Core Neural Dimensions of Functionally Selective Areas in the Human Visual Cortex

Leonard E. van Dyck (leonard.van-dyck@psychol.uni-giessen.de)
Department of Psychology, Justus Liebig University Giessen
Otto-Behaghel-Str. 10F, Giessen, 35394, Germany
Max Planck Institute for Human Cognitive and Brain Sciences
Stephanstr. 1A, Leipzig, 04103, Germany

Martin N. Hebart (hebart@cbs.mpg.de)*
Department of Medicine, Justus Liebig University Giessen
Ludwigsplatz 13-15, Giessen, 35390, Germany
Max Planck Institute for Human Cognitive and Brain Sciences
Stephanstr. 1A, Leipzig, 04103, Germany
Center for Mind, Brain and Behavior, Universities of Marburg, Giessen, and Darmstadt
Hans-Meerwein-Str. 6, Marburg, 35032, Germany

Katharina Dobs (katharina.dobs@psychol.uni-giessen.de)*
Department of Psychology, Justus Liebig University Giessen
Otto-Behaghel-Str. 10F, Giessen, 35394, Germany
Center for Mind, Brain and Behavior, Universities of Marburg, Giessen, and Darmstadt
Hans-Meerwein-Str. 6, Marburg, 35032, Germany

*equal contribution

Abstract

Prior research has extensively documented functional selectivity for categories within visual cortical areas, primarily by contrasting neural responses to images from various categories. However, such categorical approaches are less suitable to capture the diversity of neural representations within these areas. Do category-selective areas encode holistic categories, or are they instead tuned to multifaceted features? To address this question, we employed non-negative matrix factorization (NMF) for analyzing human fMRI responses to natural images in face-, body-, and scene-selective areas, which uncovered a consistent set of interpretable neural dimensions across participants. These dimensions not only aligned with the areas' respective category preferences, but also revealed finer within-category distinctions, indicating selective tuning to diverse visual input features. Mapping these dimensions onto the cortical surface showed both clustered and distributed topographies, which accounted for overlaps between areas. Our results suggest that category-selective areas show multifaceted feature tuning, challenging traditional views and highlighting the complex interplay of neural dimensions in encoding visual information.

Keywords: fMRI; functional selectivity; visual cortex; neural representations; dimensions

Introduction

In the visual cortex, specific functional regions of interest (fROIs) are known to respond selectively to categories such as faces, bodies, and scenes (Downing et al., 2005; Grill-Spector & Weiner, 2014). However, a central question remains largely unanswered: do these fROIs represent holistic categories or multifaceted features? Several studies have identified nested functional selectivity within fROIs, revealing within-category distinctions and corresponding subclusters (Bracci et al., 2015; Çukur et al., 2013, 2016; Weiner & Grill-Spector, 2011). Other research proposes continuous organizing principles–such as animacy, real-world size (Konkle & Caramazza, 2013), aspect ratio (Bao et al., 2020), and semantics (Huth et al., 2012)–that span much of the visual cortex, including fROIs, challenging the conventional notion of category selectivity.

While categorical approaches typically examine behavioral or neural responses averaged across predefined categories, dimensional approaches decompose these responses into their parts, enabling a more detailed understanding of the driving features (Contier et al., 2023; Hebart et al., 2020). Here, we leverage such a data-driven approach, inspired by Khosla et al. (2022), to extract neural dimensions from fMRI responses within multiple fROIs in the human visual cortex. This allows us to explore the diversity of neural representations and uncover their naturally emerging organizing principles.

Methods

We analyzed large-scale human fMRI responses from the Natural Scenes Dataset (NSD; Allen et. al 2022), focusing on 4 participants (10,000 individual images, 1,000 shared images, 3 repetitions each). Our analyses targeted 6 fROIs–the fusiform face area (FFA), occipital face area (OFA), fusiform body area (FBA), extrastriate body area (EBA), parahippocampal place area (PPA), and occipital place area (OPA)–identified using a standard functional localizer (Stigliani et al., 2015; $t > 2.5$).

We employed Bayesian non-negative matrix factorization (BNMF) with Gibbs sampling (Schmidt et al., 2009) to extract neural dimensions from fMRI responses of fROIs combined across hemispheres. BNMF decomposes the data matrix (V) into a response matrix (W) and a weight matrix (H), facilitating the identification of part-based, sparse, and interpretable dimensions. After baseline shifting voxel-wise responses to ensure the required non-negativity, we optimized the dimensionality (k^*) for each fROI and participant by minimizing the Akaike Information Criterion (AIC). In a first-level consensus approach (**Figure 1A**), we identified consistent dimensions within participants through multiple BNMF runs (N=50) with k^* and random initializations, using outlier detection and k^*-means clustering. In a second-level consensus approach (**Figure 1B**), we identified consistent dimensions across participants by comparing dimensions via pairwise correlation of shared images, employing a greedy selection for those with mean correlations above $r > .3$.

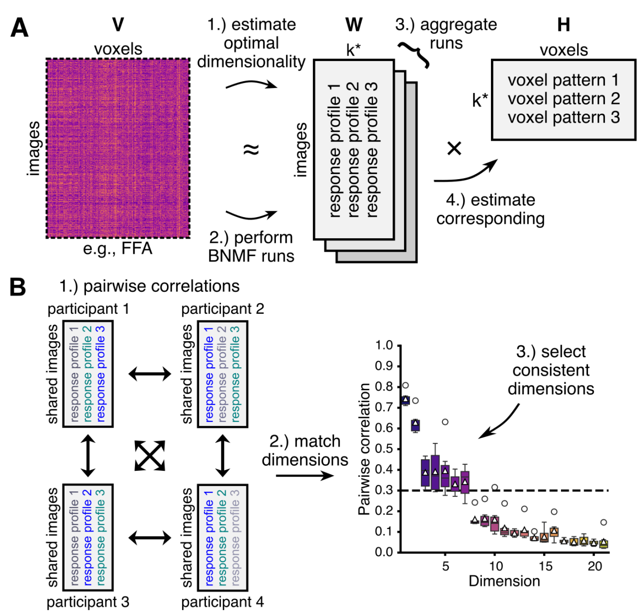

Figure 1: Overview of the BNMF approach inspired by Khosla et al. (2022).

Results

We discovered a diverse set of neural dimensions within each fROI, ranging from 3 to 15, with high consistency across participants (mean r=.47, SD=.14).

Are these neural dimensions interpretable? To explore this question, we conducted an online behavioral experiment in which participants (N=9) labeled the dimensions based on samples of highest and lowest scoring images. Remarkably, the dimensions appeared to be interpretable. As expected, these dimensions primarily aligned with the areas' respective category preferences (**Figure 2**). However, in line with previous findings, the dimensions also revealed finer distinctions, such as indoor and outdoor settings in scene-selective areas (Epstein & Baker, 2019), various body parts in body-selective areas (Bracci et al., 2015), and food in several areas (Jain et al., 2023; Khosla et al., 2022). This indicates that fROIs encode information beyond category labels, and that the neural dimensions identified are comprehensible.

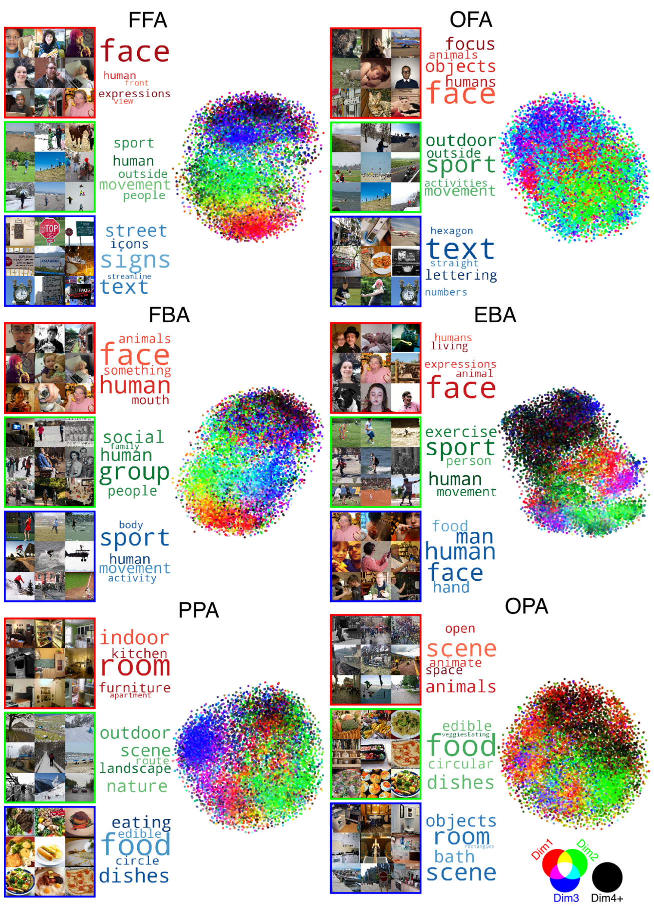

Figure 2: Top 3 consistent dimensions per fROI from participant P1 (for brevity). Highest scoring images, most frequent labels, and 2D t-SNE visualization (perplexity=100) of the neural similarity space. Dots represent images arranged by neural similarity and colored according to independently rescaled response profiles that correspond to each dimension.

But do these dimensions fully capture how fROIs represent visual information beyond the highest scoring images? We used RSA to determine how similar the elicited neural responses to individual images were in each fROI. We found that images with similar neural responses were also tuned to similar dimensions. This suggests that the dimensions are effective in capturing the diverse neural activities within fROIs and fine-tuned to specific features of the visual input (**Figure 2**).

How does the tuning of these dimensions manifest within the cortical topographies of fROIs? To address this question, we projected the voxel weights onto cortical surfaces, enabling visualization of the spatial organization of the dimensions. This analysis revealed a mixture of spatially clustered and distributed dimensions (**Figure 3**) with patterns often consistent across participants, indicating a fine-grained heterogeneity within fROIs. Interestingly, our analysis also uncovered overlaps between fROIs, particularly between face- and body-selective areas (Taubert et al., 2022). This suggests a complex interplay of unique and shared neural patterns in the encoding of visual information.

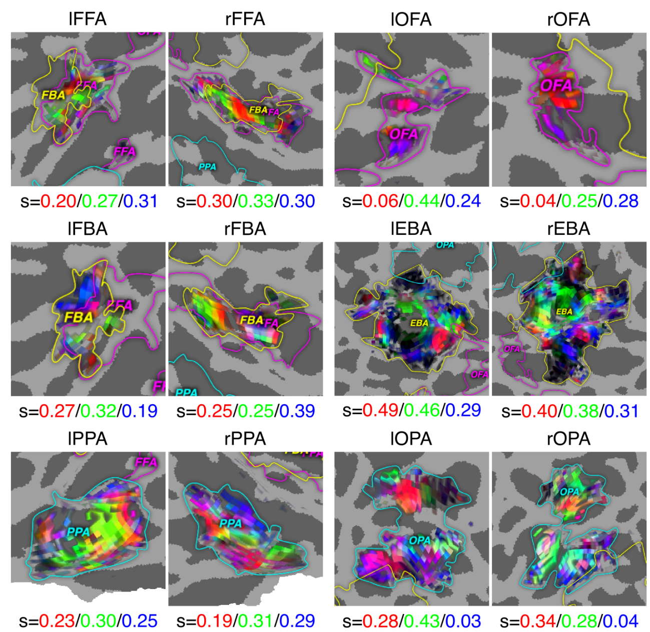

Figure 3: Cortical flatmaps colored by independently rescaled voxel weights, with sparseness of dimensions (Hoyer, 2004) ranging from perfectly distributed (s=0) to perfectly clustered (s=1) for each hemisphere.

Conclusion

Our data-driven dimensional approach reveals the inherent multidimensionality of functionally selective areas in the visual cortex. Neural representations within these areas are interpretable, multifaceted, and efficiently organized. This intricate interplay of neural dimensions likely plays a crucial role in how information is represented and processed in the brain.

Acknowledgements

This work was supported by the German Academic Scholarship Foundation (doctoral scholarship to L.v.D.), the European Research Council (ERC Starting Grant DEEPFUNC, ERC-2023-STG-101117441, to K.D.), the Hessian Ministry of Higher Education, Research, Science and the Arts (LOEWE Professorship to M.N.H. and Research Cluster 'The Adaptive Mind' to M.N.H. and K.D.), and the German Research Foundation (Collaborative Research Center SFB/TRR 135 to M.N.H. and K.D.).

References

Allen, E. J., St-Yves, G., Wu, Y., Breedlove, J. L., Prince, J. S., Dowdle, L. T., Nau, M., Caron, B., Pestilli, F., Charest, I., Hutchinson, J. B., Naselaris, T., & Kay, K. (2022). A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nature Neuroscience*, *25*(1), 116-126.

Bao, P., She, L., McGill, M., & Tsao, D. Y. (2020). A map of object space in primate inferotemporal cortex. *Nature*, *583*(7814), 103-108.

Bracci, S., Caramazza, A., & Peelen, M. V. (2015). Representational Similarity of Body Parts in Human Occipitotemporal Cortex. *Journal of Neuroscience*, *35*(38), 12977-12985.

Contier, O., Baker, C. I., & Hebart, M. N. (2023). *Distributed representations of behaviorally relevant object dimensions in the human visual system*. bioRxiv.

Çukur, T., Huth, A. G., Nishimoto, S., & Gallant, J. L. (2013). Functional Subdomains within Human FFA. *Journal of Neuroscience*, *33*(42), 16748-16766.

Çukur, T., Huth, A. G., Nishimoto, S., & Gallant, J. L. (2016). Functional Subdomains within Scene-Selective Cortex: Parahippocampal Place Area, Retrosplenial Complex, and Occipital Place Area. *Journal of Neuroscience*, *36*(40), 10257-10273.

Downing, P. E., Chan, A. W.-Y., Peelen, M. V., Dodds, C. M., & Kanwisher, N. (2005). Domain Specificity in Visual Cortex. *Cerebral Cortex*, *16*(10), 1453-1461.

Epstein, R. A., & Baker, C. I. (2019). Scene Perception in the Human Brain. *Annual Review of Vision Science*, *5*(1), 373–397.

Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, *15*(8), 536-548.

Hebart, M. N., Zheng, C. Y., Pereira, F., & Baker, C. I. (2020). Revealing the multidimensional mental representations of natural objects underlying human similarity judgements. *Nature Human Behavior*, *4*(11), 1173-1185.

Hoyer, P. O. (2004). *Non-negative matrix factorization with sparseness constraints*. arXiv.

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain. *Neuron*, *76*(6), 1210-1224.

Jain, N., Wang, A., Henderson, M. M., Lin, R., Prince, J. S., Tarr, M. J., & Wehbe, L. (2023). Selectivity for food in human ventral visual cortex. *Communications Biology*, *6*(1), 1-14.

Khosla, M., Murty, N. A. R., & Kanwisher, N. (2022). A highly selective response to food in human visual cortex revealed by hypothesis-free voxel decomposition. *Current Biology*, *32*(19), 4159-4171.e9.

Konkle, T., & Caramazza, A. (2013). Tripartite Organization of the Ventral Stream by Animacy and Object Size. *The Journal of Neuroscience*, *33*(25), 10235.

Schmidt, M. N., Winther, O., & Hansen, L. K. (2009). Bayesian Non-negative Matrix Factorization. In T. Adali, C. Jutten, J. M. T. Romano, & A. K. Barros (Eds.), *Independent Component Analysis and Signal Separation* (pp. 540–547). Springer.

Stigliani, A., Weiner, K. S., & Grill-Spector, K. (2015). Temporal Processing Capacity in High-Level Visual Cortex Is Domain Specific. *Journal of Neuroscience*, *35*(36), 12412-12424.

Taubert, J., Ritchie, J. B., Ungerleider, L. G., & Baker, C. I. (2022). One object, two networks? Assessing the relationship between the face and body-selective regions in the primate visual system. *Brain Structure and Function*, *227*(4), 1423-1438.

Weiner, K. S., & Grill-Spector, K. (2011). Not one extrastriate body area: Using anatomical landmarks, hMT+, and visual field maps to parcellate limb-selective activations in human lateral occipitotemporal cortex. *NeuroImage*, *56*(4), 2183–2199.