Unsupervised dimensionality reduction of fMRI reveals widely distributed and massively redundant representation of gender during face processing

Jarrod A. Lewis-Peacock, Timothy T. Rogers, and Bradley R. Postle
Department of Psychology, University of Wisconsin-Madison
lewispeacock@wisc.edu

1 Introduction

Human face identification is a complex, yet highly efficient, computation known to involve a network of brain regions along the ventral visual processing stream. Recent evidence using supervised multi-voxel pattern classification has suggested activity in the fusiform face area (FFA) corresponds to the detection of faces, while anterior portions of IT are engaged for identification (Kriegeskorte et al., 2007).

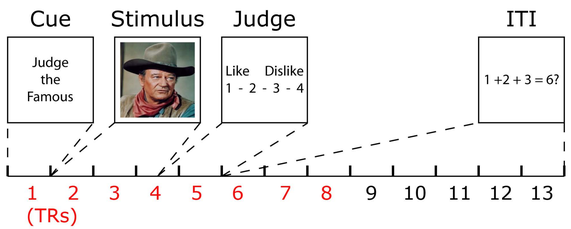
Supervised techniques force a pre-defined structure onto the data. If the representational structure in the data doesn't match the enforced structure, the analysis can fail. **Unsupervised** (e.g., clustering) methods, can extract structure in an unbiased manner, and so may better reflect the true underlying structures of neural representations. Here, we applied unsupervised dimensionality reduction to functional imaging data (3T fMRI) and distinguished cortical responses elicited by the presentation of pictures of famous men and women.

Our analyses recovered a widely distributed representational structure for gender. Region of interest analyses revealed that gender-specific representational structure is not limited to ventral visual processing areas. We discovered that this structure is recovered best by analyzing group-averaged BOLD signal rather than obtaining similarity measures from single-subject data sets, and then averaging them.

Together the results suggest a widely distributed and redundant neural code for gender that cannot be discovered by standard GLM or supervised multivariate techniques.

2 fMRI task design

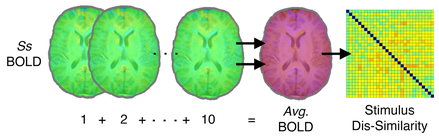
N=10 (7 men) from *Lewis-Peacock & Postle (2008)*



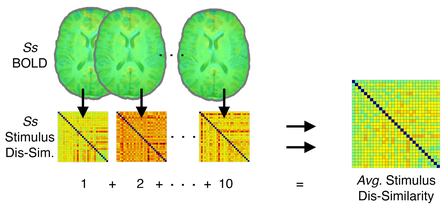
Stimuli: 30 famous faces (15 women / 15 men)

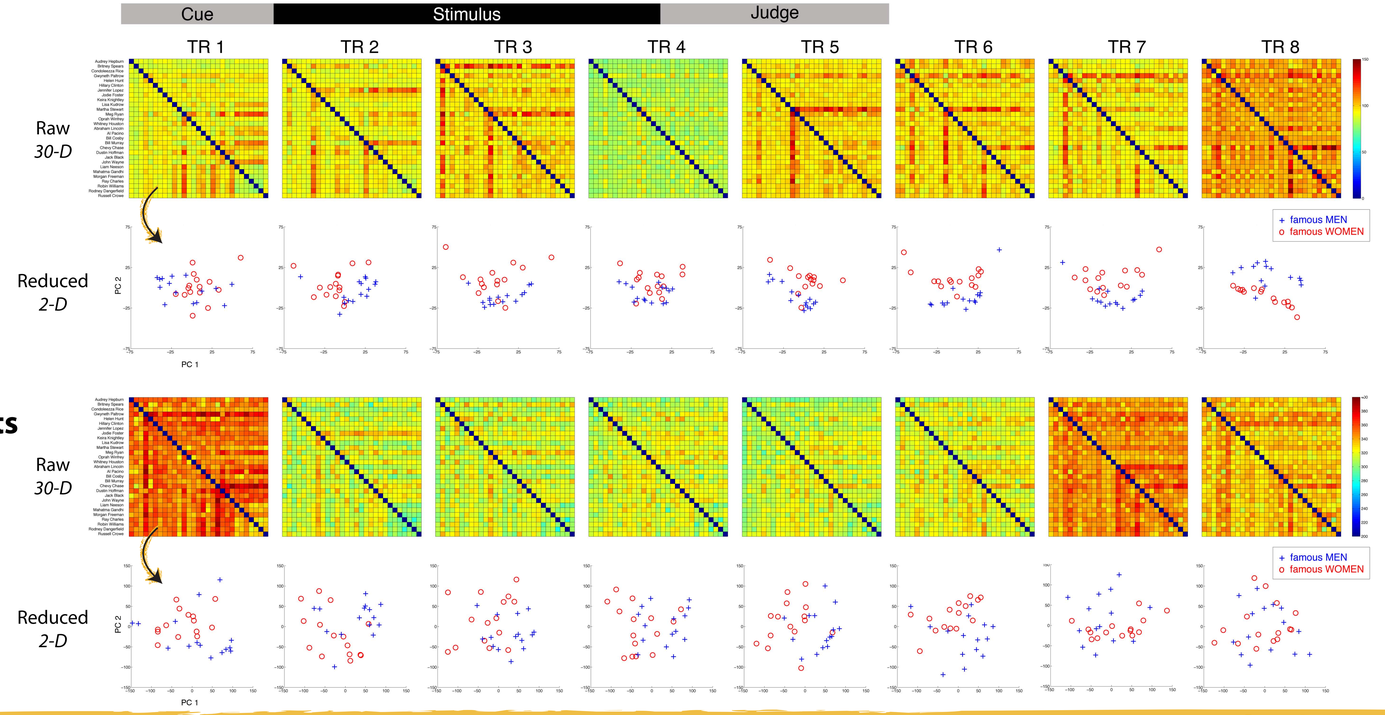
3 Dimensionality reduction of stimulus similarity matrix: avgBOLD vs. avgSimilarity

Avg. *BOLD signal* **across subjects**

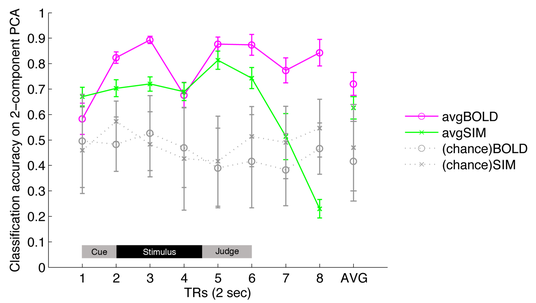


Avg. *Similarity matrix* **across subjects**





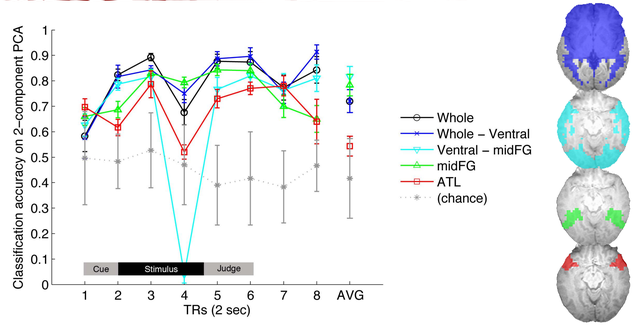
4 avgBOLD > avgSimilarity



Result - Multi-dimensional scaling of the similarity matrix created from the group-averaged BOLD signal -- rather than from the average of single-subject similarity analyses -- yielded the best separation by gender in the neural representations of faces.

This result suggests that there is shared representational structure for gender, coded in similar brain areas, across individuals.

5 Redundant representation

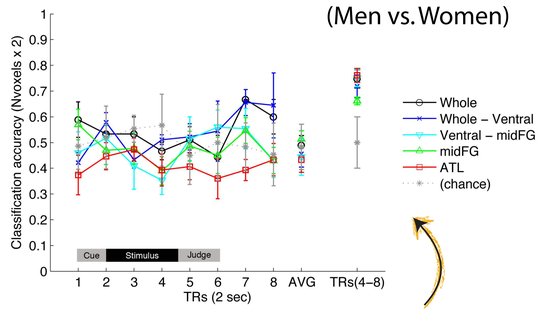


Result - Structured representation of gender was observed in brain areas outside of the ventral visual processing stream. E.g., when the ventral stream was masked out, the dimensionality reduction analysis extracted, with above-chance accuracy, a men/women gender distinction from the group-averaged BOLD data (see 'Whole - Ventral' time series above).

Interestingly, only at TR=4 did the middle fusiform gyrus (midFG) show the best gender dissociation. This suggests a distributed representation of gender that includes, but is not limited to the fusiform face area (FFA).

6 vs. Other approaches

Supervised pattern classification of avgBOLD (Men vs. Women)



Result - Pattern classification on voxels carrying the avgBOLD signal did NOT reliably perform above chance in leave-one-out cross-validation analyses, likely as a result of over-fitting the training data due to spurious correlations among the large number of voxels.

Result - Traditional univariate GLM analysis on group-averaged data revealed NO significant brain regions in the Men vs. Women contrast.

Society for Neuroscience Annual Meeting, Chicago, IL, October 2009

NIH R01 MH064498 (B.P.)