Occipitotemporal Representations are Modulated by Conceptual Knowledge and Interact with a Frontoparietal Network

Kurt Braunlich (k.braunlich@ucl.ac.uk)
University College London, 26 Bedford Way, London, UK

Bradley C. Love (b.love@ucl.ac.uk)
University College London, 26 Bedford Way, London, UK
The Alan Turing Institute, 96 Euston Rd., London, UK

Abstract

Through selective attention, knowledge of abstract concepts can emphasize informative stimulus features. For example, the size of garments is critical when choosing what to purchase, whereas weight may be more important when deciding how to ship them. In two fMRI studies, we investigated whether neural representations of stimulus features vary according to these latent attentional effects. During scanning, participants categorized stimuli according to strategies that were learned through trial and error. Formal categorization models were fit to the behavioral data from each participant, and provided attentional weight estimates for each binary stimulus dimension. We found that when greater attention was devoted to a particular visual dimension (e.g., color), its value (e.g., red) was more easily decoded from occipitotemporal cortex; indicating that conceptual knowledge can modulate representations of individual perceptual features. To better understand this effect, we conducted a multivariate connectivity analysis, which allowed us to identify regions involved in reading out and/or modulating these visual feature representations. The results suggest that several frontoparietal regions integrate or modulate occipitotemporal stimulus feature representations.

Keywords: Concepts; Attention; fMRI; Perception

Introduction

A classic idea in the categorization literature is that decision makers learn to distribute attention across stimulus features in a way that optimizes the discriminability of behaviorally-relevant information. Previous work has shown that the behavioral relevance of perceptual attributes can modulate the discriminability of their representation in occipitotemporal cortex (Reynolds & Chelazzi, 2004), and that categorization training can have lasting influences on occipitotemporal representations: increasing the discriminability of relevant stimulus dimensions (Folstein, Palmeri, & Gauthier, 2013). This implies that occipitotemporal representations might closely reflect the learned concepts and strategies used by decision-makers during visual categorization, and might thus covary with attentional parameters derived from formal categorization theory.

To test this hypothesis, we re-analyzed two previously published datasets. In the first (Mack, Preston, & Love, 2013), participants categorized abstract stimuli that varied according to four binary perceptual dimensions (shape, size, position and color), according to the classic 5/4 categorization structure (Medin & Schaffer, 1978). In the second (Mack, Love, & Preston, 2016), participants categorized pictures of insects that varied according to three binary perceptual dimensions (antenna shape, leg thickness, and mandible shape), according to the type 1, 2, and 6 categorization structures described by Shepard, Hovland and Jenkins (1961). An important aspect of both experiments was that the category structure defined the relevance of each visual attribute to the categorization decision; and the mapping of perceptual attributes to their role in the category structure was randomized for each participant. This allowed us to differentiate effects associated with perceptual representation from those associated with conceptual knowledge. Additionally, in both experiments, participants learned how to categorize the stimuli through trial and error, mirroring the way concepts are often learned in the real-world.

Attention in Occipitotemporal Cortex

To investigate whether attention influenced the discriminability of the stimulus dimensions, we used a cross-validated searchlight analysis to identify regions representing each of the perceptual stimulus dimensions. The resultant occipitemporal ROI from the 5/4 dataset is illustrated in Figure 1 (bottom left). We used a similar ROI for the second dataset. In both datasets, mixed-effects linear regression analyses indicated that the discriminability of the individual perceptual dimensions positively covaried with attentional parameters derived from formal categorization theory (Love, Medin, & Gureckis, 2004; Nosofsky, 1986).

Multivariate Pattern Connectivity

As this attentional effect implies that occipitotemporal representations interact with higher-order brain regions, we conducted a multivariate connectivity analysis. Using the 5/4 dataset, we projected the occipitotemporal neuroimaging data into a lower-dimensional subspace spanned by a set of basis vectors that provided estimates of both the sign (i.e., the evidence towards one particular dimensional attribute over the other, e.g., red vs. green) and the discriminability of each perceptual dimension (Figure 1). We then used a cross-validated searchlight analysis to identify regions that might read-out this occipitotemporal information.

To confirm that our findings reflected the communication of behaviorally-relevant information, we conducted a permu-

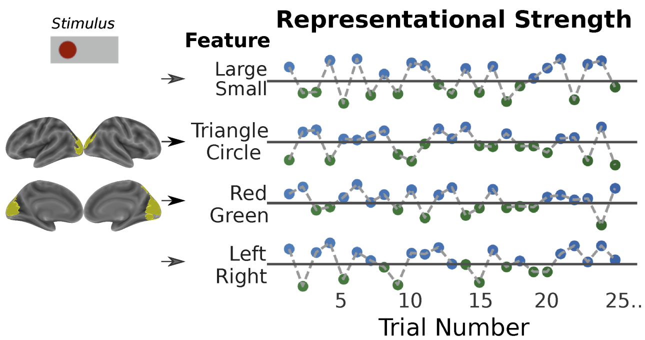

Figure 1: **Top Left:** Example Stimulus from the 5/4 Experiment. The stimuli varied according to four perceptual dimensions (position, shape, color and size). **Bottom Left:** The Occipitotemporal ROI. **Right:** To characterize OTC representations of individual stimulus dimensions, we projected the neuroimaging data onto a low-dimensional subspace spanned by four basis vectors. Each vector provided an estimate of the value (e.g., red vs. green) and the strength of a particular visual dimension.

tation test, which involved randomly shuffling the values of each visual dimension, and then projecting the occipitotemporal neuroimaging data onto the resulting random subspace. The results illustrated in Figure 2 reflect a searchlight analysis in which this procedure was used to obtain a null distribution for each voxel.

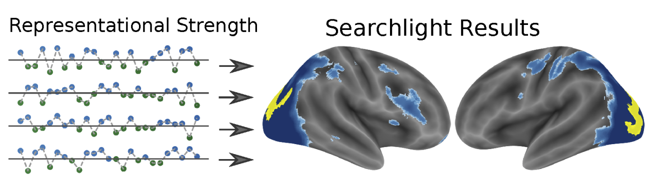

Figure 2: Full-Brain Connectivity Analysis. **Yellow:** The occipitotemporal ROI shown in Figure 1. **Blue Colormap:** Results from the permuted searchlight analysis (voxelwise control for multiple comparisons: $p < 0.05$).

This analysis indicated that activity in the right frontal eye fields (FEF), right inferior frontal sulcus (IFS), bilateral intraparietal sulcus (IPS), and left pre- and post-central motor cortices supported the decoding of the occipitotemporal basis vectors. The FEF, IFS and IPS have been previously associated with top-down modulation of visual activity. Our results may thus reflect mutually-informative interactions between bottom-up and top-down signals. Interestingly, the right IFS (which is known to play an important role in domain-general cognitive control), also represented the overall entropy of the category choice, suggesting that it might modulate attention to distinct visual attributes through consideration of abstract decisional uncertainty. The left motor cortex represented the strength of evidence for specific behavioral responses (which were made via fingers of the right hand).

Conclusion

A feature common to many categorization models is a selective attention parameter that governs the influence of particular visual dimensions on the decision outcome. In the current work, we found that the discriminability of occipitotemporal representations covaried with attentional parameters derived from formal categorization theory. This indicates that occipitotemporal representations closely reflect conceptual knowledge. To better understand how visual representations interacted with abstract knowledge, we used a multivariate connectivity analysis, and found that activity in several frontoparietal regions was consistent with a role in integrating abstract meaning from occipitotemporal visual representations, and/or modulating attention to specific visual dimensions.

Acknowledgments

We thank the authors of the original studies for sharing their data. This work was supported by NIH Grant 1P01HD080679, Leverhulme Trust grant RPG-2014-075 and Wellcome Trust Senior Investigator Award WT106931MA to BCL.

References

Folstein, J. R., Palmeri, T. J., & Gauthier, I. (2013). Category learning increases discriminability of relevant object dimensions in visual cortex. *Cerebral Cortex*, *23*(4), 814–23.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *111*(2), 309–32.

Mack, M. L., Love, B. C., & Preston, A. R. (2016). Dynamic updating of hippocampal object representations reflects new conceptual knowledge. *PNAS*, *113*(46), 13203–13208.

Mack, M. L., Preston, A. R., & Love, B. C. (2013). Decoding the brain's algorithm for categorization from its neural implementation. *Current Biology*, *23*(20), 2023–2027.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207–238.

Nosofsky, R. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology General*, *115*(1), 39–61.

Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience*, *27*(1), 611–647.

Shepard, R., Hovland, C., & Jenkins, H. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, *75*(13).