in press (May 2020) @ *Visual Cognition* special issue on “Current Perspectives on Visual Working Memory Research”; S. van der Stigchel and C. Olivers, Eds.

Neuroimaging and the localization of function in visual cognition

Bradley R. Postle[1,2,*] and Qing Yu[2]
Departments of [1]Psychology and [2]Psychiatry
University of Wisconsin–Madison
Madison, WI USA

*corresponding author:
postle@wisc.edu
1202 West Johnson St.
Madison, WI 53706
Tel. 608-262-4330
Fax. 608-262-4029

Abstract

Several recent studies have interpreted multivariate evidence for stimulus-specific patterns of activity in parietal and/or frontal cortex as evidence for a representational function in those regions that is qualitatively similar to the representational functions of the visual system. Here we argue that although evidence that a brain system takes on a different configuration for each stimulus in a stimulus set is a necessary property for that system having a role in perceptual representation, such evidence is not specific for this function. Drawing on several recent examples from the recent literature, we offer alternative accounts for understanding stimulus-specificity in parietal and frontal cortex that are consistent with longstanding ideas that activity these regions is best understood as implementing control-related, rather than inherently representational, functions.

The application of multivariate "information-based" analyses to functional magnetic resonance imaging (fMRI) datasets has led to important advances in our understanding of visual attention and working memory. In this commentary we will focus on the question of how to interpret instances of successful discrimination of stimulus-specific patterns of activity in regions of parietal and frontal cortex corresponding to the dorsal attentional network (DAN) and/or the frontoparietal control network (FPC, e.g., Power et al. 2011, Yeo et al. 2011). This is an important question because it has implications for longstanding conceptualizations of a distinction between perceptual and representational functions supported by occipital and temporal regions versus control functions supported by frontoparietal regions. In this commentary we will be considering research that has implications for theoretical frameworks in two domains of visual cognition: feature-based attention and visual working memory. Importantly, however, our goal here is not to provide full-throated advocacy for or against, for example, the biased-competition model of selective attention or the sensorimotor-recruitment model of visual working memory[1]. Rather, our intent is that the specific points made about the small set of studies that we focus on in this commentary will help illustrate a more general question: Does the discovery of equivalent levels of stimulus decoding/encoding in two or more regions necessarily mean that these regions are all supporting the same function? That they are all "doing the same thing"?

Several recent studies have interpreted multivariate evidence for stimulus specificity in regions of the DAN and/or the FPC as evidence for a representational function that is qualitatively similar to the representational functions of regions more classically associated with visual perception and visual semantics. What's common across the specific cases that we will consider here is that successful discrimination of stimulus-specific information in regions of the DAN/FPC has led to conclusions with important implications for our understanding of core

---

[1] For a recent review that more thoroughly covers a broad range of human and nonhuman primate studies, and related theories, the reader is referred to Postle and Oberauer (in press). For full-throated advocacy of sensorimotor recruitment in visual working memory, see Postle (in press).

principles of the organization of two domains of visual cognition: feature-based attention and visual working memory. This commentary will be organized into three sections. In the first we will briefly summarize the methods, results, and conclusions from these studies of feature-based attention and visual working memory. The results from each of these studies have been interpreted by their authors with the "same function" assumption. In the second, we will use logical argumentation to describe alternative possibilities that don't make the "same function" assumption. In the third, we will offer empirical evidence to demonstrate the plausibility of these proposed alternatives. Finally, a "future directions" section will sketch out the kinds of studies that might adjudicate between "same function" versus "different functions" accounts. As we embark on this effort, we want to be very clear that each of the three target studies at are the focus of this commentary used state-of-the-art methods, and we do not question the veracity of the empirical results. We do, however, have questions about some of the mechanistic and/or theoretical conclusions that the authors have drawn from these results.

Equivalent perceptual representations in frontoparietal cortex vis-a-vis occipital cortex?

In a study of feature-based attention, Ester and colleagues (2016) scanned subjects while they fixated centrally and attended to either the orientation or the luminance of a square-wave grating stimulus presented in the periphery. Analysis with inverted encoding models (IEM) indicated that, in regions of occipital, parietal, and frontal cortex, the reconstructed representation of stimulus orientation was stronger when subjects monitored the stimulus for subtle changes of orientation relative to when they monitored it for subtle changes in luminance. Because the neural representations of stimulus orientation in each of these regions were observed to demonstrate qualitatively similar sensitivity to this manipulation of attention, the authors interpreted this finding as problematic for the longstanding and widely accepted idea that frontoparietal circuits serve as a source of top-down attentional control that influences perceptual processing at occipitotemporal sites of stimulus representation (Ester et al. 2016).

In the visual working memory literature there have been several arguments, based on the results of multivariate pattern analyses (MVPA), for a key role for delay-period representation of stimulus information (a.k.a. "storage") in parietal and, in some cases, frontal regions. For example, Xu (2017) has argued that parietal rather than occipital cortex is the primary site of working memory storage, because presenting distractors during the delay period of working memory can abolish evidence for representation of the remembered information in occipital cortex (Bettencourt and Xu 2016), and Christophel et al. (2018) have argued that information held in working memory, but outside the focus of attention, is held in a "low-resolution" (p. 496) format in regions of the intraparietal sulcus (IPS) and the frontal eye fields (FEF) of posterior superior prefrontal cortex (PFC), and not in occipital cortex.

Our goal with this commentary is to make the argument that, contrary to the interpretations of the authors of the papers summarized here, the ability to decode stimulus-specific information in IPS and/or PFC need not imply that these regions are involved in stimulus representation per se. Successful item-level decoding in these regions can also be consistent with other functional interpretations.

Alternative interpretations of stimulus-specific decoding/encoding in parietal and frontal cortex

Evidence that a brain system takes on a different configuration for each stimulus in a stimulus set is a necessary property if that system is to be understood as having a role in stimulus representation. Importantly, however, such evidence is not specific for this function. In the case of a putative source of the top-down selection of one feature over another, for example, it must be the case that this selection mechanism is configured differently for each possible target of selection. Consider the spotlight metaphor. The location within a hypothetical arena that is currently the focus of illumination can be deduced from two pieces of information about the spotlight: the location of the spotlight within the 3D volume of the arena, and the spotlight's pose relative to a stable landmark (e.g., a vertical post stretching from floor to ceiling). Let's say that with a particular configuration of these values one can determine that the spotlight is illuminating a group of seats in Section L of the arena, centered on seat L35. If one were now to re-aim the spotlight so that it will illuminate a seat in Section K, the values that specify its pose will necessarily be different. Thus, one can "decode" the seat that is being illuminated by knowing the location and the pose of the spotlight. Furthermore, if its location is fixed (let's say that the housing that holds the spotlight is welded to our vertical post, 10 meters up from the floor), there is a one-to-one relationship between a unique set of values of pose and each seat in the arena. Thus, we can obtain seat-level decoding from information relating exclusively to the configuration of the spotlight. Importantly, however, this ability to decode seat number from spotlight pose isn't equivalent to being in that seat. Thus, if the lighting technician has never explored the arena, they can't deduce from different spotlight configurations whether, for example, seat L35 or seat K16 is more comfortable to sit in, or that the occupant with season tickets for seat L34 is more polite than the occupant with season tickets for seat K15. This logic generalizes to feature-based attention: Even if there is a one-to-one relationship between each unique filter placed on the lens of the spotlight and each unique feature in the arena that it's capable of highlighting, the ability to decode from the configuration of the filter, for example, precisely what color is being selected is not equivalent to being a member of the sports team who is wearing a jersey whose color matches that filter.

From the argument laid out in the preceding paragraph, it is logically possible that, even though stimulus-specific orientation information can be reconstructed from regions of the DAN, these regions may have been supporting a function that is different from the perceptual function that was presumably being carried out by the occipital regions that also supported item-specific reconstruction. Empirical results that could be consistent with alternative interpretations of stimulus-specific encoding in the DAN will be considered in the next section.

Turning now to the working memory literature, one can see that the argument would be quite similar if one were to accept the overly simplistic proposition that working memory results from temporally extended attentional selection of representations of percepts, knowledge, and/or action plans (e.g., Postle 2006). Even when one broadens the conceptualization of the control of working memory to such constructs as the representation of rules, goals, and plans by frontostriatal circuits (reviewed, e.g., in Postle and Oberauer in press), the logic is similar: The precise configuration of activity supporting a control process will differ as a function of the stimulus information that is being controlled. Thus, for example, the neural configurations of the processes recruited to track the order in which stimuli are presented, to guide behavior according to the rules of the AX-CPT task, or to reorder an arbitrary set of letters into alphabetical order, would be expected to be influenced by the specific identities of the stimuli that these processes are operating on.

Empirical results consistent with alternative interpretations of stimulus-specific decoding/encoding in parietal and frontal cortex

In this section we will again begin with feature- and object-based attention. An important factor to take into account when interpreting the feature-based attention data from Ester et al. (2016) is that this study's procedure entailed the processing of one stimulus per trial. This may have limited its ability to discriminate source-like from site-like patterns of activity. In the biased-competition framework (Desimone and Duncan 1995), a cardinal function of top-down attention is to influence the competition for perceptual representation that is inferred from the fact that activity associated with the processing of a single stimulus item is decreased when a second item is present in the display (e.g., Chelazzi et al. 1998). One example of an empirical consequence of this fact is that studies employing microstimulation of the frontal eye field (a putative source of the endogenous control of spatial attention) have found that attention-like biasing of visual responses in V4 are more pronounced when a distractor is also present in the display (Moore and Armstrong 2003). Results from a more recent neuroimaging study of object-based attention (Sheldon et al. in press), one that did incorporate the need to resolve competition for representation between two discrete objects, suggests a way that the results of Ester et al. (2016) can be reconciled with an anatomically based distinction between sources and sites of attentional control[2].

On each trial in the Sheldon et al. (in press) study, two items drawn from two of three possible categories (face, doughnut, abacus) flashed at a 1 Hz rate and changed state unpredictably and independently, and a cue indicated which item's changes to monitor. Classifier evidence from MVPA (logistic regression) was used as a proxy for the strength of a representation. The pattern of results in frontal and parietal cortex were very similar, and if one were to only consider the effects of cuing in these regions, one would observe a pattern consistent with the interpretation given by Ester et al. (2016) to their data: The representation of the cued item transitioned from intermediate strength during an initial uncued portion of the trial, to elevated strength while it was attended, to baseline-level decoding on trials when a second cue indicated that it was irrelevant for the remainder of the trial. Importantly, although this item remained in one visual field throughout the trial, this pattern was observed in both contralateral and ipsilateral regions of interest (ROI) in frontal and parietal cortex. In ventral occipitotemporal (VOT) cortex, however, the pattern was qualitatively different: Contralateral to the cued item, the representation of the cued item strengthened and that of the uncued item dropped to baseline; ipsilateral to the cued item, the representation of the two items took

---

[2] There are many important aspects of the Sheldon et al. (in press) study that differ from that of Ester et al. (2016), including that one addressed object-based and the other feature-based attention, and that one carried out analyses at the category level and the other at the item level. These will be considered in the section on **Future directions**, but for the purposes of this section of this commentary the results from Sheldon et al. (in press) lend some empirical support to the plausibility of our alternative interpretation of the results of Ester et al. (2016).

on the same strength (i.e., the strength of the contralateral (and unattended) item became the same as that of the ipsilateral (and attended) item). Therefore, the patterns observed in VOT cortex illustrated the biasing effects of attentional selection on the competition for representation at a site of attentional control. The patterns observed in frontal and parietal cortex, in contrast, were identical in each hemisphere. Based on these results, we speculate that had Ester et al. (2016) included a second stimulus in their display, they, too, may have observed important differences of the effects of attention on stimulus reconstructions in frontoparietal versus in temporooccipital regions (See **Future directions**, below).

Returning again to visual working memory, results from previous studies have been interpreted as evidence that distraction (Bettencourt and Xu 2016) and withdrawal of attention (Christophel et al. 2018) result in a loss of delay-period stimulus information from occipital cortex, and that this information therefore also needs to be actively represented in a putatively more robust store in more anterior regions, such as IPS (Xu 2017, Christophel et al. 2018) or PFC (Leavitt et al. 2018). Endorsing this perspective, however, requires the acceptance of null results from Bettencourt and Xu (2016) and from Christophel et al. (2018). In both instances, however, other studies have found positive evidence for delay-period representation of stimulus information in occipital cortex under comparable conditions. In the case of distraction, Rademaker, Chunharas, and Serences (2019) have provided evidence for coexisting representations of both remembered and distracting information in visual cortex. In the case of the withdrawal of attention (Christophel et al., 2018), other studies (van Loon et al. 2018, Yu et al. in press, Wan et al. in-principle accepted) have found evidence for a “priority-based remapping” (Yu et al. in press) of stimulus information in occipital and temporal cortex.

And what about successful stimulus-specific decoding in IPS (Bettencourt and Xu 2016, Christophel et al. 2018) and FEF (Christophel et al., 2018)? Other studies have also reported multivariate evidence for stimulus-specific information in parietal and frontal regions, but some of these have been interpreted as evidence for processes that are not mere duplications of the stimulus-representation functions of visual regions. In frontoparietal cortex, but not in early visual cortex, the delay-period representation of the most recently presented item is stronger than those of items presented earlier in a series (Yu and Shim 2019) (see also Konecky et al. 2017), a pattern that suggests encoding of ordinal context. Relatedly, in IPS and PFC, but not in occipital cortex, stimulus decoding is accompanied by fluctuations in signal intensity that track the demands of binding ordinal context to stimulus identity (Gosseries et al. 2018). (See **Future directions**, below, for more consideration of the distinction between the representation of *content* versus *context* in working memory.) Stimulus specific decoding in occipital cortex, but not IPS, declines monotonically as a function of memory load (Emrich et al. 2013, Gosseries et al. 2018), whereas activity in IPS, but not occipital cortex, scales monotonically with estimates of load-related levels of diffusion in task performance (Yu et al. unpublished). This double dissociation may indicate that qualitatively different functions are supported by activity in these two regions.

Future directions

Although we have articulated some possible alternative explanations of some published examples of stimulus-specific decoding/encoding in parietal and frontal cortex, these ideas will remain nothing more than that until they can be tested empirically. To test the function of attention-sensitive, stimulus-specific patterns of activity in the DAN during feature-based attention, one could carry out a repeated-measures study that included the one-item condition from Ester et al. (2016) and a two-item condition. If one of the two stimuli were presented in each visual field, a 'source-versus-site' account would predict that, during attend-orientation trials, only the orientation of the attended stimulus could be reconstructed from all frontal and parietal ROIs (those ipsilateral as well as those contralateral to the attended stimulus); whereas in the occipital ROI ipsilateral to the attended stimulus, the prediction would be that the orientation of both stimuli could be reconstructed.

To assess the functions supported by stimulus-specific patterns of delay-period activity in visual working memory, in occipital cortex versus in IPS and FEF, one could design a study that varied demands on a control process that these regions are hypothesized to support. One such process is the binding of stimulus identity (e.g., a Gabor patch tilted at 30°: the *content* of working memory) to the trial-unique context in which it was presented (e.g., the location on the screen, or the ordinal position in a series of individually presented stimuli). Some theories hold this function to be essential of what it means for information to be held "in working memory" (Oberauer and Lin 2017), and it is compatible with ideas that IPS encodes a map of priority in the visual environment (e.g., Bisley and Mirpour 2019). A hypothetical experiment, then, might vary the level and/or nature of context binding required in different conditions. For example, every trial might begin with the serial presentation of three Gabor patches, each in a different location on the screen. During some blocks of trials, context is irrelevant, because subjects need only judge whether or not a centrally presented memory probe matches any of the three samples from that trial. The encoding and retention of location context would be required during blocks when the discrimination to be made is whether the identity of the probe, which appears at one of the three sample locations, matches that of the sample that had also appeared at that location. The encoding and retention of ordinal context would be required during blocks when the discrimination to be made is whether the identity of the probe, which is accompanied by the digit "1," "2," or "3," matches that of the sample that had appeared in the ordinal position indicated by the digit. A "same function" account might predict that stimulus-specific patterns of delay-period activity in occipital cortex and in IPS would be equally sensitive to the manipulation of context, whereas a "separation of functions" account might predict that these patterns in occipital cortex would be insensitive to the manipulation of context, but that in IPS they would vary with the nature of the task-critical contextual information.

Conclusion

The fact that one can identify stimulus-specific patterns of activity in multiple regions of the brain is not sufficient to conclude that the same function is being carried out by each of these regions. The incorporation into our research of powerful multivariate techniques has led to many important advances in our understanding of the neural bases of visual attention and working memory. When interpreting the result produced by these techniques, however, it is important to take into account the constraints inherent in the design of the experiment in question, and to contextualize the interpretation of any single finding with the broader body of knowledge about the system(s) under study.

References

Bettencourt, K. C. and Y. Xu (2016). "Decoding the content of visual short-term memory under distraction in occipital and parietal areas." Nature Neuroscience **19**: 150-157.

Bisley, J. W. and K. Mirpour (2019). "The neural instantiation of a priority map." Current Opinion in Psychology **29**: 108-112.

Chelazzi, L., J. Duncan, E. K. Miller and R. Desimone (1998). "Responses of neurons in inferior temporal cortex during memory-guided visual search." Journal of Neurophysiology **80**: 2918-2940.

Christophel, T. B., P. Iamshchinina, C. Yan, C. Allefeld and J.-D. Haynes (2018). "Cortical specialization for attended versus unattended working memory." Nature Neuroscience **21**: 494-496.

Desimone, R. and J. Duncan (1995). "Neural mechanisms of selective visual attention." Annual Review of Neuroscience **18**: 193-222.

Emrich, S. M., A. C. Riggall, J. J. Larocque and B. R. Postle (2013). "Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory." Journal of Neuroscience **33**: 6516-6523.

Ester, E. F., D. W. Sutterer, J. T. Serences and E. Awh (2016). "Feature-selective attentional modulations in human frontoparietal cortex." Journal of Neuroscience **36**: 8188-8199.

Gosseries, O., Q. Yu, J. J. LaRocque, M. J. Starrett, N. Rose, N. Cowan and B. R. Postle (2018). "Parieto-occipital interactions underlying control- and representation-related processes in working memory for nonspatial visual features." Journal of Neuroscience **38**: 4357-4366.

Konecky, R. O., M. A. Smith and C. R. Olson (2017). "Monkey prefrontal neurons during Sternberg task performance: full contents of working memory or most recent item? ." Journal of Neurophysiology **117**: 2269-2281.

Leavitt, M. L., D. Mendoza-Halliday and J. C. Martinez-Trujillo (2018). "Sustained activity encoding working memories: not fully distributed." Trends in Neurosciences **40**: 328-346.

Moore, T. and K. M. Armstrong (2003). "Selective gating of visual signals by microstimulation of frontal cortex." Nature **421**: 370-373.

Oberauer, K. and H.-Y. Lin (2017). "An interference model of visual working memory." Psychological Review **124**: 21-59.

Postle, B. R. (2006). "Working memory as an emergent property of the mind and brain." Neuroscience **139**: 23-38.

Postle, B. R. (in press). Cognitive neuroscience of visual working memory. Working Memory: State of the Science. R. H. Logie, V. Camos and N. Cowan. Oxford, U.K., Oxford University Press.

Postle, B. R. and K. Oberauer (in press). Working memory: Theoretical, computational, and neural considerations. The Oxford Handbook of Human Memory. M. J. Kahana and A. D. Wagner. Oxford, U.K., Oxford University Press.

Power, J. D., A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Miezen, B. L. Schlaggar and S. E. Petersen (2011). "Functional network organization of the human brain." Neuron **72**: 665-678.

Rademaker, R. L., C. Chunharas and J. T. Serences (2019). "Coexisting representations of sensory and mnemonic information in human visual cortex." Nature Neuroscience **22**: 1336-1344.

Sheldon, A. D., E. Saad, M. I. Sahan, E. Meyering, M. J. Starrett, J. J. LaRocque, N. S. Rose and B. R. Postle (in press). "Attention biases competition for visual representation via enhancement of targets and inhibition of nontargets." Journal of Cognitive Neuroscience.
van Loon, A. M., K. Olmos-Solis, J. J. Fahrenfort and C. N. L. Olivers (2018). "Current and future goals are represented in opposite patterns in object-selective cortex." eLife **7**: e38677.
Wan, Q., Y. Cai, J. Samaha and B. R. Postle (in-principle accepted). "Tracking stimulus representation across a 2-back visual working memory task." Royal Society Open Science **in-principle accepted registered report**.
Xu, Y. (2017). "Reevaluating the sensory account of visual working memory storage." Trends in Cognitive Sciences **27**: 794-815.
Yeo, B. T., F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, B. Fischl, H. Liu and R. L. Buckner (2011). "The organization of the human cerebral cortex estimated by intrinsic functional connectivity." Journal of Neurosphysiology **106**: 1125-1165.
Yu, Q., M. F. Panichello, B. R. Postle and T. J. Buschman (unpublished). "Delay-period activity in frontal, parietal, and occipital cortex tracks different attractor dynamics in visual working memory." bioRxiv.
Yu, Q. and W. M. Shim (2019). "Temporal-order-based attentional priority modulates mnemonic representations in parietal and frontal cortices." Cerebral Cortex **29**: 3182-3192.
Yu, Q., C. Teng and B. R. Postle (in press). "Different states of priority recruit different neural codes in visual working memory." PLoS Biology.