A Highly Selective Neural Response to Food in Human Visual Cortex Revealed by Hypothesis-Free Voxel Decomposition

Meenakshi Khosla (mkhosla@mit.edu), N Apurva Ratan Murty, Elizabeth Ann Mieczkowski and Nancy Kanwisher
McGovern Institute for Brain Research and Department of Brain and Cognitive Sciences,
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract:
Prior work has identified regions of high-level visual cortex selectively responsive to faces, places, bodies, and words. However, this largely hypothesis-driven work cannot reveal how prominent these category selectivities are in the overall functional organization of the visual cortex, or what other un-hypothesized selectivities exist. Further, standard methods cannot detect selective neural populations that coexist with functionally distinct populations within voxels. To overcome these limitations, we applied a data-driven voxel decomposition analysis to identify a robust set of component response profiles consistent across subjects in a recently released public data set of fMRI responses to thousands of complex photographic stimuli (Allen et al., 2021). Four of the five top components revealed by our analysis were clearly selective for people, faces, scenes, and words. The analysis also revealed a novel component with a distinct anatomy that responded highly selectively to images of food. Alternative accounts based on lower-level visual features like color, shape or texture failed to account for the high-level category selectivity of this component. Analyses of independent data revealed the same top components, replicating these dominant dimensions, including a food-selective component, in new subjects.

Keywords functional organization; category selectivity; naturalistic; high-resolution fMRI; hypothesis-free; food

Introduction

The last few decades of research in human cognitive neuroscience have revealed the functional organization of the cortex in rich detail. This organization features a set of regions that are selectively engaged in single mental processes, from perceiving faces or places or music, to understanding the meaning of a sentence or inferring the content of another person's thoughts. Why do our brains have these specializations, and apparently not others? To answer these questions, we need to know whether the specializations we have discovered to date are the foundational ones, or whether other important specializations exist that have not yet been discovered.

Our inventory of selective neural responses may be biased and/or incomplete for several reasons. First, most prior research on the ventral visual pathway has tested a relatively small number of stimulus categories and may not have subtended the relevant part of stimulus space preferred by some neural populations. Second, this work has proceeded largely in a hypothesis-driven fashion, so neural selectivities that defy conventional wisdom or are less suspected may have been missed. Third, prior research based on voxel-wise contrasts is not well suited to discover neural populations that may be highly selective, but whose selectivity is masked by the blurring inherent in the fMRI BOLD signal. Here, we overcome all three limitations by analyzing fMRI responses to a very broad and large set of naturalistic stimuli with a data-driven analysis method that can de-mix responses from neural populations that are spatially intermingled within voxels.

Results

We present a hypothesis-neutral approach for analyzing fMRI recordings from 8 participants while they each viewed 9,000-10,000 images in a recently released dataset (Allen et al., 2021). Our goal is to identify the dominant functionally distinct neural populations in the ventral visual pathway, and their anatomical locations, including novel selectivities that may have evaded standard analyses due to the spatial overlap of distinct neural populations within voxels.

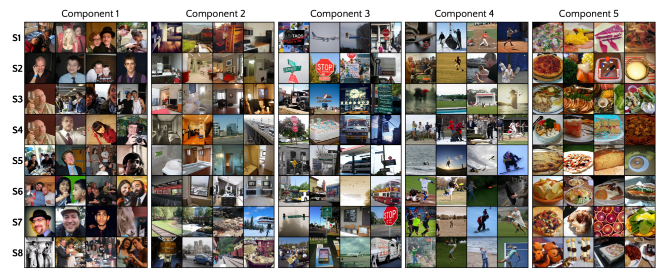

Figure 1: Top images for each component in all 8 Ss.

Our approach relies on Bayesian non-negative matrix factorization (Schmidt et al., 2009) that expresses the fMRI responses to natural scene images as the product of a response profile matrix that characterizes the response of each component (distinct neural populations) to all stimuli and a voxel weight matrix that represents the anatomical proportions of different neural populations within every voxel. This approach allows us to computationally de-mix voxel responses to recover the activity of overlapping neural populations. We first performed this analysis on 4 subjects who each viewed 10,000 stimuli with 3 repetitions each. Subject-specific components were derived by applying the algorithm on individual subjects' data matrices. We then selected the number of components that led to the lowest Bayesian Information Criterion (BIC), which was ~20 in all participants. Next, we searched for components shared across participants, using the 1,000 images viewed by each participant to align the components derived separately for each participant. We then repeated the same analysis on the held-out group of subjects from the dataset. This analysis revealed the same top 5 components, providing an independent replication of the discovered organization.

Data-driven component modeling reveals dominant selectivities in the ventral visual cortex. We first examined the images producing the strongest response in each of the top five components (Fig. 1). These images strongly suggested that the first four components were selectively responsive to faces, places, words and bodies, consistent with extensive prior research and hence providing strong positive controls for our method. Intriguingly, the fifth component appeared to respond in a highly selective fashion to images of food and hinted towards a novel specialization. We subsequently collected food salience ratings on the shared image set to test the preferred dimension of this component. As shown in Fig. 2, most salient food images are concentrated at the top for this component, producing a ramp-shaped tuning curve.

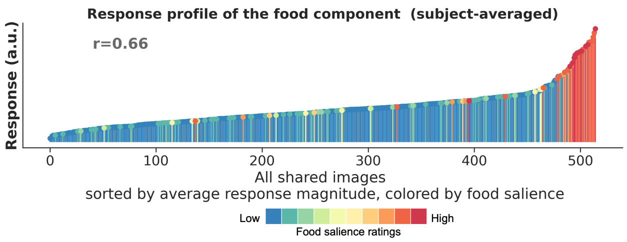

Figure 2: Response profile of the food component, averaged across all 8 participants for the shared 515 images, colored by food salience ratings.

Food selectivity of the novel component cannot be explained by alternative lower-level visual features.

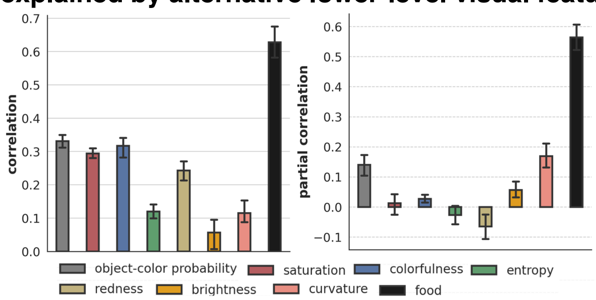

Figure 3: Correlation (left) and partial correlation (right) of image-computable properties and food salience ratings with the novel component's responses.

We next tested whether the novel component's responses might be explained by low-level features. Besides food, the top images for this component also had other common attributes, including warmer colors, higher curvature, and a complex spatial structure with rich texture. However, most of the variance explained by these properties was shared with food salience ratings (Fig. 3, right) and these properties, by themselves, failed to explain much of the response variation of this novel component. Interestingly, when presented with computationally matched food and non-food image pairs that elicit similar activations in a pre-trained DNN (and are thus visually very similar), this component still responds more strongly to food than the corresponding matched non-food images. High-throughput testing and further control experiments on a highly accurate computational model of this component additionally validated its highly food-selective nature.

Voxel decomposition reveals the anatomical organization of components. We next characterized the anatomical distribution of each component by projecting its voxel weights back into anatomical coordinates. For known selectivities, the component anatomies exhibited remarkable agreement with the corresponding regions of interest identified with an independent functional localizer: the place component was largely localized within the parahippocampal place area (PPA), the face component produced highest voxel weights in the fusiform face area (FFA), the word component was concentrated within the visual word form area (VWFA) and the bodies component in parts of FFA and the extrastriate body area (EBA). The food component weights appeared spatially patchy like the other components, with two main clusters, one medial and one lateral to the FFA, though with substantial spatial heterogeneity across subjects.

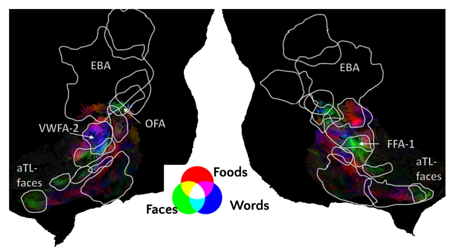

Figure 4: Voxel weights of the Face, Food, and Word components for one subject on their cortical flatmap.

De-mixing reveals stronger selectivity for components than voxels. The novel selectivity for food begs the question of why it was not discovered before, particularly in previous hypothesis-driven investigations (e.g., Downing et al., 2006)? We speculated that prior studies employing voxel-wise comparisons might have missed food selectivity due to the spatial overlap of food-selective populations with other distinctive neural populations within voxels. To investigate this possibility, we computed the food selectivity of raw voxels by correlating the response of all voxels in the ventral visual stream with the food salience ratings on the shared image set. In all 8 subjects, the food selectivity of all voxels was substantially lower than the food selectivity of their respective component, suggesting that de-mixing is critical for discovering this strong food selectivity, which is much less strong in raw voxel responses.

Conclusion

Our hypothesis-neutral investigation of a large-scale dataset revealed that selective neural response for faces, places, words and bodies not only exist in the ventral visual pathway, but are prominent features of the neural response of this region. This analysis further revealed a novel selectivity for food that was not explainable by low-level or mid-level visual features like color, shape and texture.

Acknowledgements

NIH Pioneer Award NIH DP1HD091957; NSF Science and Technology Center—Center for Brains, Minds, and Machines Grant NSF CCF-1231216, NIH K00/R00 Pathway to Independence (NEI, Award Number 1K99EY032603 to N.A.R.M)

References

Allen, E. J., St-Yves, G., Wu, Y., Breedlove, J. L., Prince, J. S., Dowdle, L. T., ... & Kay, K. (2022). A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. Nature neuroscience, 25(1), 116-126.

Downing, P. E., Chan, A. Y., Peelen, M. V., Dodds, C. M., & Kanwisher, N. (2006). Domain specificity in visual cortex. Cerebral cortex, 16(10), 1453-1461.

Schmidt, M. N., Winther, O., & Hansen, L. K. (2009, March). Bayesian non-negative matrix factorization. In International Conference on Independent Component Analysis and Signal Separation (pp. 540-547). Springer, Berlin, Heidelberg.