BOLD predictions: automated simulation of fMRI experiments

Leila Wehbe[1] **(lwehbe@berkeley.edu), Alexander G Huth**[2] **(huth@cs.utexas.edu),**

Fatma Deniz[1] **(fatma@berkeley.edu), James S Gao**[1] **(jamesgao@berkeley.edu),**

Marie-Luise Kieseler[3] **(mlkieseler@gmail.com) and Jack L Gallant**[1] **(gallant@berkeley..edu)**

1- Helen Wills Neuroscience Institute
University of California, Berkeley
2- Neuroscience and Computer Science Department
University of Texas at Austin
3- Department of Psychological and Brain Sciences
Dartmouth University

Abstract

In a typical fMRI experiment responses are recorded under a few conditions (e.g. abstract words and concrete words) and then contrasts are performed between conditions. Locations of significant differences are reported, usually in a table listing peak locations in standardized space. However, statistical thresholds are usually not directly comparable across experiments because of differences in design and analysis. Thus, it is difficult to replicate experiments or synthesize results across them. Naturalistic experiments and voxel-wise modeling provide one alternative to the contrast-based approach. These studies sample the stimulus space broadly and characterize the relationship between linearized stimulus features and brain activity in single voxels. Here, we provide a means to bridge between contrast-based studies and naturalistic studies. Specifically, we present a web-based replication engine that uses data derived from naturalistic voxel-wise modeling experiments to simulate any simple language contrast that can be expressed in terms of a list of words reflecting each of two conditions.

The replication engine is based on data from (Huth, de Heer, Griffiths, Theunissen, & Gallant, 2016). In that study seven subjects listened to hours of recorded narrative stories. Semantic features were extracted from the stories by projecting the stimulus words into a 985 word co-occurrence space computed over a large text corpus. Ridge regression was used to fit a linearized model to each voxel, predicting activity through time as a function of the semantic vectors. Our replication engine uses these voxel-wise models to simulate the results that would be expected to occur given a new contrast-based language experiment. A contrast is first defined between two conditions by providing two lists of stimulus words. Brain activity for each stimulus word is then predicted by multiplying its semantic vector with the semantic model weights for each subject, and averaging predicted activity across words for each condition. The contrast is simulated by computing the difference between predicted activities and a contrast map is obtained for each subject, and for an MNI-normalized average of the subjects. A non-parametric two-sample test is used to establish the significance of the difference in activity at each voxel. Finally, we produce a series of visualizations of the contrasts, including interactive Pycortex brain viewers (Gao, Huth, Lescroart, & Gallant, 2015). For contrast words drawn from previous studies that are included in our database, we compare the replicated results against the original published results.

The automated replication engine is available at `https://boldpredictions.gallantlab.org`. It can be used to (a) discover new networks representing semantic concepts, (b) to plan a new experiment by simulating the results for different stimuli and (c) to assess whether the results of a published contrast-based fMRI experiment generalize to the naturalistic context. The website includes several experiments published previously. Inspection of the published ROIs with our simulated replication results shows that the peaks of the simulated contrasts align with many of the reported ROIs. The website also shows which reported ROIs failed to be replicated.

The online replication engine can thus be used to simulate a replication of existing experiments, and to simulate any language contrast that can be tested in a contrast-based design. We invite the community to use the engine and contribute to the database of published studies that can be replicated.

Keywords: reproducibility, fMRI, language

Acknowledgments

Place acknowledgments (including funding information) in a section at the end of the paper.

References

Gao, J. S., Huth, A. G., Lescroart, M. D., & Gallant, J. L. (2015). Pycortex: an interactive surface visualizer for fmri. *Frontiers in neuroinformatics*, *9*, 23.

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453–458.