object recognition is associated with schizophrenia. *Biol. Psychiatry Cogn. Neuroimaging* 6, 1193–1201. doi: 10.1016/j.bpsc.2020.09.018
- Qiu, C., Kersten, D. J., and Olman, C. A. (2016). Responses in early visual cortex to contour integration are context-dependent. *J. Vis.* 16:19.
- Rao, R. P., and Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580
- Reynolds, J. H., and Heeger, D. J. (2009). The normalization model of attention. *Neuron* 61, 168–185.
- Roberts, M., Delicato, L. S., Herrero, J., Gieselmann, M. A., and Thiele, A. (2007). Attention alters spatial integration in macaque V1 in an eccentricity-dependent manner. *Nat. Neurosci.* 10, 1483–1491. doi: 10.1038/nn1967
- Schallmo, M. P., Grant, A. N., Burton, P. C., and Olman, C. A. (2016). The effects of orientation and attention during surround suppression of small image features: A 7 Tesla fMRI study. *J. Vis.* 16:19. doi: 10.1167/16.10.19
- Schindler, A., and Bartels, A. (2017). Connectivity reveals sources of predictive coding signals in early visual cortex during processing of visual optic flow. *Cereb. Cortex* 27, 2885–2893. doi: 10.1093/cercor/bhw136
- Schumacher, J. F., and 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. *Vis. Res.* 50, 1214–1224. doi: 10.1016/j.visres.2010.04.005
- Seymour, K., Clifford, C. W. G., Logothetis, N. K., and Bartels, A. (2010). Coding and binding of color and form in visual cortex. *Cereb. Cortex* 20, 1946–1954.
- Shushruth, S., Nurminen, L., Bijanzadeh, M., Ichida, J. M., Vanni, S., and Angelucci, A. (2013). Different orientation tuning of near- and far-surround suppression in macaque primary visual cortex mirrors their tuning in human perception. *J. Neurosci.* 33, 106–119. doi: 10.1523/JNEUROSCI.2518-12.2013
- Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., and Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature* 378, 492–496.
- Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., et al. (2013). Resting-state fMRI in the human connectome project. *Neuroimage* 80, 144–168.
- Vizioli, L., Moeller, S., Dowdle, L., Akcakaya, M., De Martino, F., Yacoub, E., et al. (2021). Lowering the thermal noise barrier in functional brain mapping with magnetic resonance imaging. *Nat. Commun.* 12:5181. doi: 10.1038/s41467-021-25431-8
- Westlin, C., Theriault, J. E., Katsumi, Y., Nieto-Castanon, A., Kucyi, A., Ruf, S. F., et al. (2023). Improving the study of brain-behavior relationships by revisiting basic assumptions. *Trends Cogn. Sci.* 27, 246–257. doi: 10.1016/j.tics.2022.12.015
- Williams, M. A., Baker, C. I., Op de Beeck, H. P., Shim, W. M., Dang, S., Triantafyllou, C., et al. (2008). Feedback of visual object information to foveal retinotopic cortex. *Nat. Neurosci.* 11, 1439–1445.
- Zenger-Landolt, B., and Heeger, D. J. (2003). Response suppression in V1 agrees with psychophysics of surround masking. *J. Neurosci.* 23, 6884–6893. doi: 10.1523/JNEUROSCI.23-17-06884.2003
- Zipser, K., Lamme, V. A. F., and Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *J. Neurosci.* 16, 7376–7389.