Supplementary methods and results for "Decoding the large-scale structure of brain function by classifying mental states across individuals"

Russell A. Poldrack, Yaroslav Halchenko, and Stephen José Hanson

March 4, 2009

1 Supplementary Methods

1.1 Participants

A total of 130 right-handed, healthy, English-speaking individuals participated in the studies included in this dataset. All participants were free of neurological and psychiatric history and gave informed consent to participate according to protocols approved by the University of California, Los Angeles Institutional Review Board.

1.2 fMRI acquisition

Imaging was performed using a 3T Siemens AG (Erlangen, Germany) Allegra MRI scanner at the UCLA Ahmanson-Lovelace Brain Mapping Center. For all subjects, whole-brain functional imaging was performed using a T2*-weighted echoplanar imaging (EPI) sequence with common parameters [repetition time (TR), 2 s; echo time (TE), 30 ms; flip angle, 90; matrix, 64 x 64; field of view (FOV), 200 mm; slice thickness, 4 mm; interleaved acquisition]. The only acquisition parameters that varied between subjects were the interslice interval, which was either 1mm or 0mm across different studies, and number of slices, which varied from 25 to 34 slices across studies. In addition to the functional images, a T2-weighted matched-bandwidth high-resolution anatomical scan (same slice prescription as EPI) and magnetization-prepared rapid-acquisition gradient echo (MPRAGE) were acquired for each subject for registration purposes.

1.3 Task designs

The design of each of the tasks is described below.

Task 1: Risk-taking (Stover et al., 2006). Subjects performed the Balloon Analog Risk Task (Lejuez et al., 2002), in which they decide on each trial whether to continue taking

risk (in this case, inflating a balloon) in order to potentially receive a greater reward. Decision-making trials on which subjects accepted further risk were compared to visual fixation baseline.

Task 2/2a: Probabilistic classification (Foerde et al., 2006). Subjects performed a probabilistic classification task, in which they classified items without feedback after having previously learned the task by trial and error. Classification trials were compared to a visual fixation baseline.

Task 3: Pseudoword rhyme judgment (Xue & Poldrack, 2007). Subjects judged whether pairs of visually presented pseudowords rhymed or not, in a blocked design. Rhyme judgment trials were compared to a visual fixation baseline

Task 4: Tone counting (Foerde et al., 2006). Subjects were presented with a series of high- and low-frequency tones, and asked to keep a running count of the number of high tones, with a probe at the end of each 36-second block. Tone-counting blocks were compared to blocks of rest.

Task 5: Mixed gamble decision making (Tom et al., 2007). Subjects performed a gamble acceptability judgment on mixed (gain/loss) gambles that varied in the amount of potential gain and loss. Gamble trials were compared to a visual fixation baseline.

Task 6: Semantic decision on mirror-reversed text (Cazalis et al., 2004). Subjects performed a living/nonliving decision on words presented in mirror-reversed text. Mirror-reading trials were compared to a visual fixation baseline.

Task 7: Reading pseudowords aloud (Xue et al., 2008). Subjects were presented with pronounceable pseudowords and asked to read them aloud. Reading trials were compared to a visual fixation baseline.

Task 8/8a: Successful response inhibition (Aron & Poldrack, 2006; Xue et al., 2008). Subjects performed the stop-signal task, which requires inhibition of a motor (finger) response when an infrequent auditory cue is presented. Stop trials on which subjects successfully inhibited the response were compared to a visual fixation baseline.

2 Analysis procedures

2.1 Preprocessing

Initial analysis was performed using the FSL toolbox from the Oxford Centre for fMRI of the Brain (www.fmrib.ox.ac.uk/fsl). The image timeseries was first realigned to compensate for small head movements using FSL MCFLIRT. Data were spatially smoothed using a 5 mm full-width-half-maximum Gaussian kernel. Registration was conducted through a 3-step procedure, whereby EPI images were first registered to the matched-bandwidth high-resolution structural image, then to the MPRAGE structural image, and finally into standard [Montreal Neurological Institute (MNI)] space (MNI avg152 template), using 12-parameter affine transformations. Statistical analyses were performed in native space, with the statistical maps normalized to standard space prior to higher-level analysis.

2.2 First-level analysis

Statistical modeling was performed separately for each imaging run. Regressors of interest were created by convolving a delta function representing trial onset times (for event-related studies) or a boxcar (for blocked-design studies) with a canonical (double-gamma) hemodynamic response function. Time-series statistical analysis was carried out using FILM (FMRIB's Improved Linear Model) with local autocorrelation correction after highpass temporal filtering (Gaussian-weighted LSF straight line fitting, with sigma=33.0s). For each study, contrasts were performed to compare the task condition to baseline, and Z (Gaussianized t) statistic images were created based on these contrasts. The mask for each individual (which was used to determine the mask for the classifier analyses) was determined by thresholding the brain-extracted image at 10% of its maximum value to exclude background voxels; the masking did not employ any information about task-related activation.

2.3 Classifier analysis

Classifier analyses were performed on the Z statistic images created for each subject in the first-level analysis. SVM was performed using libsvm (Chang & Lin, 2001), with a linear kernel and cost parameter C optimized by limited search. Voxels were selected by taking the intersection of the first-level masks in standard space, which resulted in a mask with a total of 214,940 voxels. Analysis of the entire dataset was performed using one-versus-one classification (all pairs, decision based on majority vote) with generalization accuracy determined using n-fold crossvalidation. The empirical null distribution for this analysis was obtained by performing the classification analysis 200 times with the task labels randomly permuted in each run. Mean chance performance across the 200 runs was 13.3%; the 95th percentile of the distribution of accuracy (18.5%) was taken as the $p < .05$ threshold.

2.4 ROI anlaysis

SVM was performed separately within each region of interest within the Harvard-Oxford Probabilistic Atlas, which is included in FSL version 4.0. which includes 48 cortical regions. The analysis was performed separately for left and right hemispheres, as well as combining the two hemispheres. The SVM analysis was performed as in the whole-brain analysis, using a linear kernel, with an optimized C value, and N-fold crossvalidation.

2.5 Localized SVM analysis

The local kernel classifier (fixed radius radial basis or "searchlight") analysis was performed at each voxel by extracting data from an 8 mm sphere around that voxel (including only

voxels that fall within the whole-brain mask). These data were submitted to 10-fold cross-validation using a linear multiclass (one-vs-one) SVM, and the crossvalidation accuracy value was assigned to the center voxel.

2.6 Diagnosticity analysis

We implemented a sensitivity/perturbation approach (Hanson & Halchenko, 2008; Hanson et al., 2004), which measures the error for a given TASK category (see Table 1). Specifically, to estimate SVM-based sensitivity, we used one of the simplest criteria proposed (Guyon et al., 2002; Ishak & Ghattas, 2005; Rakotomamonjy, 2003), which is the reciprocal of the separating margin width $W = 1/w_i$, where $w_i = \sum \alpha_i y_i x_i$. Rakotomamonjy (2003) has shown that these voxel weights W in a linear SVM are equivalent to the contribution of that voxel to the generalization performance altogether, as if it were removed. Minimization of this criterion leads to maximization of the margin width. In the case of linear SVM, the squared values of the separating plane normal coefficients (i.e., w_i^2), as stated, effectively correspond to the change of the criteria W as if the voxel i is removed. Therefore, the classifier is less sensitive to the features with low w_i^2. Additionally, in order to increase diagnostic selectivity, we derived weights for each TASK category by using only $TASK_i$ SVs. Thus, higher voxel value tends toward typical regions in the classification space for the SV appropriate TASK category. Given cross validation, SVM maps were produced for all possible classifiers, and aggregated by voting with the presence or absence of a voxel sensitivity per map, creating an aggregate map, thresholded at a vote value associated with $p < .01$.

2.7 Neural network analysis

Voxels were selected for the neural network (additive sigmoidal kernel approximator) analysis by computing the relative entropy ($-[p * log(p)]$ where p indicated relative frequency of a given voxel bandwidth) of each voxel over all brains and tasks and then thresholding them at $p < 1\%$. This selected a sparse set of 2173 voxels throughout the brain (see figure S4). These voxels were used as inputs to a neural network using a sigmoidal activation function, with 2173 input units, 8 logistic output units, and a number of logistic hidden units varying from zero (essentially equivalent to linear discriminant analysis or logistic regression) to a total of 14. The networks also generally possessed skip weights, allowing the network to extract low order statistics (e.g. 1st few eigenvectors) and was trained with a regularizer (weight decay) to punish high variance solutions. Training was done using standard BFGS nonlinear optimization over mulitple error criteria, including least squares and softmax.

2.8 Dimensional visualization

Cluster analysis was performed using an agglomerative hierarchical clustering algorithm with euclidian distance measure between vectors and centroid linkage group membership rule, implemented in the hclust function in R. Slice images of the dimensional loadings on anatomical images were obtained by smoothing the image with weightings for each voxel by an 8 mm Gaussian kernel.

2.9 Ontology projection

To obtain projections of the neural network dimensions onto the ontology terms, the dimensions were first averaged across subjects within each task. This 6 (dimension) X 8 (task) matrix was then multiplied by the 8 (task) X 22 (cognitive concept) matrix to obtain a matrix that provided the weights of each dimension on each concept. These were treated as distances and scaled to [0,1], and were then exponentiated to the 8th power to enhance the largest loadings. These values were then used to the create the tag cloud representations, using the online tool at http://TagCrowd.com.

3 Supplementary results

Supplementary Table 1. Generalization accuracy (via leave-one-out cross-validation) for a selected set of parameter combinations, as a function of model, regularization, loss function, and feature selection.

Mask	Model	Loss Function	Accuracy
Whole-brain (215K voxels)			
	SVM	Risk Minimization	80%
Entropy mask (2137 voxels)			
	SVM	Risk Minimization	72%
	NN (4 hidden)	Softmax	61%
	NN (6 hidden)	Softmax	71%
	NN (14 hidden)	Softmax	70%
	NN (6 hidden)	Least Squares	66%
	NN (14 hidden)	Least Squares	68%
Standard Deviation mask (6175 voxels)			
	LDA	Least Squares	44%
	NN (4 hidden)	Softmax	45%

Supplementary Table 2a. Confusion matrix for generalization across runs (classifier trained on first run and tested on second)

True	RISK	CLS	DEC	SEM	READ	INH	Accuracy
RISK	15	0	0	0	1	0	93.75%
CLS	1	18	0	0	1	0	90.00%
DEC	1	1	14	0	0	0	87.50%
SEM	1	1	0	11	1	0	78.57%
READ	0	0	0	0	18	1	94.74%
INH	1	0	0	0	0	14	93.33%

Supplementary Table 2b. Confusion matrix for generalization to new subjects on different versions of tasks 2 and 8 (data for those subjects on tasks 4 and 7, respectively, excluded from training

True	RISK	CLS	DEC	SEM	INH	Accuracy
CLS	2	13	1	2	0	72.22%
INH	0	1	0	0	19	95.00%

Supplementary Table 2c. Confusion matrix for generalization to new subjects on different versions of tasks 2 and 8 (data for those subjects on tasks 4 and 7, respectively, included in training set)

True	RISK	CLS	DEC	SEM	READ	INH	Marginal
CLS	2	13	1	2	0	0	72.22%
INH	0	1	0	0	7	12	60.00%

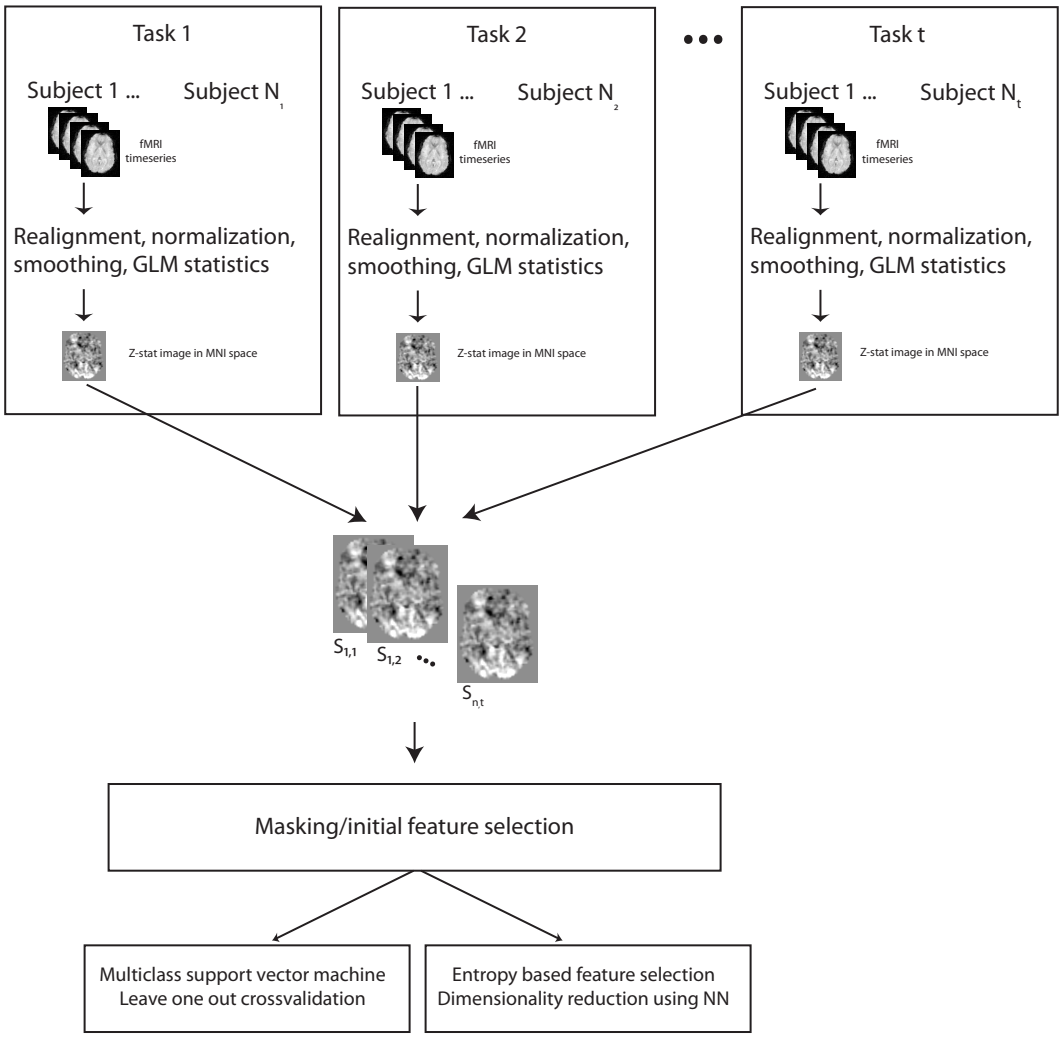
Supplementary Table 3. Results from anatomical ROI-based SVM analysis, sorted by decreasing classification accuracy. Anatomical regions based on Harvard-Oxford Probabilistic Atlas, which is distributed as part of FSL Version 4.0. B: bilateral, R: right hemisphere, L: left hemisphere.

Region	B accuracy	R accuracy	L accuracy
Precentral Gyrus	56.15	43.08	37.69
Occipital Pole	50.77	47.69	43.85
Lingual Gyrus	43.85	32.31	43.08
Lateral Occipital Cortex, superoir division	41.54	32.31	42.31
Occipital Fusiform Gyrus	40.77	31.54	36.15
Superior Frontal Gyrus	39.23	33.85	26.92
Postcentral Gyrus	37.69	30.77	31.54
Angular Gyrus	37.69	33.08	29.23
Precuneous Cortex	36.92	39.23	30.77
Lateral Occipital Cortex, inferior division	36.15	29.23	37.69
Temporal Fusiform Cortex, posterior division	36.15	28.46	30.00
Supramarginal Gyrus, posterior division	34.62	30.77	25.39
Intracalcarine Cortex	34.62	33.85	35.38
Frontal Orbital Cortex	34.62	32.31	31.54
Temporal Occipital Fusiform Cortex	34.62	32.31	28.46
Superior Temporal Gyrus, posterior division	33.85	33.08	23.85
Superior Temporal Gyrus, anterior division	33.08	22.31	29.23
Paracingulate Gyrus	33.08	28.46	23.85
Planum Temporale	33.08	30.77	23.08
Middle Temporal Gyrus, posterior division	32.31	22.31	37.69
Middle Temporal Gyrus, temporooccipital part	32.31	28.46	17.69
Cingulate Gyrus, posterior division	32.31	32.31	23.08
Cuneal Cortex	31.54	37.69	25.39
Inferior Temporal Gyrus, temporooccipital part	30.77	30.00	23.08
Heschl's Gyrus (includes H1 and H2)	30.77	25.39	20.77
Frontal Pole	30.00	36.92	21.54
Insular Cortex	30.00	26.92	26.15
Temporal Pole	30.00	21.54	23.85
Juxtapositional Lobule/Supplementary Motor Cortex	30.00	26.92	23.85
Parahippocampal Gyrus, posterior division	29.23	26.15	21.54
Supracalcarine Cortex	28.46	33.85	27.69
Middle Frontal Gyrus	27.69	23.08	23.85
Inferior Frontal Gyrus, pars triangularis	26.92	19.23	17.69
Superior Parietal Lobule	26.92	22.31	24.61

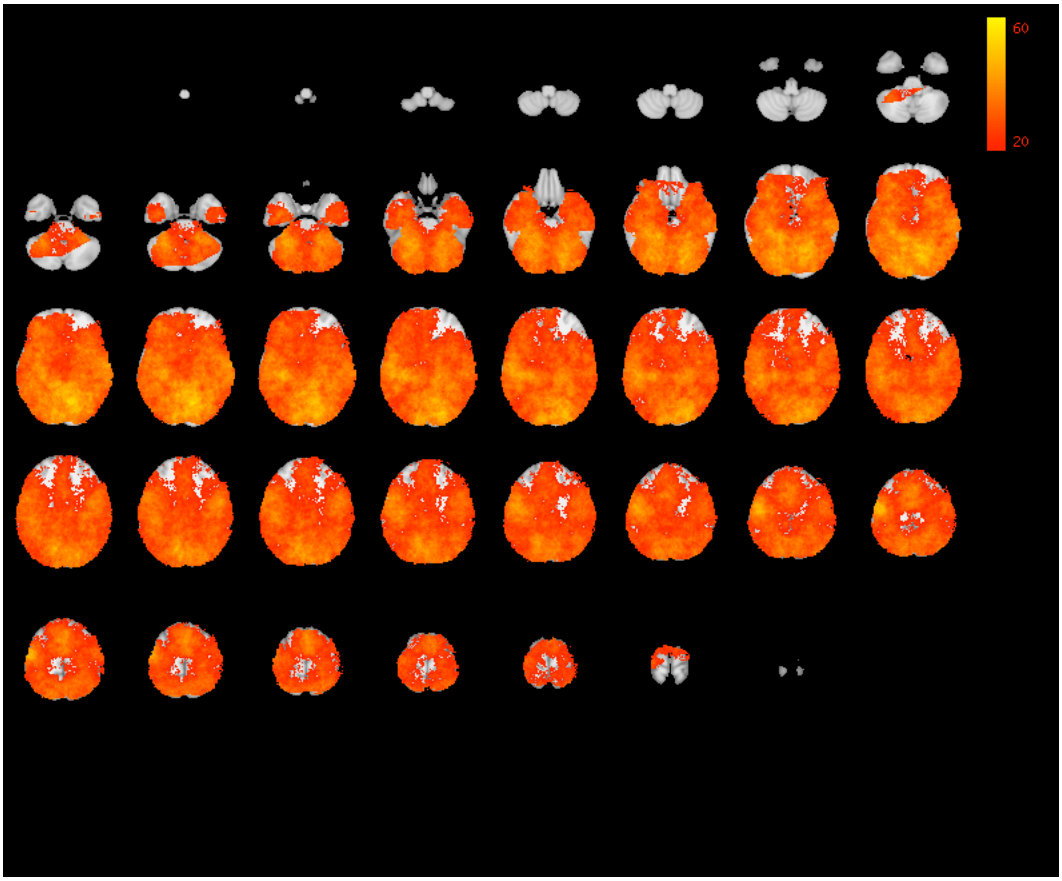
Continued on next page

Supplementary Table 3, continued from last page.

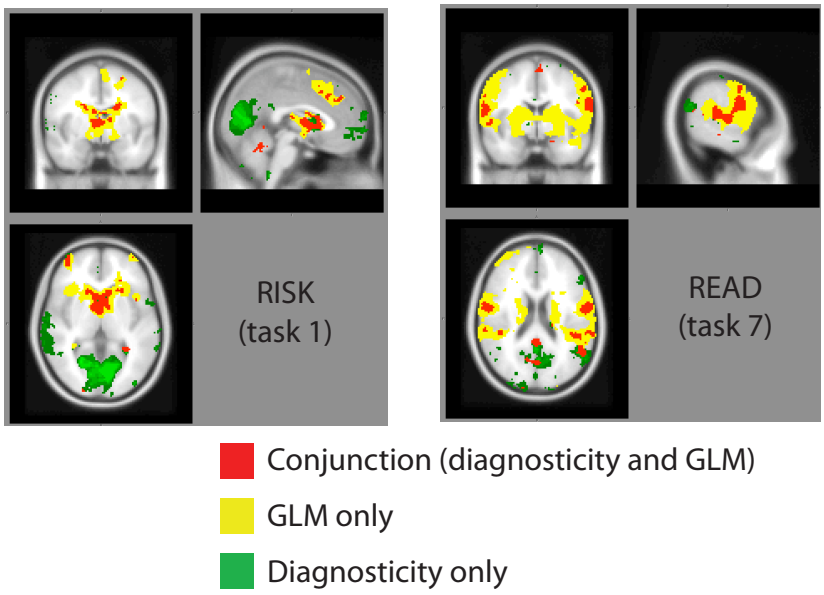
Region	B accuracy	R accuracy	L accuracy
Parietal Operculum Cortex	26.92	29.23	22.31
Central Opercular Cortex	26.15	25.39	24.61
Supramarginal Gyrus, anterior division	25.39	21.54	16.15
Cingulate Gyrus, anterior division	24.61	26.15	19.23
Planum Polare	23.85	21.54	25.39
Subcallosal Cortex	23.08	21.54	16.92
Inferior Frontal Gyrus, pars opercularis	21.54	24.61	23.08
Middle Temporal Gyrus, anterior division	21.54	17.69	23.08
Frontal Medial Cortex	20.77	23.08	23.08
Inferior Temporal Gyrus, posterior division	20.00	19.23	21.54
Frontal Operculum Cortex	16.92	20.77	14.62
Parahippocampal Gyrus, anterior division	15.38	21.54	15.38
Inferior Temporal Gyrus, anterior division	14.62	15.38	19.23
Temporal Fusiform Cortex, anterior division	13.85	12.31	15.38



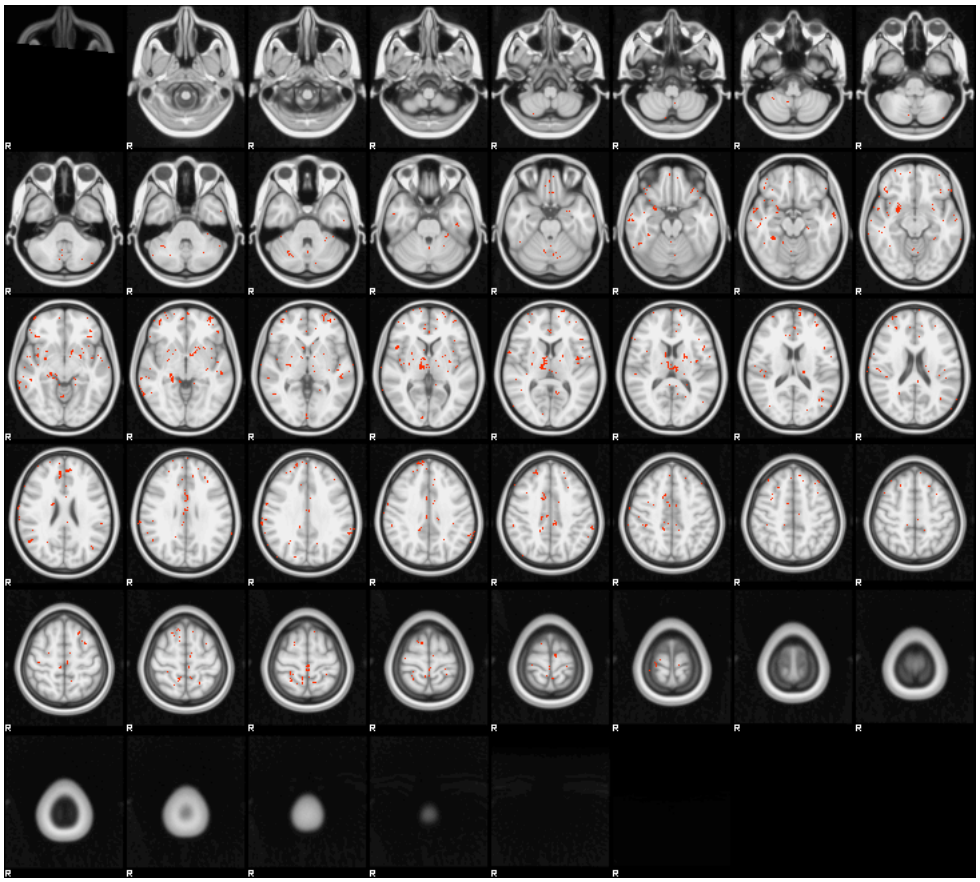
Supplementary Figure 1: Overview of data analysis pipeline



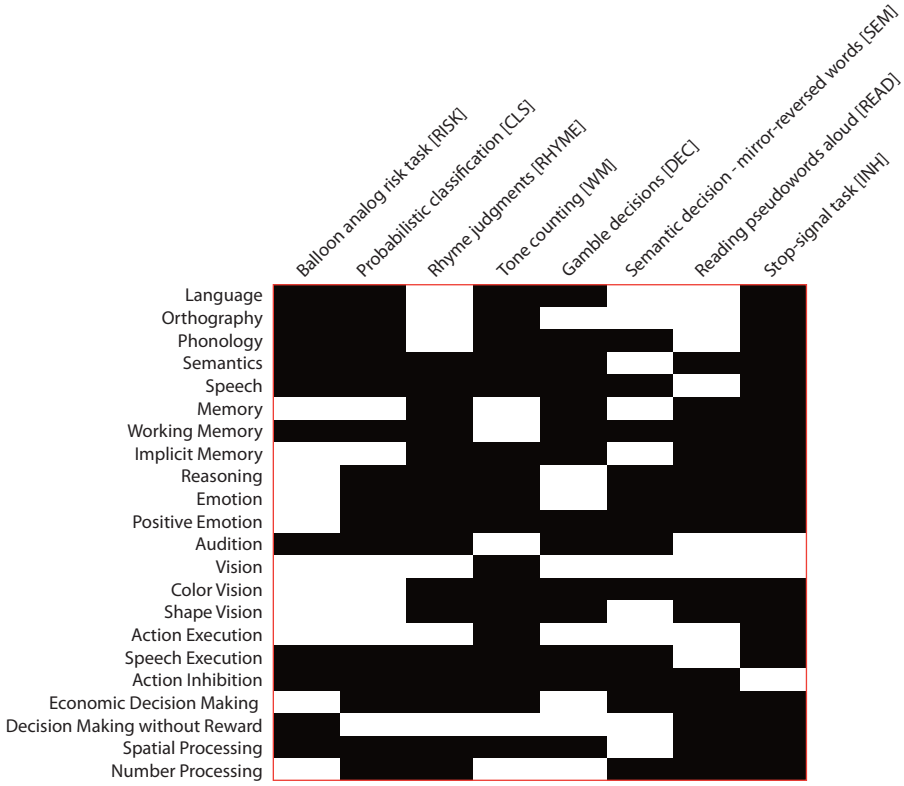
Supplementary Figure 2: Voxelwise classification accuracy from localized SVM analysis with 8 mm radius.



Supplementary Figure 3: Comparison of diagnosticity analysis to standard general linear model analysis for two tasks. Regions in red represent the conjunction of the diagnosticity analysis (2 or more votes for the specific task) and the GLM across all subjects in the relevant task (p¡.05 corrected using cluster-based randomization test).



Supplementary Figure 4: Features identified using entropy measure across the 130 brain/8 task data.



Supplementary Figure 5: Coding of each task in terms of the presence (white) or absence (black) of a set of cognitive concepts from a coarse cognitive ontology. The ontology was derived from the BrainMap.org behavioral domains framework, with a number of concepts added.

References

Aron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus. *J Neurosci*, *26*(9), 2424–2433.

Cazalis, F., Tom, S., Reger, M., Stover, E., Turner, K., & Poldrack, R. (2004). Event-related fmri study of mirror reading skill acquisition. In *Society for Neuroscience Abstracts*, (p. 369.363).

Chang, C.-C., & Lin, C.-J. (2001). Libsvm : a library for support vector machines. URL `http://www.csie.ntu.edu.tw/ cjlin/libsvm`

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proc Natl Acad Sci U S A*, *103*(31), 11778–11783.

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, *46*(1-3), 389–422. URL `http://citeseer.ist.psu.edu/guyon02gene.html`

Hanson, S. J., & Halchenko, Y. O. (2008). Brain reading using full brain support vector machines for object recognition: there is no "face" identification area. *Neural Comput*, *20*(2), 486–503.

Hanson, S. J., Matsuka, T., & Haxby, J. V. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? *Neuroimage*, *23*(1), 156–166.

Ishak, B., & Ghattas, B. (2005). An efficient method for variable selection using svm based criteria. *Journal of Machine Learning Research*, *6*, 1357–1370.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the balloon analogue risk task (bart). *J Exp Psychol Appl*, *8*(2), 75–84.

Rakotomamonjy, A. (2003). Variable selection using svm-based criteria. *Journal of Machine Learning Research*, *3*, 1357–1370.

Stover, E., Fox, C., Trepel, C., & Poldrack, R. (2006). The neural correlates of decision making under risk: an fmri study. In *Organization for Human Brain Mapping Abstracts*, (p. 379).

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*(5811), 515–518.

Xue, G., Aron, A. R., & Poldrack, R. A. (2008). Common neural substrates for inhibition of spoken and manual responses. *Cereb Cortex*, *18*(8), 1923–1932.

Xue, G., & Poldrack, R. A. (2007). The neural substrates of visual perceptual learning of words: implications for the visual word form area hypothesis. *J Cogn Neurosci*, *19*(10), 1643–1655.