Task-free MRI Predicts Individual Differences in Brain Activity During Task Performance

I Tavor[1,2], O Parker Jones[1], RB Mars[1,3], SM Smith[1], TE Behrens[1,4], S Jbabdi[1*]

[1]Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, Oxford OX3 9DU, UK.

[2]Department of Diagnostic Imaging, Sheba Medical Center, Tel Hashomer, 52621, Israel.

[3]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, 6525 EZ Nijmegen, The Netherlands.

[4]Wellcome Trust Centre for Neuroimaging, University College London, London, WC1N 3BG, UK.

*Correspondence to: Saad Jbabdi; email: saad@fmrib.ox.ac.uk; tel: 44 1865 222466, fax: 44 1865 222717, address: FMRIB Centre, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK

When asked to perform the same task, different individuals exhibit markedly different patterns of brain activity. This variability is often attributed to volatile factors such as task strategy or compliance. We propose that individual differences in brain responses are, to a large degree, *inherent* to the brain, and can be predicted from *task-independent* measurements collected at rest. Using a large set of task conditions, spanning several behavioral domains, we train a simple model that relates task-independent measurements to task activity and evaluate the model by predicting task activation maps for unseen subjects. Our model can accurately predict individual differences in brain activity highlighting a coupling between brain connectivity and function that can be captured at the level of individual subjects.

One Sentence Summary: Resting-state connectivity predicts individual task activation maps across several behavioral domains.

We all differ in how we perceive, think, and act. Our brains also differ in how they solve tasks. Understanding these individual differences in brain activity is an important goal in neuroscience, as it provides a route for linking brain and behavior.

Often individual differences in brain activity induced by an experimental task are attributed to two possible factors. Firstly, they may be accounted for by differences in gross brain morphology. The vast majority of brain imaging studies rely on the spatial alignment of different brains (registration) to account for between-subject discrepancies in gross anatomy (*1*). Secondly, subjects may use different strategies or cognitive processes that involve different brain circuits. Psychologists design their tasks with great care to limit this source of variability. Nevertheless individual variations are seen in all behavioral domains (*2-6*). These between-subject differences are often treated as 'noise' in imaging studies and discarded through the process of averaging individual responses (*7*). Importantly, it is thought that such individual differences are mainly explained by volatile factors related to the behavior.

We investigated the possibility that individual differences in brain activation are *inherent* features of individuals, and to a large degree independent of volatile factors. We explored the extent to which individual differences in task-evoked brain activity can be predicted by differences in the functional connectivity of the brain, acquired in an MRI scanner while the subjects are at rest, and not performing *any explicit task*. We aimed to predict *several* task-evoked activity maps matching individual subjects' maps in multiple behavioral domains based on a single task-free scan (i.e., unconstrained cognition during rest, with no explicit experimental task) of any given subject.

We designed a set of regression-based models that use *task-independent* features to predict individual *task-evoked* responses. Based on the hypothesis that functional differentiation in the brain can be understood in terms of the underlying long-range brain connections and interactions (*8, 9*), we used predictors based on functional connectivity at rest (*10*). Brain networks, extracted from resting-state data sets, qualitatively resemble task-evoked networks at the group level (*11*). We therefore hypothesized that using functional connectivity at rest we could predict individual variations in task responses. We also used predictors encoding individual brain morphology (gross structure) and microstructure to yield a total of 107 predictors (see Materials and Methods). All of our predictors were based on imaging subjects at rest, and were independent of any given task. The model was trained to map between the predictors and task activations in a cohort of subjects for each task, and subsequently the trained model was applied to out-of-sample (unseen) subjects to predict their task activations (leave-one-out approach, see Materials and Methods for details). Throughout, we focused on predicting task activations on the cortex, though the approach can easily be extended to incorporate subcortical gray matter.

Our data were a subset (98 subjects) of the Human Connectome Project (HCP) database (*12*). HCP data were chosen for their inclusion of resting-state measurements, diffusion-weighted MRI, structural MRI as well as task-evoked data spanning several behavioral domains. We could therefore use the same set of task-free data to test predictions of several different tasks. The HCP task data include 7 behavioral domains, and each task set comprises several statistical maps pertaining to different aspects of each task (all tasks and derived parametric maps are described in (*13*)). A total of 47 independent task maps (z-scores) were available for each subject (see Materials and Methods).

We aimed to predict not how subjects activate on average in a given task, but how they *differ* (from each other) in their activation patterns. Figure 1 illustrates this aspect of the model in four different task conditions (maps thresholded as described in Materials and Methods). The model could predict qualitative differences between subjects, in terms of shape, position, size and topography of activations. Figures S1-S3 and supplementary movies SM1-SM6 show similar individual variations predicted across more subjects and all behavioral domains used by the HCP. The model could precisely predict individual task variations across all behavioral domains, based on a single task-free dataset. To get an impression of how similar our predictions are to the actual task maps, we overlapped the predictions for three subjects with their actual activation maps (Fig. 2A) for the contrast MATH-STORY from the LANGUAGE task. For comparison, we also overlapped the same subjects with the actual activation map of a canonical (median) subject. The maximum overlap was obtained when comparing a given subject's prediction to their actual map. This was despite our use of a leave-one-out approach: the predictive model for subject X had not seen the activation map of subject X during training, but it had seen the maps of all other subjects. Nevertheless, the prediction was more similar to the map it had not seen (subject X) than to the maps it had seen (all but subject X). The model did not learn what the activations for a given task looked like, but rather it learned how to *map* from the features (resting-state connectivity and morphology) to the task maps in individual subjects.

To quantitatively assess the performance of the model, we estimated the spatial correlation between the (unthresholded) predictions and actual maps for all pairs of subjects (Fig. 2B and 2C). Each entry in the matrix is the Pearson correlation between the task map of one subject and the predicted map of another (off-diagonal) or the same (on-diagonal) subject. The correlation matrix is noticeably diagonal-dominant, indicating that on average the model prediction for any given subject is more similar to the subject's own map than to other subjects' maps (see also the histograms shown in figures 2B-C comparing on-diagonal to off-diagonal entries). Normalising rows and columns of the correlation matrix (which removes the overall mean correlation and accounts for the fact that the actual maps are more variable than the predicted maps) shows the diagonal-dominance even more clearly (see also figures S4-6 in which we show the same results for all contrast maps). The diagonal-dominance is apparent for all tasks and contrast maps, with the exception of one contrast map (GAMBLING PUNISH-REWARD). The reason is that activations for this contrast are restricted to sub-cortical gray matter, whereas our model only makes predictions for the cortex. This result, i.e., the diagonal dominance of the spatial correlation matrix, is non-trivial considering that our leave-one-out procedure ensures that whenever a model is applied to predict a given subject, it has never seen that subject's task map during training. Yet, the prediction matches that subject's task data better than the subjects that it has seen during training. The model's ability to generalize beyond the training subjects has important implications. It can predict activations in individuals for which there are no available task data (e.g. patients who cannot perform the task).

The model can precisely capture inter-individual variability (Table S3). There are different types of such variations. Functional brain areas in individual subjects can greatly vary in size, location, and shape (*14*). Primary visual cortex, for instance, can vary (across different subjects) by a factor of two or more in surface area (*15*). Such variability still allows one-to-one mapping of brain areas across subjects, provided that individual areas can be mapped in individual subjects. But another type of variation is topological variability, which for example means that different

subjects may have different patterns of activity with no obvious one-to-one correspondence between subjects. Both types of variability (shape/size vs topological) can be captured by this model (Fig. 3A). Importantly, topological variability cannot be accounted for using spatial alignment between subjects, and thus comparing or averaging subjects' activations when they have different topologies is an open problem. The fact that we can predict such variability from task-free data may enable new strategies for matching brain networks, as opposed to brain areas, across subjects.

Our model can also make predictions on the distribution of activity across different systems, such as in the case of hemispheric lateralization. Lateralization of function is seen commonly in the domain of language (*16*), but also in attentional processing (*17*). Figure 3B shows how our model can predict such differences across individuals in a language task. We estimated a lateralization index defined as the difference between right hemisphere and left hemisphere peak Z (averaged over a 10 mm radius sphere around the predicted peak) from the predictions and the actual data. The model is able to precisely predict individual subjects' lateralization index. We also predicted two different contrasts from the same task (MATH-STORY and STORY-MATH) where the former tends to show an approximately equal distribution among left and right lateralized individuals, whereas the latter shows more left lateralized subjects. The model can capture the lateralized language trend, even for the minority of right-lateralized subjects, despite the fact that, for the STORY-MATH contrast, it has been trained on subjects the majority of whom are left-lateralized.

The model could also predict atypical activation patterns. In Fig. 4, we show in five different behavioral domains examples of subjects that either do or do not conform to the pattern seen in the average subject. The model predicted activations in regions that were not activated on average. Conversely the model correctly predicted the *lack* of activations in regions that were activated on average.

Overall, a simple set of models trained to learn a mapping between task-free and task-evoked maps could be used to predict individual differences in activation maps accurately across a wide range of behavioral domains. The predicted maps matched even large individual variations in the spatial layout of brain activity, including position, shape, size and topology of functional activations (e.g. whether the activity is localised or spread-out, or whether it consists of a single region or multiple sub-regions). The model could also predict individual differences in the *amount* of activation in a given task from a task-free data set.

The model mainly used resting-state connectivity in addition to a few structural features capturing local morphology and micro-structure (see Materials and Methods). On average across all tasks, all features participated in the predictions except for sub-cortical connectivity features (excluding the cerebellum - see Figs. S7 and S8). Although the structural features were exploited by the model (Fig. S7), removing them did not affect the performance of the model (Fig. S9), and using a model purely based on structural features abolishes inter-individual variability in the predictions (Fig S10). This suggests that resting-state connectivity alone is sufficient to predict individual variability in task maps, independently from variations of morphology as captured by our handful of structural features. It is however possible that additional anatomical variability may also be indirectly captured by the resting-state data.

Why can we predict task-evoked activity from resting-state data? A good match between resting state networks and task networks at the group level has been shown before (*11, 18*), where it has

been argued that functional networks are continuously interacting with each other at rest, with the same functional hierarchy that is seen during action and cognition. Our model provides an explicit mapping from resting-state data to task maps that corroborates this assumption, but it is important to consider alternative explanations of our result. Inter-subject alignment may be sub-optimal when using anatomy alone. Our model is therefore possibly accounting for functional misalignment between subjects (*19*). However, the model not only tracks variation in the position and shape of functional areas, but also captures variations in the topology of the brain networks activated in individual subjects, and can predict activations of atypical subjects. Therefore, it is unlikely that the standard approach of aligning subjects while preserving the brain topology will be able to fix these types of misalignments. Our model can make quantitative predictions, in terms of the amount of activity that individual subjects display during task. Such predictions are unlikely to arise from misalignment considerations. Resting-state data may provide a means for "calibration" of the BOLD signal (*20*). However, our predictors are based on measures of connectivity rather than raw BOLD signal strength per se (although signal-to-noise may still affect functional connectivity at rest (*21*)). Determining what information in the resting-state signal is driving our predictions will be important for understanding its nature and potential. Because all the predictors that were used in this study are based on scans acquired at rest, our model is blind to different strategies that are chosen by the participants in performing a given task. We refer to these features as "inherent", but we acknowledge that these can be "structurally inherent" (related to brain organization and connectivity) or "functionally inherent" (related to the cognitive state of subjects during the resting state scan).

The idea that brain connectivity can predict activation has previously been reported for different modalities (*22*), where diffusion MRI tractography was used (*23*) to measure connectivity. This study was limited to a specific cognitive task and a pre-defined anatomical region. More recently, resting-state connectivity has been shown to be predictive of subjects' identity, in a similar way to a fingerprint (*24*). Rather than simply identifying subjects, our goal was to predict the entire layout of brain activity for each subject. Moreover, we also aim to predict such layout of activity in a number of different cognitive domains, from a single task-free scan, including in subjects that show patterns of activation that are different from the group average (perhaps most strikingly in right-lateralized subjects when the majority of training subjects are left-lateralized).

There are significant practical implications to the proposed framework in basic research and translational neuroscience. It provides a method for inferring *multiple individualized* functional localizers based on a *single* resting state scan. Such a tool could be used to investigate in detail the response profiles of localized brain regions without the need to acquire often time-consuming task localizers. Importantly, such a tool, if generalizable beyond the young, healthy population that makes up the HCP database, could be used to investigate functional regions in subjects who cannot perform tasks, such as paralyzed patients or infants.

Task-free MRI Predicts Individual Differences in Brain Activity During Task Performance

I Tavor[1,2], O Parker Jones[1], RB Mars[1,3], SM Smith[1], TE Behrens[1,4], S Jbabdi[1*]

Supplementary Materials:

Materials and Methods

Human Connectome Project data

We used the minimally pre-processed data (*25*) provided by the Human Connectome Project (www.humanconnectome.org). The dataset includes functional and structural scans (September 2013 "Q3" release). All acquisition parameters and processing pipelines are described in detail in (*13, 25-28*).

Briefly, resting-state and task-based functional MRI scans were acquired using simultaneous multi-slice (or multi-band) acceleration (*29*), with an acceleration factor of 8 and image resolution of 2.0mm isotropic. Each subject had four 15-minute resting FMRI runs, with a TR of 0.72s giving 1,200 time points per run. The functional data were FIX-cleaned (*30, 31*) to automatically remove the effect of structured artefacts and then resampled onto the set of 91,282 "grayordinates" (*25*) in standard space, which use a surface representation of the data on the two cortical hemispheres and voxel representation for sub-cortical grey matter and the cerebellum. Subjects' brains were aligned on the basis of their cortical folding as described in (*25*).

Task FMRI data included 86 contrasts from 7 task domains, which are labeled here and in the HCP data base as: WM (working memory), GAMBLING, MOTOR, LANGUAGE, SOCIAL RELATIONAL, and EMOTION. The tasks are described in detail in (*13*). As most of the originally released contrasts are paired with a related negative contrast which contains the same information (for the purpose of our modelling), we extracted 47 unique contrast maps by excluding all redundant contrasts. The contrast maps that we kept are described in table S1. Figure S11 shows correlations between the 47 contrast maps used in this study.

Structural T1 and T2 weighted data were also acquired (0.7mm isotropic voxels) and processed as described in (*25*). These structural data were primarily used to extract cortical surface and associated features (see below). They were also used to align subjects to standard space using surface-based alignment. Finally, we used the diffusion MRI data that are distributed by the HCP (*27*). These consist of 3 shells (b-values=1000, 2000, and 3000 s/mm^2) with 270 diffusion directions equally spread amongst the shells, and six b=0 acquisitions within each shell, with a spatial resolution of 1.25mm isotropic voxels.

In total, a group of 98 subjects were used in all our analyses (i.e., about half of the subjects with all modalities available at the time of analysis). We excluded half of all subjects at random to allow independent researchers to replicate our results. A list of the HCP subjects used in this article is provided in table S2.

Feature extraction

Resting-state FMRI data were processed using group-PCA (*32*) followed by group-ICA (for cortex) and separate parcellations (for sub-cortical structures). The purpose of the group ICA was to define a set of brain maps to serve as "seeds" for creating connectivity maps. The connectivity maps in turn served as features for our predictions. First, group ICA with a fixed number of 40 parcels was performed separately on each hemisphere using fast ICA (*33*). Then spatial maps were compared across hemispheres to determine which maps had a symmetric counterpart in the other hemisphere. A total of 33 cortical maps per hemisphere were retained through this procedure. For the sub-cortical structures, we used customized parcellation techniques. Whole-brain ICA tends to favor cortical regions because of the lower signal to noise in the sub-cortex. To parcellate the sub-cortical structures, we used the following: 1 cluster per hemisphere (Nucleus Accumbens, Pallidum); spectral clustering with 2 clusters per hemisphere (Caudate nucleus, Hippocampus, Putamen, Amygdala); and ICA with 3 clusters per hemisphere (Cerebellum and Thalamus).

We used dual regression (*34*) against individual time series to produce subject-specific cortical ICA maps. The first (multiple) regression uses the cortical group maps described above as regressors to get individual time series for each subject and each map. The second regression uses the individual time series as regressors to get individual spatial maps (Fig S14). The individual resting-state time-series data are demeaned and variance-normalized (*35*), then concatenated across the 4 runs prior to dual regression.

Following dual regression, the resulting individual maps were regressed against individual time series to obtain one time course per spatial map. These time courses were then correlated against the individual data (equivalent to single regression) to produce individual subjects' connectivity maps for each map and each subject. In total, this procedure produced 33 connectivity maps per hemisphere for the cortical seed maps, and 32 connectivity maps from the subcortical seeds.

Structural features included cortical curvature, sulcal depth, and cortical thickness. Micro-structural features included "myelin maps" (*36*) and the following DTI-based indices: fractional anisotropy (FA); mean diffusivity for each of three b-values (MD1, MD2, MD3); and a map of the dot-product between the principal diffusion direction and the cortical surface at the interface between grey and white matter. (This last index represents the relationship between white matter fibre orientations and cortical folds.) In total we produced 107 features (98 functional connectivity features and 9 structural features). Figure S12 shows example features in a single, representative subject. Figure S13 reports correlations between all features used for our predictions (averaged across subjects).

Note that we opted not to include features derived using diffusion MRI tractography. Although anatomical connectivity would be a natural addition to this model, we did not include tractography-derived features because of the so-called *gyral bias (37, 38)*: estimates of anatomical connectivity tend to be stronger on gyral crowns than gyral walls, and they tend to be even weaker towards sulcal fundi. This is primarily due to the spatial resolution of in vivo diffusion MRI relative to the length-scale at which fibres branch out from main pathways towards the cortex (*38*). Since all our modelling and analyses are performed on the cortical sheet, we believe that a model with features that contain a strong gyral bias would be poorly informative on individual variability in functional localisation, and instead be highly correlated with variability in gyrification (which are already accounted for with some of the structural features).

Model

We used a piece-wise linear model to map the resting-state and structural features to the task data. Here, "task data" refers to the z-score maps derived from the contrasts described in the "Human Connectome Project data" section. In this article, we focused on predicting task data on the *cortex*, which is where we observe the majority of the inter-subject variability. However the model could easily be extended to predict sub-cortical activations. The task data were modelled using the multiple regression (Fig S15), broken down spatially into 50 non-overlapping regions of interest created using group ICA and a winner-takes-all parcellation on the resulting ICA maps. Within each of these 50 parcels, a general linear model was used to fit features to activation data in each subject. Prior to learning the model parameters, all features were normalized to zero mean and unit norm across the whole brain. Feature normalization is important both to average regression results across subjects and to compare regression coefficients between features. A constant regressor was added and the model fitted for each of the 50 parcels. To cover the whole cortex, we thus fitted 50 models for each subject (Figure S18-19 show how the prediction accuracy changes with the number of parcels or parcellation method). To predict an individual's task activation map, regression coefficients were averaged across the 97 other subjects (Fig S15). This process was repeated for all 98 subjects to produce predictions for individual task data and for each of 47 contrast maps. Since our model uses a parcellation of the cortex, all our predictions were restricted to the cortex. However, note that the features used for the predictions include connectivity maps between the cortex and sub-cortical structures (32 features).

During training, our model learns the regression betas for each subject and each brain parcel separately, then these betas are averaged across n-1 subjects for out-of-sample predictions. To demonstrate the stability of the betas we performed a split-half reproducibility analysis, by randomly splitting the group of 98 subjects into two halves and plotting one set of betas against the other. Fig. S16 shows split-half reproducibility per task contrast, but pooled over brain parcels and subjects. Additionally, we also compared split-half reproducibility (correlation coefficient between the two split halves for each task) to the performance of the out-of-sample test. Fig. S17 shows that reproducibility is higher when model accuracy is higher.

Thresholding

In all figures, whenever a thresholded map is shown, the threshold values are based on a Gaussian-two-Gammas mixture model (*35*), where the Gaussian represents the noise and the two Gamma distributions represent positive and negative activations. The positive and negative thresholds are chosen to correspond to the medians of the two Gamma distributions.

Table S1. Nomenclature for all 7 tasks and 47 contrast maps. The same nomenclature is used in the HCP data base, and all the tasks and contrasts are described in great detail in (*13*).

task	behavioral domains/attributes	contrasts
EMOTION	Valance judgments (faces), shape recognition	[1] FACES [2] SHAPES [3] FACES-SHAPES

GAMBLING	Reward, punishment, decision making	[4] PUNISH [5] REWARD [6] PUNISH-REWARD
LANGUAGE	Sentences, stories, mental arithmetic (auditory)	[7] MATH [8] STORY [9] MATH-STORY
MOTOR	Hand, foot, tongue movements	[10] CUE [11] LF [12] LH [13] RF [14] RH [15] T [16] AVG [17] CUE-AVG [18] LF-AVG [19] LH-AVG [20] RF-AVG [21] RH-AVG [22] T-AVG
RELATIONAL	Higher-order cognition	[23] MATCH [24] REL [25] MATCH-REL
SOCIAL	Interpret social vs random interactions	[26] RANDOM [27] TOM [28] RANDOM-TOM
WM	N-back working memory, body parts, tools, places	[29] 2BK_BODY [30] 2BK_FACE [31] 2BK_PLACE [32] 2BK_TOOL [33] 0BK_BODY [34] 0BK_FACE [35] 0BK_PLACE [36] 0BK_TOOL [37] 2BK [38] 0BK [39] 2BK-0BK [40] BODY [41] FACE [42] PLACE [43] TOOL [44] BODY-AVG [45] FACE-AVG [46] PLACE-AVG [47] TOOL-AVG

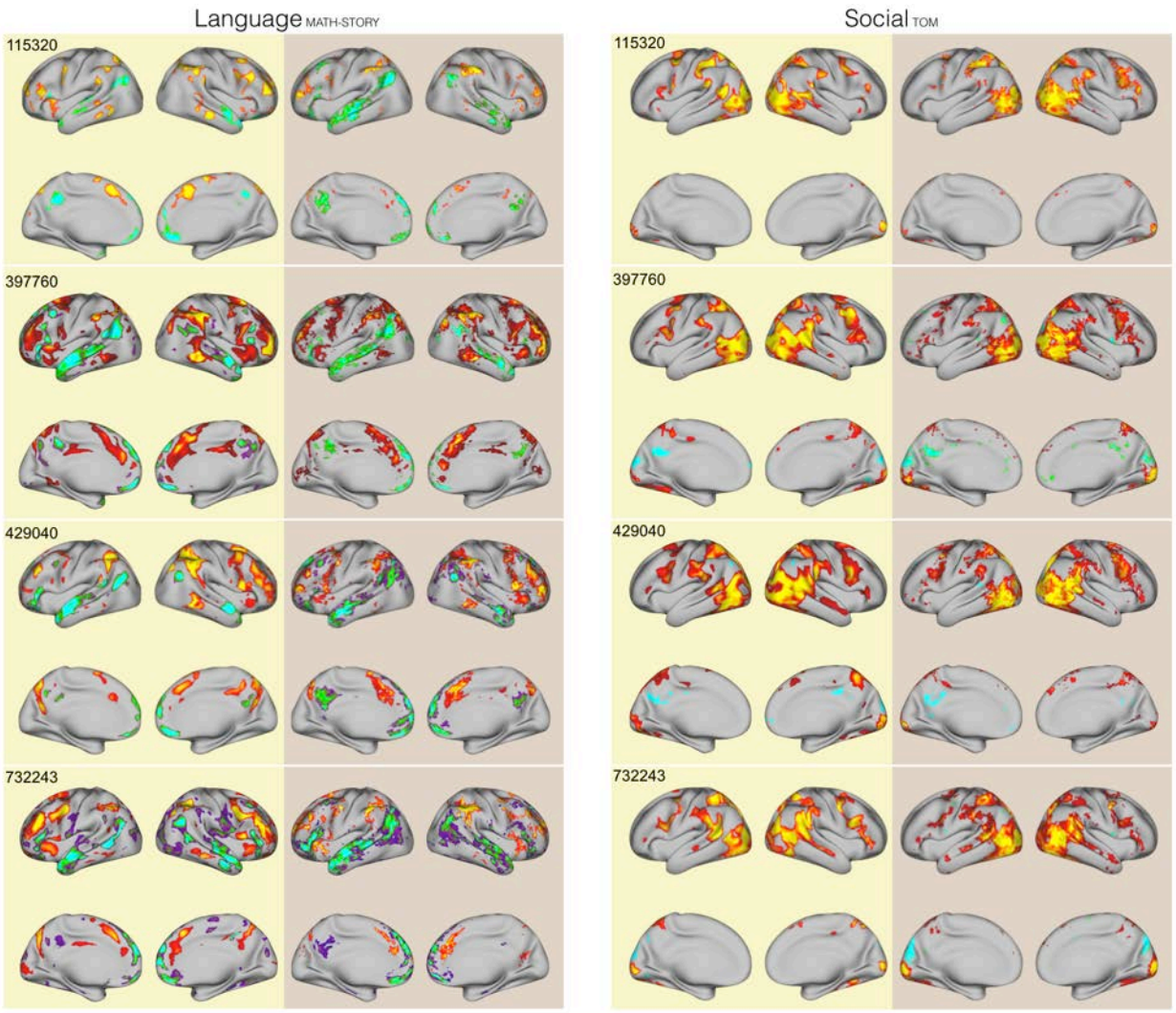
Table S2. List of HCP subjects used in this study.

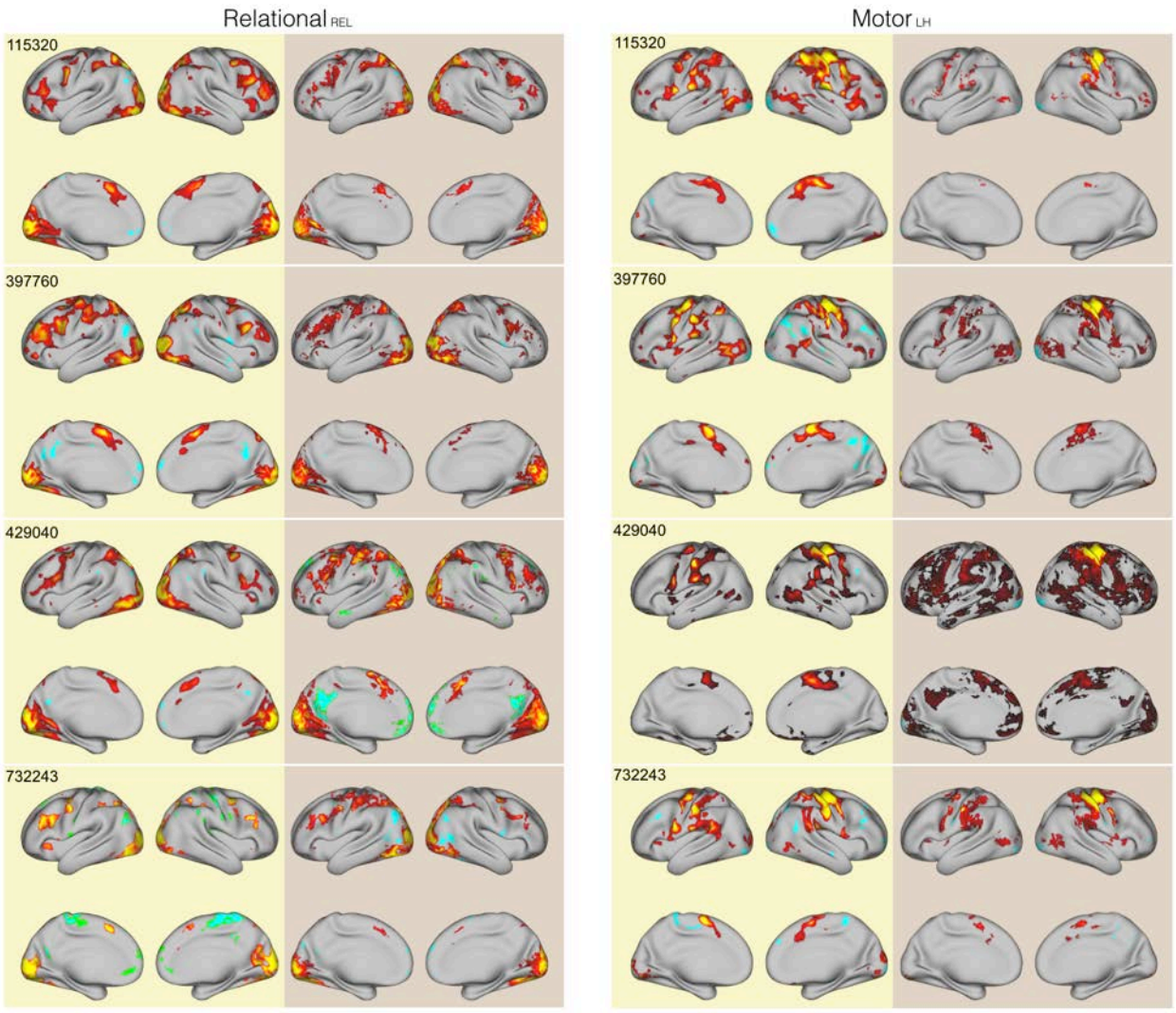
100307	101915	103414	103818	106319	110411	111009	111514	115320	117122	117324	118932	123117
124422	125525	127933	129028	130013	130316	130922	133928	136833	138231	138534	139637	140824
142828	143325	148032	148335	151627	153429	154431	157336	158035	158540	159239	159340	162733
163129	163432	172332	175439	176542	188347	190031	192540	194140	196144	198451	200614	201111
201414	205119	210617	212318	221319	224022	239944	245333	249947	250427	255639	280739	293748
298051	307127	366446	397760	414229	429040	448347	497865	530635	541943	545345	559053	570243
579665	627549	645551	654754	680957	702133	729557	732243	748258	756055	761957	765056	792564
856766	887373	894673	896778	901139	984472	992774						

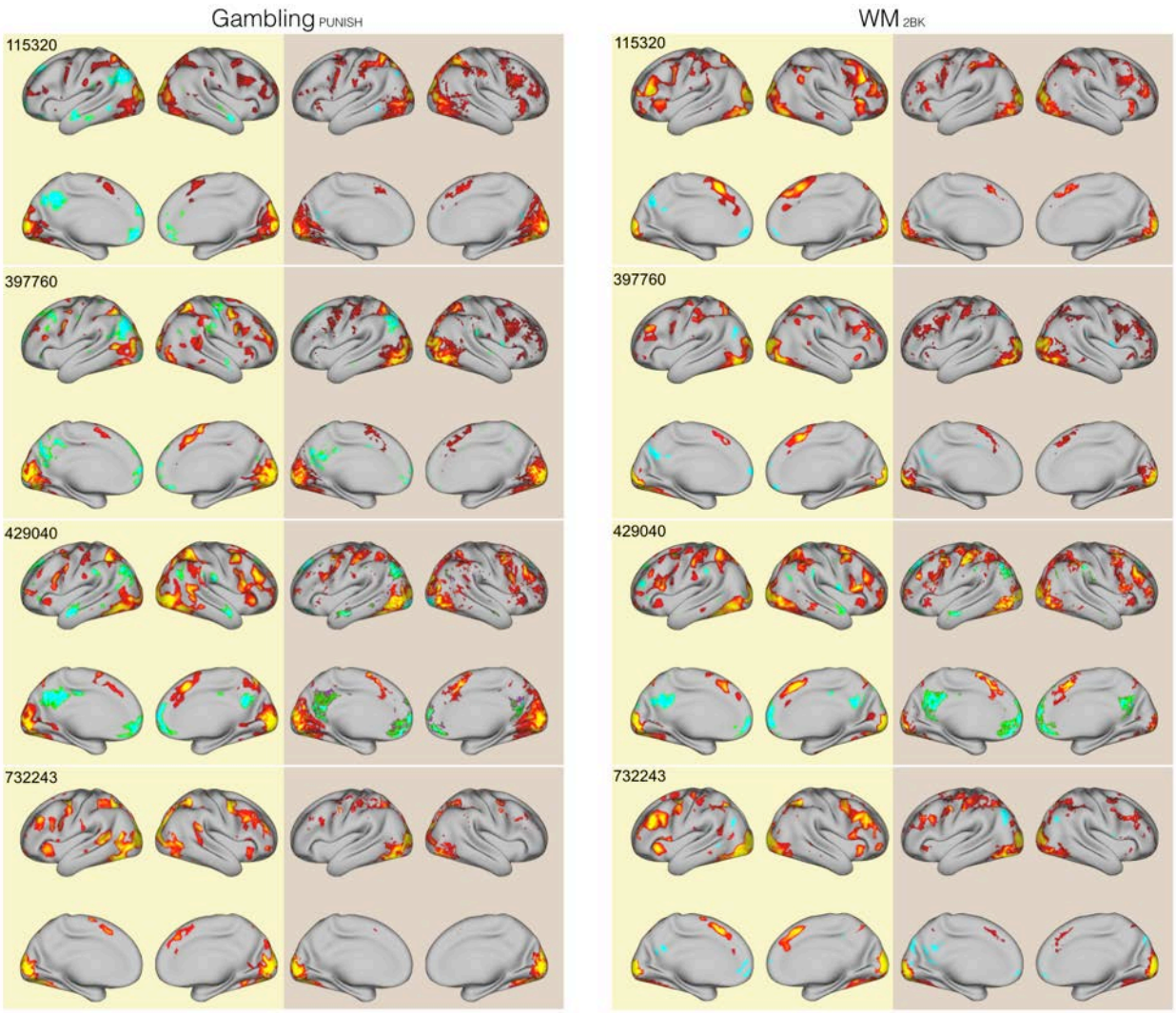
Table S3. Prediction scores. The score is calculated as the correlation coefficient between actual and predicted activation maps. Scores are averaged across subjects (standard deviation shown between brackets).

Domain	Contrast	Score (std)
EMOTION	FACES	0.66 (0.076)
EMOTION	SHAPES	0.63 (0.095)
EMOTION	FACES-SHAPES	0.57 (0.092)
GAMBLING	PUNISH	0.72 (0.064)
GAMBLING	REWARD	0.74 (0.056)
GAMBLING	PUNISH-REWARD	0.12 (0.101)
LANGUAGE	MATH	0.54 (0.099)
LANGUAGE	STORY	0.61 (0.101)
LANGUAGE	MATH-STORY	0.72 (0.068)
MOTOR	CUE	0.71 (0.077)
MOTOR	LF	0.55 (0.097)
MOTOR	LH	0.53 (0.097)
MOTOR	RF	0.57 (0.085)

MOTOR	RH	0.52 (0.100)
MOTOR	T	0.58 (0.096)
MOTOR	AVG	0.56 (0.090)
MOTOR	CUE-AVG	0.71 (0.072)
MOTOR	LF-AVG	0.55 (0.105)
MOTOR	LH-AVG	0.54 (0.103)
MOTOR	RF-AVG	0.54 (0.093)
MOTOR	RH-AVG	0.54 (0.104)
MOTOR	T-AVG	0.58 (0.093)
RELATIONAL	MATCH	0.80 (0.043)
RELATIONAL	REL	0.80 (0.041)
RELATIONAL	MATCH-REL	0.37 (0.184)
SOCIAL	RANDOM	0.73 (0.088)
SOCIAL	TOM	0.76 (0.058)
SOCIAL	RANDOM-TOM	0.52 (0.107)
WM	2BK_BODY	0.69 (0.069)
WM	2BK_FACE	0.68 (0.062)
WM	2BK_PLACE	0.70 (0.066)
WM	2BK_TOOL	0.70 (0.062)
WM	0BK_BODY	0.62 (0.089)
WM	0BK_FACE	0.59 (0.084)
WM	0BK_PLACE	0.70 (0.072)
WM	0BK_TOOL	0.67 (0.087)
WM	2BK	0.74 (0.051)
WM	0BK	0.72 (0.066)
WM	2BK-0BK	0.57 (0.170)
WM	BODY	0.70 (0.066)
WM	FACE	0.68 (0.065)
WM	PLACE	0.73 (0.062)
WM	TOOL	0.73 (0.063)
WM	BODY-AVG	0.42 (0.087)
WM	FACE-AVG	0.53 (0.096)
WM	PLACE-AVG	0.51 (0.096)
WM	TOOL-AVG	0.35 (0.122)


Language MATH-STORY
Social TOM
115320
397760
429040
732243
115320
397760
429040
732243


Relational REL
Motor LH
115320
397760
429040
732243
115320
397760
429040
732243



Figures S1-S3: Visual comparisons between predicted and actual activation maps in 4 representative subjects in the LANGUAGE/SOCIAL (S1), RELATIONAL/MOTOR (S2) and GAMBLING/WM (S3) behavioural domains. The maps are thresholded using a Gaussian-Gamma mixture model (see Methods) and color-coded according to the ammount of activation (Z-score). For each subject and domain, actual activations are presented on the left and predictions are presented on the right. The six-digit numbers are subjects IDs on the HCP database.

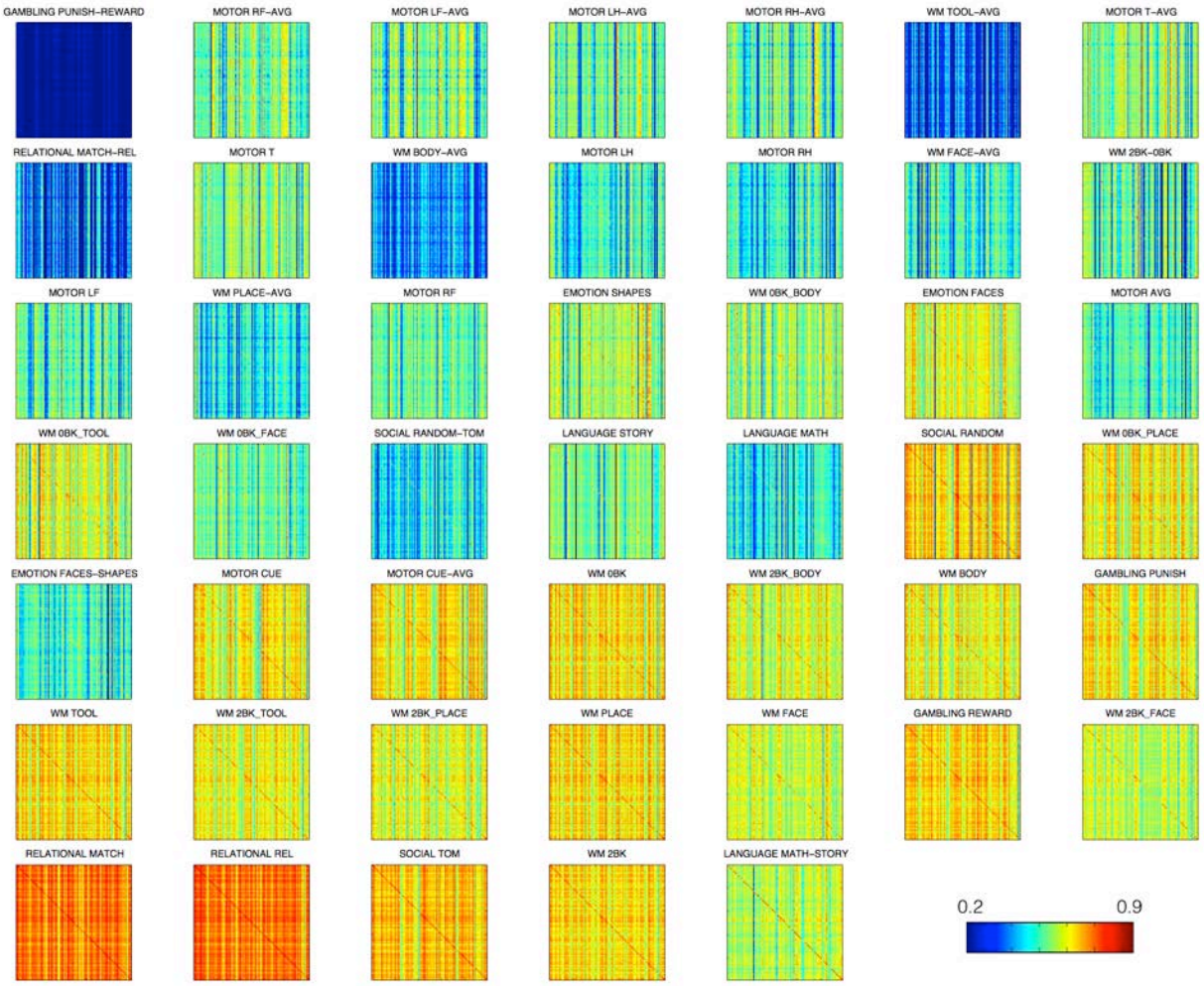


Fig. S4: Between-subjects cross-correlation matrices of predicted versus actual activation maps for all the task contrasts. The heavy diagonals indicate that for each subject the prediction matches that subject's map best. The tasks are ordered according to the mean of the diagonal divided by the standard deviation along each row.

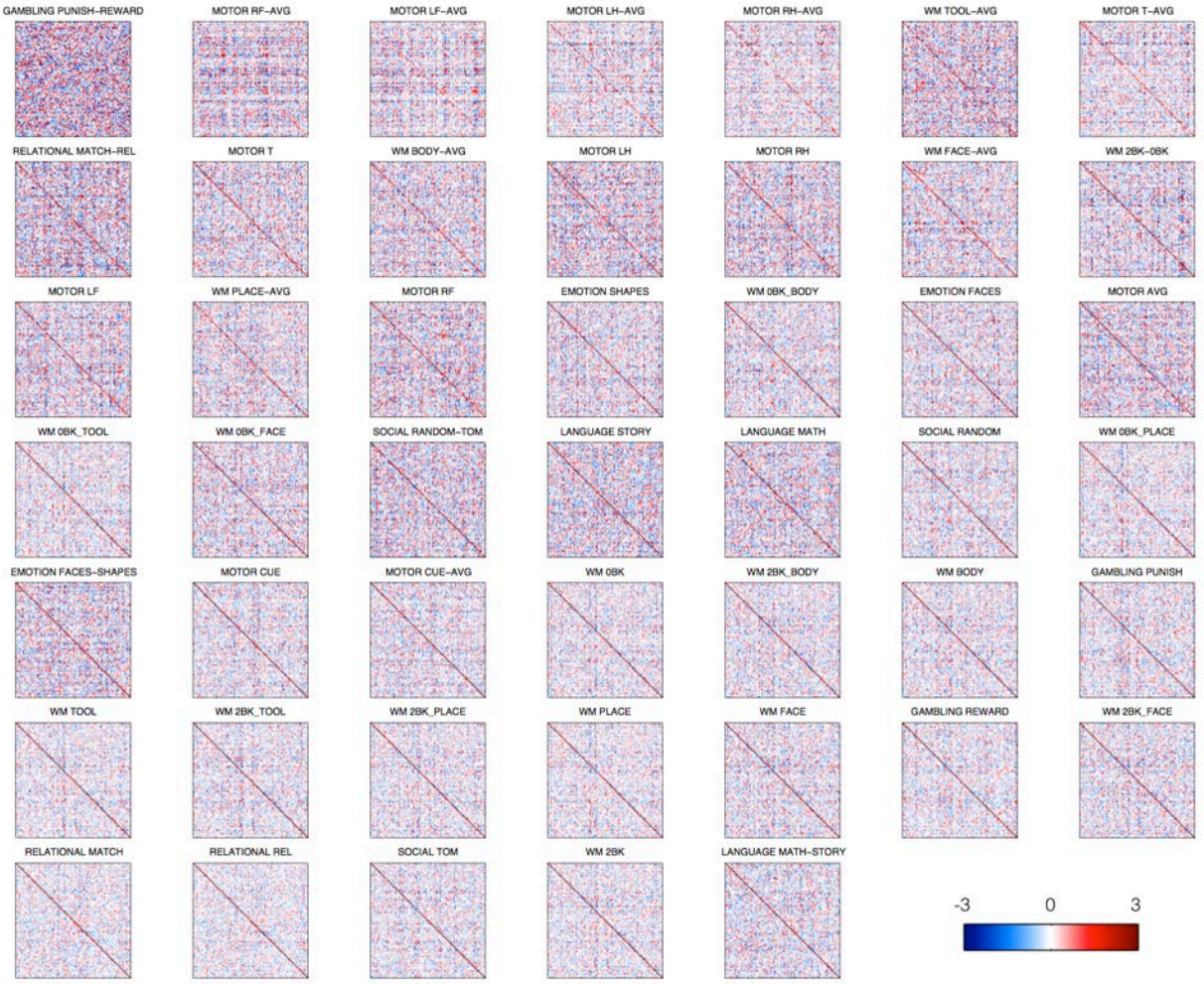


Fig. S5: Same as Fig. S4 but each matrix is row- and column- normalised. This normalization accounts for the fact that the actual maps are more variable than the predicted maps and removes the global mean correlation. The diagonal is even more prominent. Matrices are ordered as in Fig. S4.

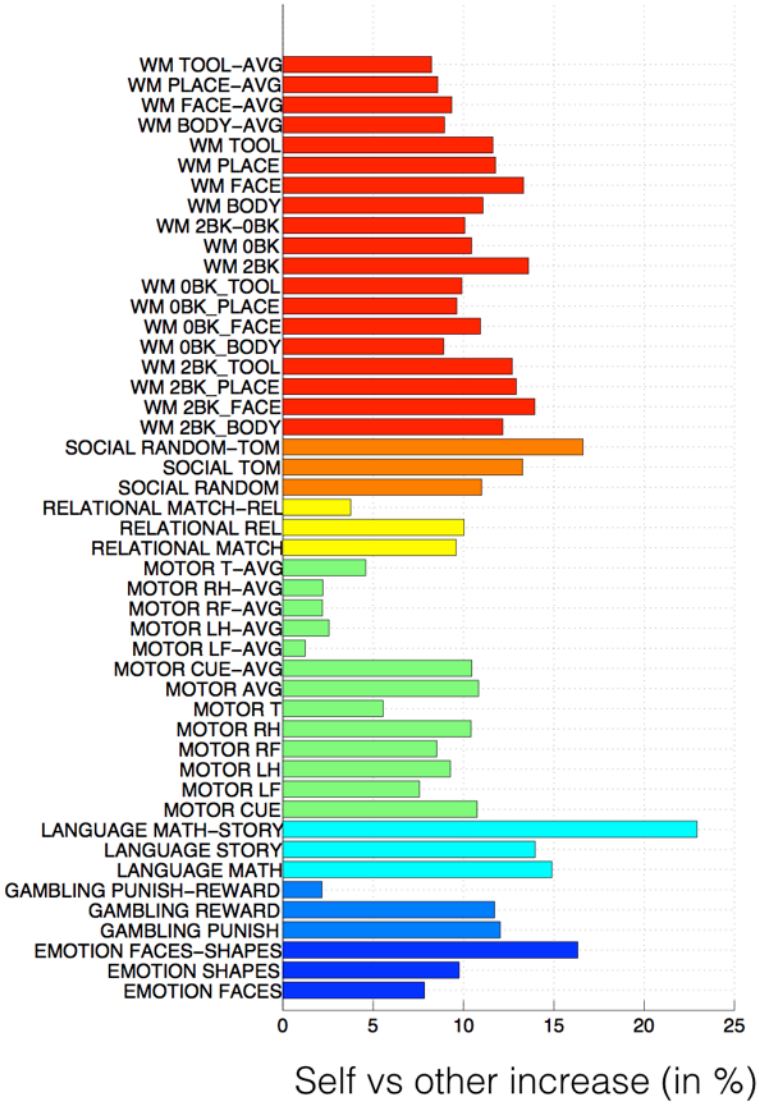


Fig. S6. Bar plots indicate the difference between the mean of the diagonal and of the extra-diagonal elements of the matrices in Fig. S4 (as a percentage relative to the mean of the extra-diagonal elements). In all cases, positive values show that the predictions, on average, match the actual maps better than the mean, particularly for the cognitive tasks.

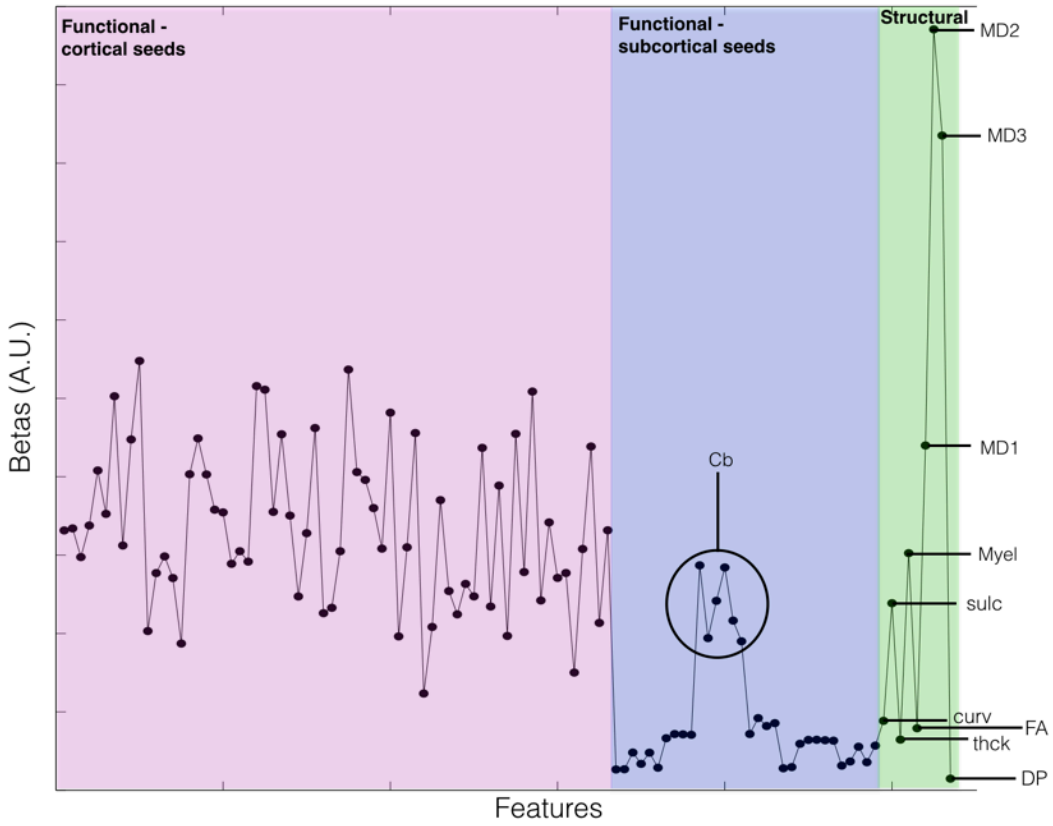


Fig. S7: Beta coefficients averaged across tasks and across brain parcels (after dividing by the maximum beta within task). Abbreviations: Cb (Cerebellum), sulc (sulcal depth), thck (cortical thickness), Myel (myelin map), FA (fractional anisotropy), MDx (Mean diffusivity for bvalue=x um^2/ms), DP (dot product between principal diffusion direction and cortical surface).

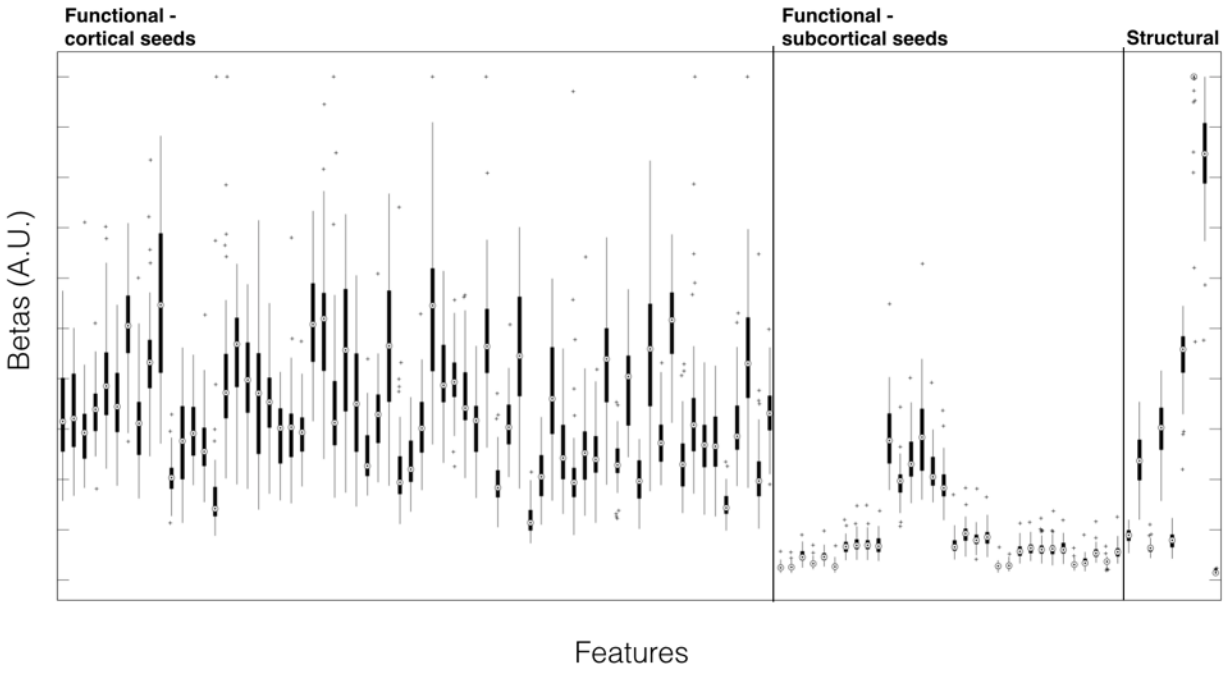


Fig. S8: As in S7 but showing median, quantiles, and outliers.

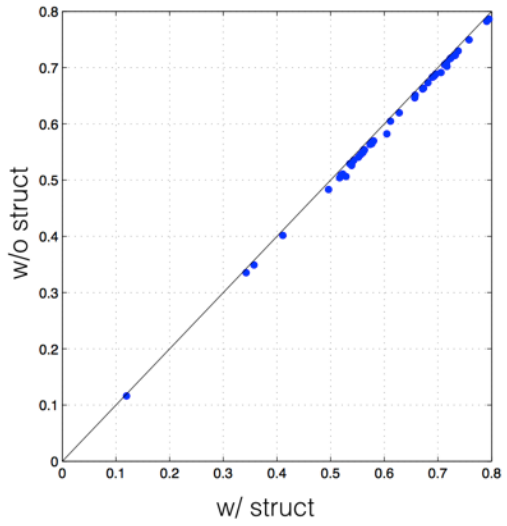


Fig. S9: Effect of excluding all structural features from the model (keeping only the resting-state predictors). The scatter plot shows the mean diagonal correlation matrix (between predictions and actual maps) for all 47 tasks with and without structural features. Overall, the addition of structural features only marginally improves the leave-one-out predictions.

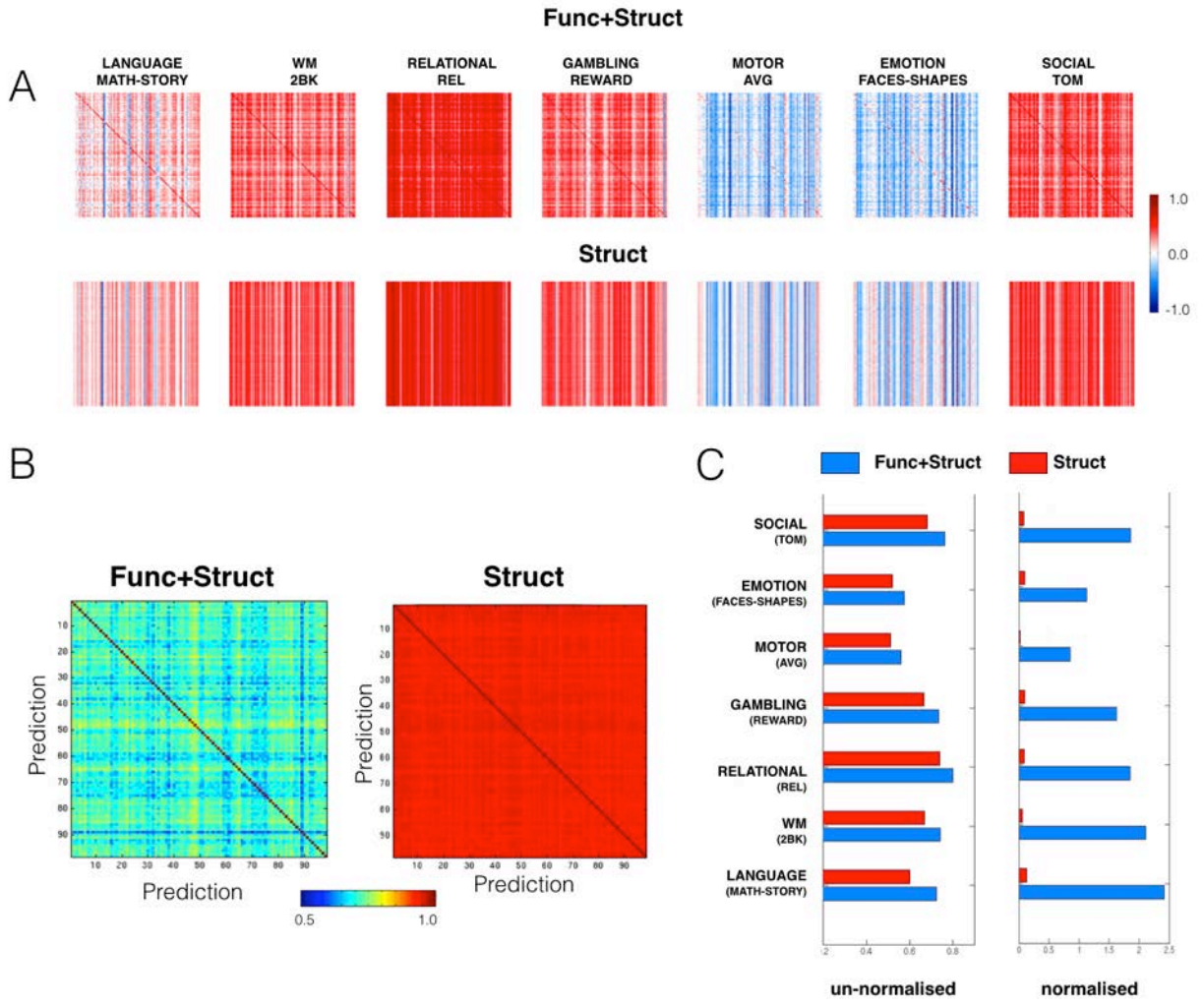


Fig. S10: Performance of a model based on structural features alone. (A) Pearson correlation matrices between actual (columns) and predicted (rows) activations in a complete model (top row) versus a structural-based model (bottom row). Removing functional features abolishes the diagonal-dominance of these matrices. (B) Cross-correlation matrix between predictions of all subjects in the LANGUAGE MATH-STORY task demonstrates the low variability of subjects' prediction in a model purely based on structural features as opposed to the complete model. (C) Quantification of the mean diagonal of the correlation matrix in a model with or without functional features. Overall, it is shown that using a model purely based on structural features abolishes inter-individual variability in the predictions.

Fig. S11: Between-tasks cross-correlation matrix.

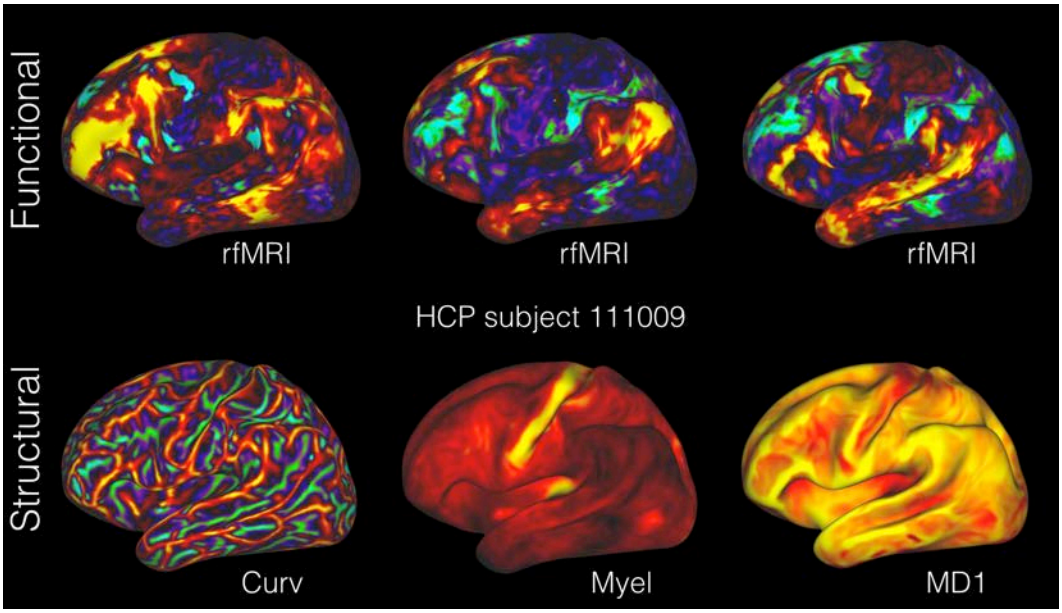


Fig. S12: Example feature maps for a single subject.

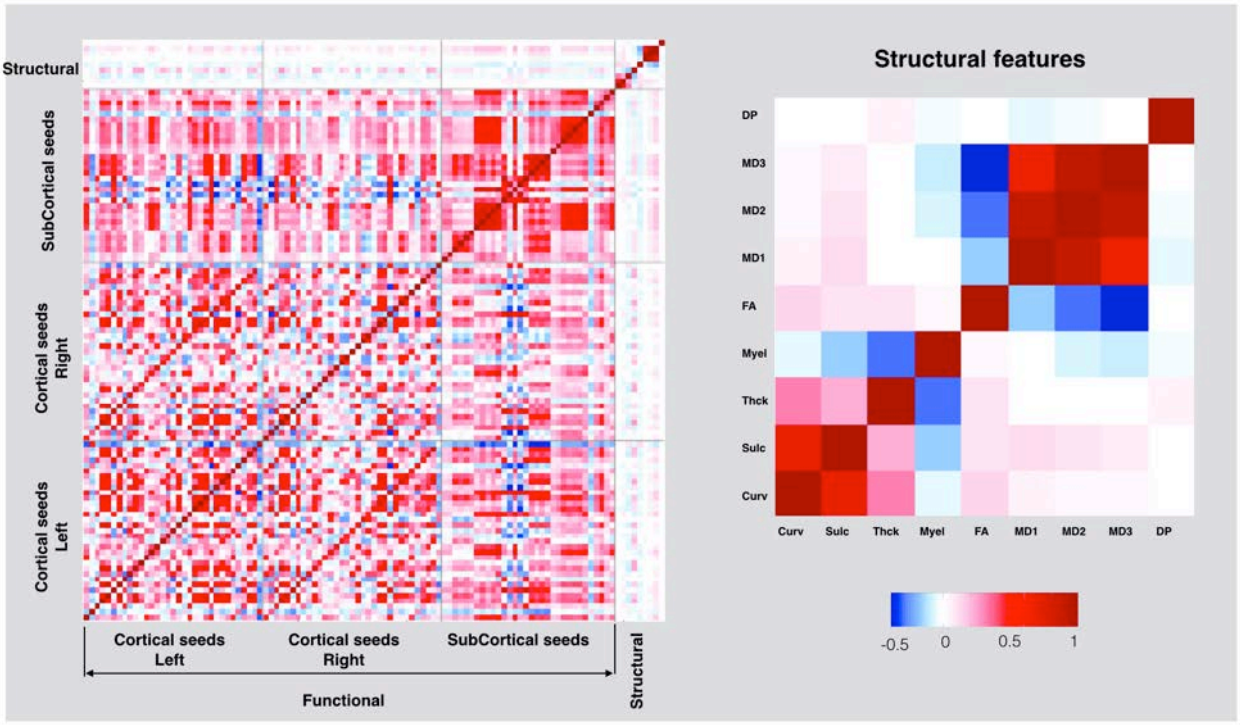


Fig. S13: Between-features cross-correlation matrix.

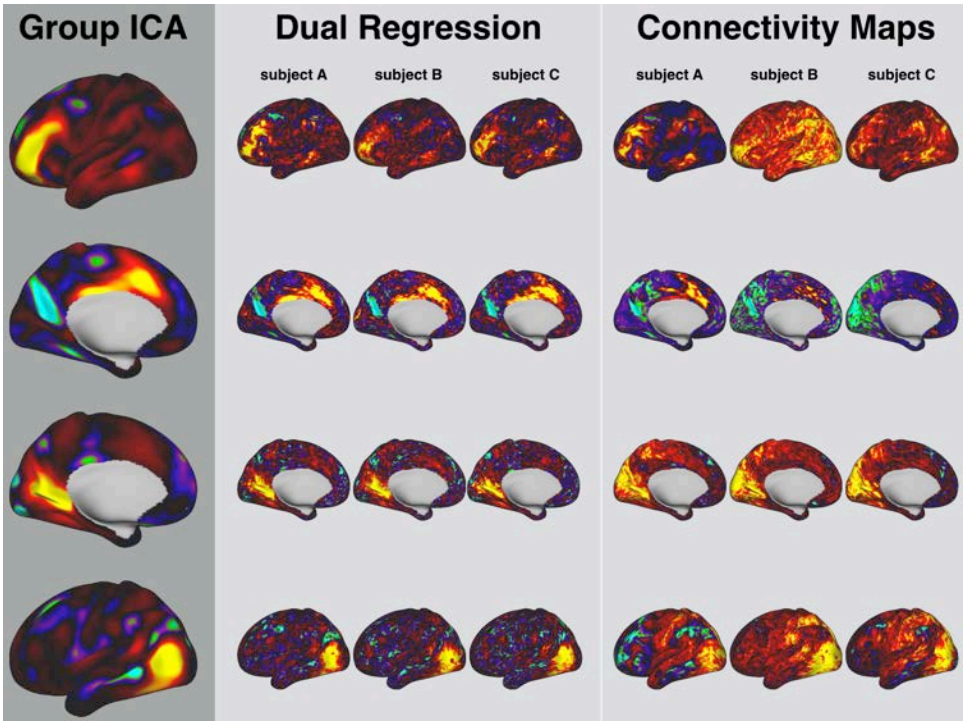


Fig S14: Examples maps from each of the feature extraction steps. The first column presents examples of 4 different independent components from the group ICA. The middle column shows the same components for three different subjects after performing dual regression to obtain

individual spatial maps. The last column shows the final features, which are a result of a regression against individual time series to produce individual connectivity maps.

Notation

$\mathbf{X}_i$	Features for subject i	91282 X 108 (column of 1's + 107 features)
$\mathbf{y}_i$	Task map for subject i	91282 X 1
$\mathbf{X}_i^{(k)}$	Features for subject i and parcel k	n_k X 108 (here 316 < n_k < 6771)
$\mathbf{y}_i^{(k)}$	Task map for subject i and parcel k	n_k X 1
$\boldsymbol{\beta}_i^{(k)}$	Regressor for subject i and parcel k	108 X 1

Learning and prediction

Regression	$\mathbf{y}_i^{(k)} = \mathbf{X}_i^{(k)} \boldsymbol{\beta}_i^{(k)}$
Leave-one-out	$\boldsymbol{\beta}_{LOO,i}^{(k)} = \sum_{i \neq j} \boldsymbol{\beta}_j^{(k)} / 97$
Prediction	$\hat{\mathbf{y}}_i^{(k)} = \mathbf{X}_i^{(k)} \boldsymbol{\beta}_{LOO,i}^{(k)}$

Fig S15: Description of the trained model.

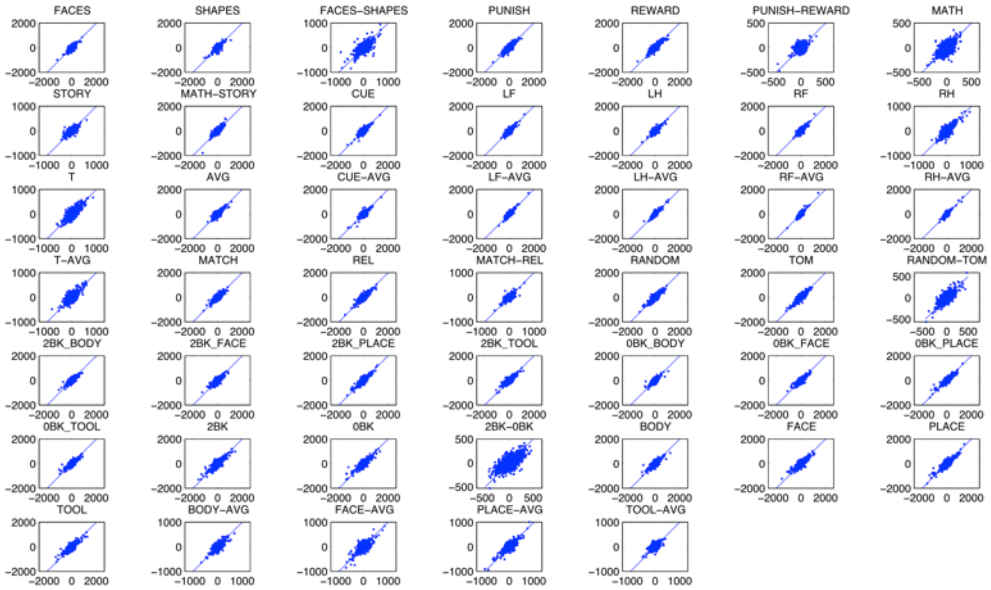


Fig S16: Split-half reproducibility. The group of 98 subjects was randomly divided into two halves and the model was trained on both halves separately. The scatter plots show one set of betas plotted against the other (pooled over brain parcels and subjects).

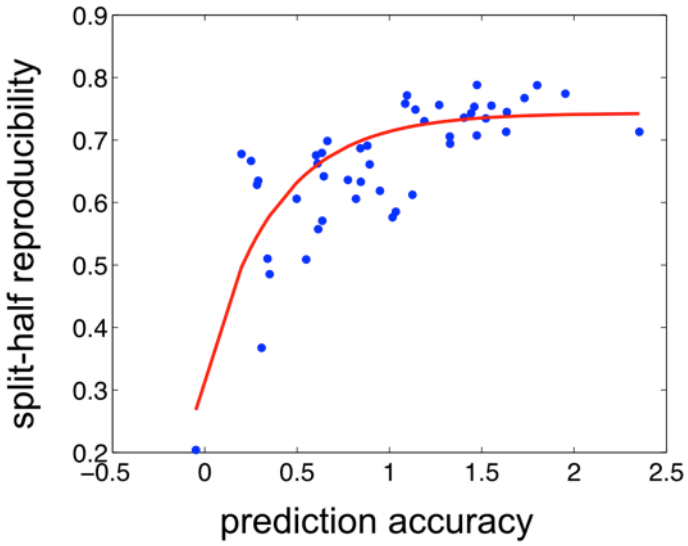


Fig S17: Comparison of split-half reproducibility (correlation coefficient between two split halves for each task) and the performance of the out-of-sample test. It is shown that reproducibility is higher when model accuracy is higher (red line is an exponential fit)

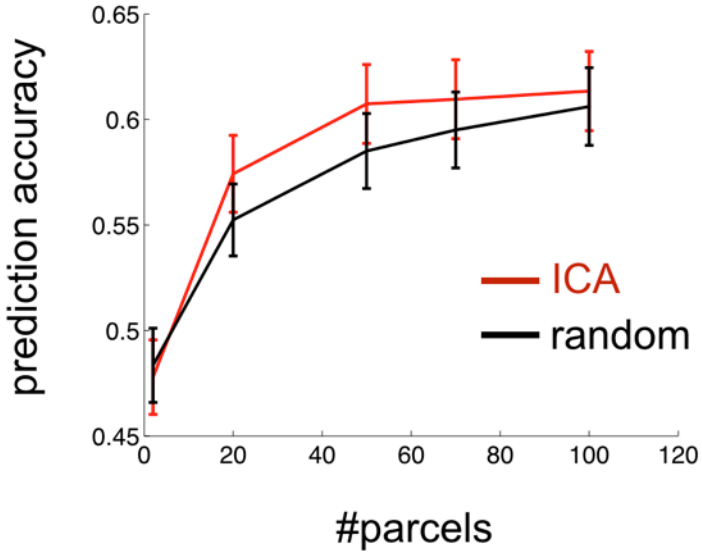


Fig S18: Prediction accuracy, averaged across subjects and tasks, using different number of parcels (K=2, 20, 50, 70, 100) and two different parcellation methods: ICA and random parcellation. Model accuracy increases with the granularity of the parcellation, but there is a

plateau at around 50 parcels. It is also shown that ICA marginally outperforms random parcellation for k>20, however that is not the case for all contrast maps.

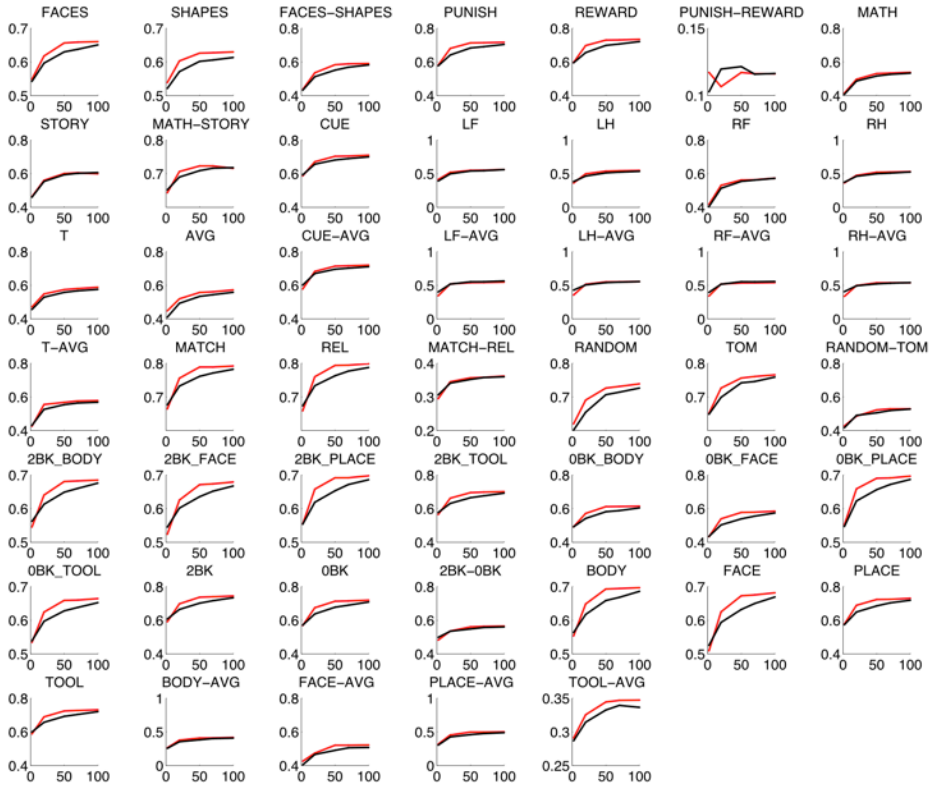


Fig S19: Per-task prediction accuracy as a function of the parcellation scheme.

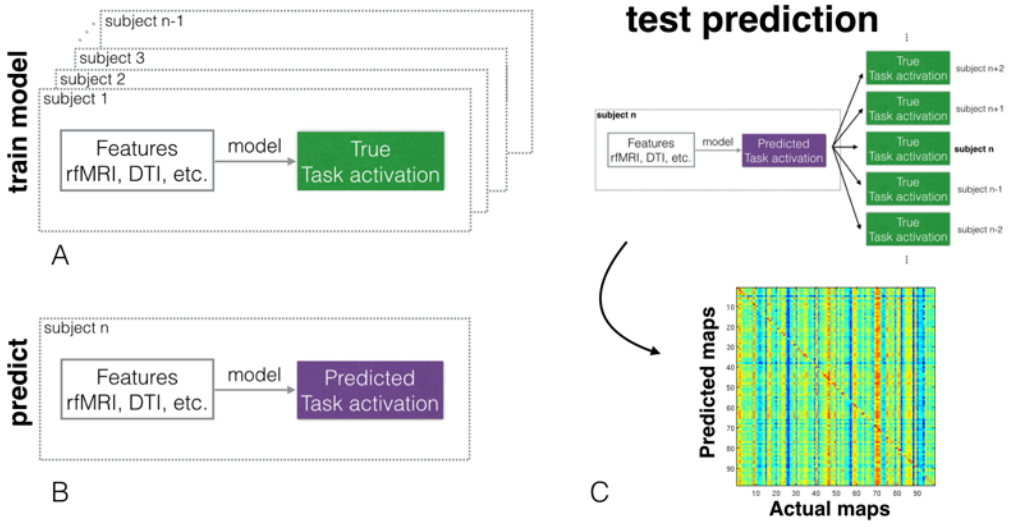


Fig. S20: Diagrams representing our strategies for (A) training the model, (B) predicting out-of-sample subjects, and (C) testing the model.

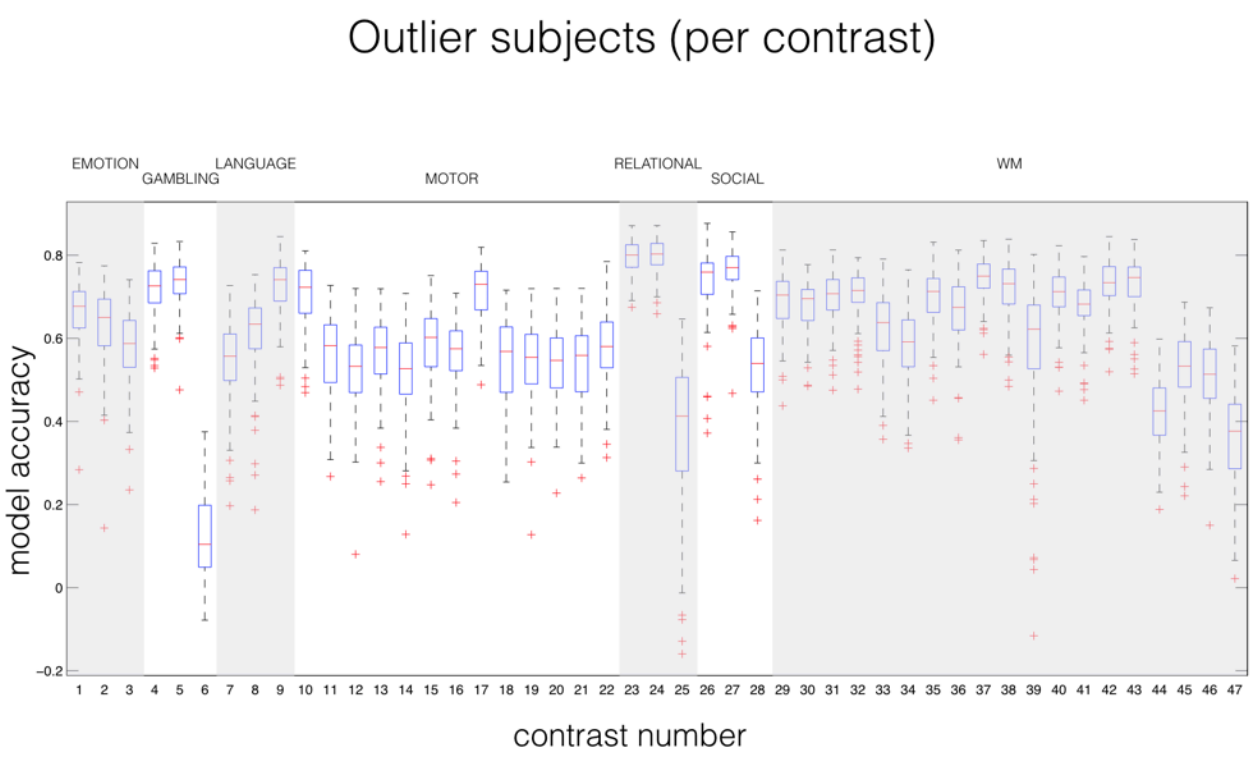


Fig. S21: Outlier Subjects. Box plots of the model performance in all 47 contrasts. Y-axis depicts the Pearson Correlation Coefficient as a measure of accuracy. It is shown that for each contrast there are only a few outlier subjects with poor prediction, while none of the subjects are outliers with high prediction accuracy.

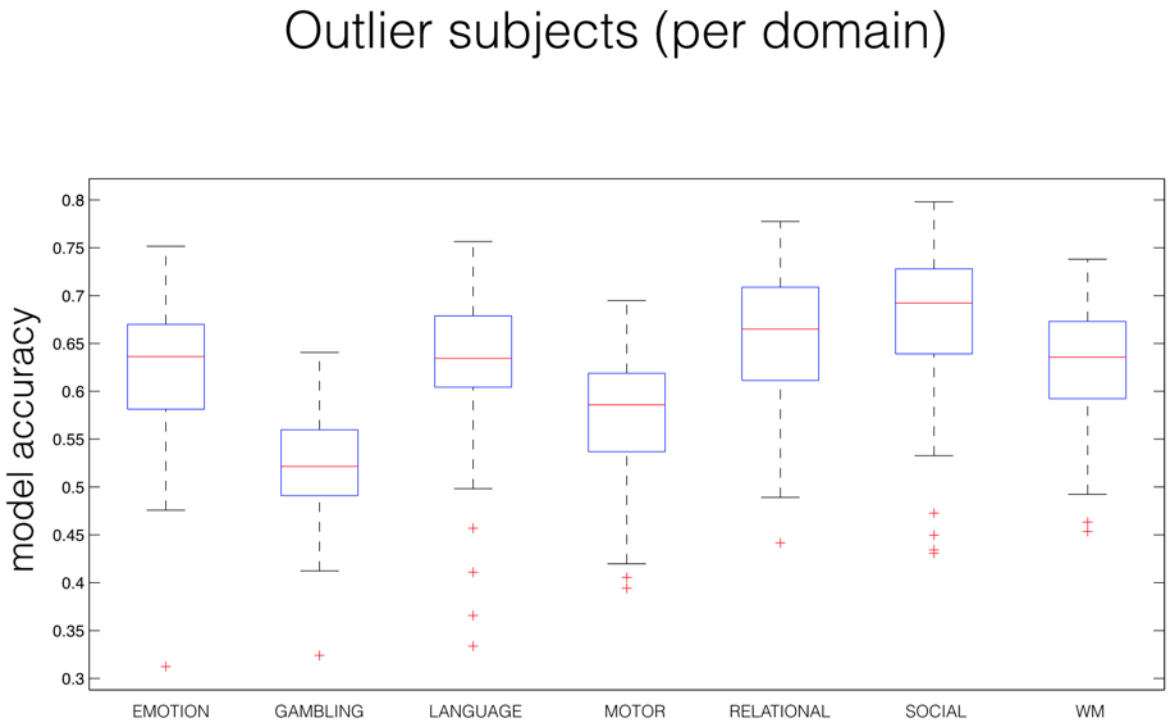


Fig. S22: This figure demonstrates the same result as Fig. S21, averaged within each cognitive domain.

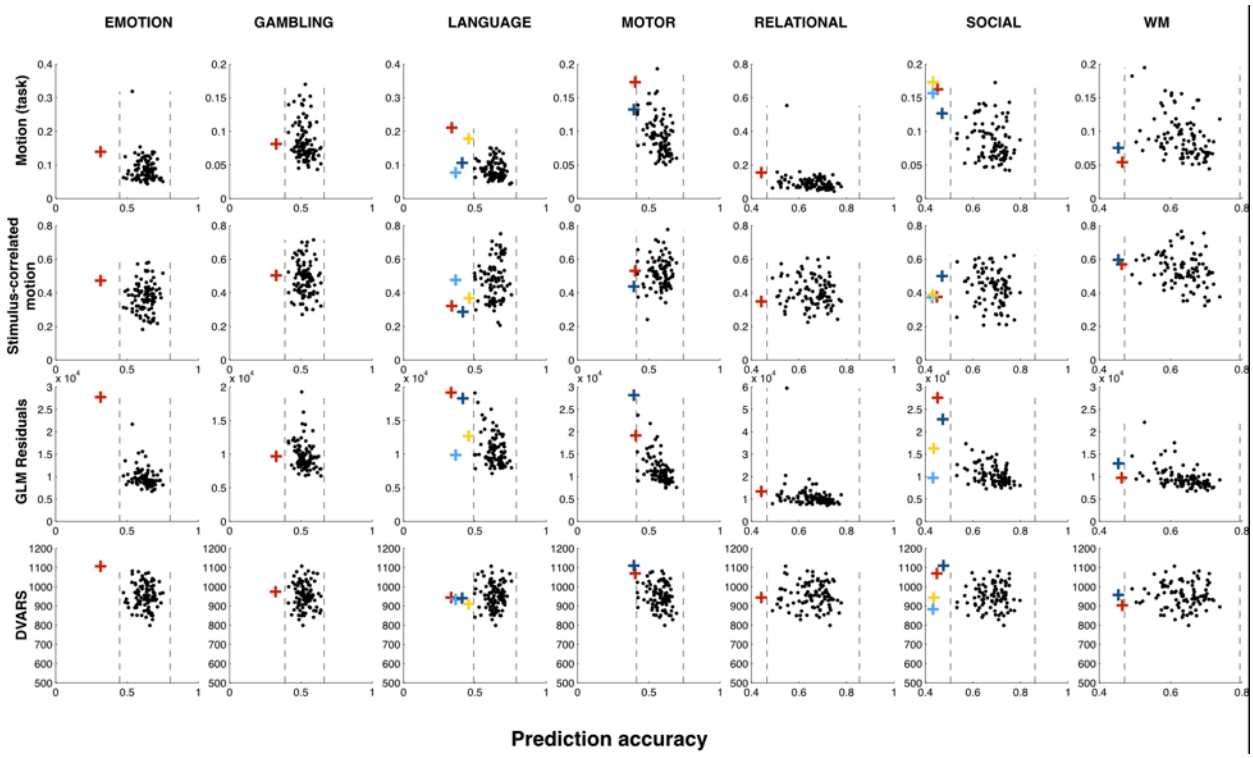


Fig. S23: Data quality versus outlier behavior. The figure demonstrates potential factors that can cause outlier behavior: subject's general motion; stimulus-correlated motion; GLM residuals; variance of the temporal derivative (DVARS) on the resting-state data. Outlier subjects (marked with a plus sign) are usually outliers in at least one of these measures as well.

Movies SM1-SM6

Movies SM1-SM6 present visual comparisons between predicted and actual activation maps in 10 subjects with the top 10 t-scores in each task (different subjects for different tasks). The maps are thresholded using a Gaussian-Gamma mixture model (see Methods) and binarised in order to see their overlap. Positive contrast maps are red/yellow; negative contrast maps are blue/light blue:

Movie SM1: LANGUAGE, contrast: MATH-STORY.

Movie SM2: behavior: GAMBLING, contrast: REWARD.

Movie SM3: behavior: RELATIONAL, contrast: MATCH.

Movie SM4: behavior: SOCIAL, contrast: TOM.

Movie SM5: behavior: WM, contrast: 2BK-0BK.

Movie SM6: behavior: EMOTION, contrast: FACES-SHAPES.

Movie SM7

This movie presents all 33 group-level ICA maps.

References and Notes

1. M. Jenkinson, S. Smith, A global optimisation method for robust affine registration of brain images. *Medical Image Analysis* **5**, 143-156 (2001); published online EpubJun (
2. J. Besle, R. M. Sanchez-Panchuelo, R. Bowtell, S. Francis, D. Schluppeck, Single-subject fMRI mapping at 7 T of the representation of fingertips in S1: a comparison of event-related and phase-encoding designs. *J Neurophysiol* **109**, 2293-2305 (2013); published online EpubMay (10.1152/jn.00499.2012).
3. F. McNab, T. Klingberg, Prefrontal cortex and basal ganglia control access to working memory. *Nat Neurosci* **11**, 103-107 (2008); published online EpubJan (10.1038/nn2024).
4. I. Mukai, D. Kim, M. Fukunaga, S. Japee, S. Marrett, L. G. Ungerleider, Activations in visual and attention-related areas predict and correlate with the degree of perceptual learning. *J Neurosci* **27**, 11401-11411 (2007); published online EpubOct 17 (10.1523/JNEUROSCI.3002-07.2007).
5. S. M. Tom, C. R. Fox, C. Trepel, R. A. Poldrack, The neural basis of loss aversion in decision-making under risk. *Science* **315**, 515-518 (2007); published online EpubJan 26 (10.1126/science.1134239).
6. G. S. Wig, S. T. Grafton, K. E. Demos, G. L. Wolford, S. E. Petersen, W. M. Kelley, Medial temporal lobe BOLD activity at rest predicts individual differences in memory ability in healthy young adults. *Proc Natl Acad Sci U S A* **105**, 18555-18560 (2008); published online EpubNov 25 (10.1073/pnas.0804546105).
7. R. Kanai, G. Rees, The structural basis of inter-individual differences in human behaviour and cognition. *Nat Rev Neurosci* **12**, 231-242 (2011); published online EpubApr (10.1038/nrn3000).
8. T. E. Behrens, H. Johansen-Berg, Relating connectional architecture to grey matter function using diffusion imaging. *Philos Trans R Soc Lond B Biol Sci* **360**, 903-911 (2005); published online EpubMay 29 (
9. R. E. Passingham, K. E. Stephan, R. Kotter, The anatomical basis of functional localization in the cortex. *Nat Rev Neurosci* **3**, 606-616 (2002); published online EpubAug (
10. B. T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zollei, J. R. Polimeni, B. Fischl, H. Liu, R. L. Buckner, The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* **106**, 1125-1165 (2011); published online EpubSep (10.1152/jn.00338.2011).
11. S. M. Smith, P. T. Fox, K. L. Miller, D. C. Glahn, P. M. Fox, C. E. Mackay, N. Filippini, K. E. Watkins, R. Toro, A. R. Laird, C. F. Beckmann, Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A* **106**, 13040-13045 (2009); published online EpubAug 4 (
12. D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W. U.-M. H. Consortium, The WU-Minn Human Connectome Project: an overview. *Neuroimage* **80**, 62-79 (2013); published online EpubOct 15 (10.1016/j.neuroimage.2013.05.041).
13. D. M. Barch, G. C. Burgess, M. P. Harms, S. E. Petersen, B. L. Schlaggar, M. Corbetta, M. F. Glasser, S. Curtiss, S. Dixit, C. Feldt, D. Nolan, E. Bryant, T. Hartley, O. Footer, J. M. Bjork, R. Poldrack, S. Smith, H. Johansen-Berg, A. Z. Snyder, D. C. Van Essen, W. U.-M. H. Consortium, Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage* **80**, 169-189 (2013); published online EpubOct 15 (10.1016/j.neuroimage.2013.05.033).
14. B. Fischl, N. Rajendran, E. Busa, J. Augustinack, O. Hinds, B. Yeo, H. Mohlberg, K. Amunts, K. Zilles, in *Cereb Cortex*. (Department of Radiology, Harvard Medical School, Charlestown, MA 02129, USA., 2007).
15. T. J. Andrews, S. D. Halpern, D. Purves, Correlated size variations in human visual cortex, lateral geniculate nucleus, and optic tract. *J Neurosci* **17**, 2859-2868 (1997); published online EpubApr 15 (
16. S. Knecht, M. Deppe, B. Drager, L. Bobe, H. Lohmann, E. Ringelstein, H. Henningsen, Language lateralization in healthy right-handers. *Brain* **123 (Pt 1)**, 74-81 (2000); published online EpubJan (
17. M. Thiebaut de Schotten, M. Urbanski, H. Duffau, E. Volle, R. Levy, B. Dubois, P. Bartolomeo, Direct evidence for a parietal-frontal pathway subserving spatial awareness in humans. *Science* **309**, 2226-2228 (2005); published online EpubSep 30 (10.1126/science.1116251).

18. M. W. Cole, D. S. Bassett, J. D. Power, T. S. Braver, S. E. Petersen, Intrinsic and task-evoked network architectures of the human brain. *Neuron* **83**, 238-251 (2014); published online EpubJul 2 (10.1016/j.neuron.2014.05.014).
19. E. C. Robinson, S. Jbabdi, M. F. Glasser, J. Andersson, G. C. Burgess, M. P. Harms, S. M. Smith, D. C. Van Essen, M. Jenkinson, MSM: a new flexible framework for Multimodal Surface Matching. *Neuroimage* **100**, 414-426 (2014); published online EpubOct 15 (10.1016/j.neuroimage.2014.05.069).
20. N. K. Logothetis, B. A. Wandell, Interpreting the BOLD signal. *Annual review of physiology* **66**, 735-769 (2004)10.1146/annurev.physiol.66.082602.092845).
21. K. Friston, Functional and effective connectivity: a review. *Brain Connectivity*, (2011).
22. Z. M. Saygin, D. E. Osher, K. Koldewyn, G. Reynolds, J. D. E. Gabrieli, R. R. Saxe, Wired for function: Anatomical connectivity patterns predict face-selectivity in the fusiform gyrus. *Nat Neurosci* **15**, (2012).
23. P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, A. Aldroubi, In vivo fiber tractography using DT-MRI data. *Magn Reson Med* **44**, 625-632 (2000); published online EpubOct (
24. E. S. Finn, X. Shen, D. Scheinost, M. D. Rosenberg, J. Huang, M. M. Chun, X. Papademetris, R. T. Constable, Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat Neurosci* **18**, 1664-1671 (2015); published online EpubNov (10.1038/nn.4135).

Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The data are available for download at www.humanconnectome.org. Data from the Q3 release (September 2013) were used in this paper. All code available upon request from the corresponding author. Funding was provided by the UK Medical Research Council (MR/L009013/1 to SJ), UK EPSRC (EP/L023067/1 to SJ and TEB), James S McDonnell Foundation (JSMF220020372 to TEB), the Wellcome Trust (WT104765MA to TEB and 098369/Z/12/Z to SMS) and the Netherlands Organization for Scientific Research NWO (452-13-015 to RBM). Author contributions were as follows: SJ and IT built the model and performed the analyses. SJ, IT, RBM, OPJ, SMS and TEB wrote the paper.

Figure legends:

Fig. 1. Predicting individual variations in task maps. Figure shows actual and predicted thresholded task maps in 3 subjects and 4 different task contrasts. The model is able to capture striking variations between subjects in the shape, topology and extent of their activation maps. Predictions and comparisons to actual maps for all behavioral domains and more subjects are shown in Figures S1-S3 and movies SM1-SM6.

Fig. 2. Specificity of the individual predictions. A subject's prediction map is more similar to the subject's actual map than to the rest of the subjects. (A) Predicted maps (zoomed on the right frontal lobe) of the MATH-STORY contrast from the LANGUAGE task of 3 subjects are overlapped with their actual activation maps (top row). We also overlap the subjects' predictions with the activation map of the median subject (bottom row). Blue represents actual activation,

red is the predicted activation and yellow is the overlap. The maximum overlap is obtained when comparing a given subject's prediction to their own actual map. (B) Pearson correlation matrix between actual (columns) and predicted activations (rows). The correlation matrix is noticeably diagonal-dominant, indicating that on average the model prediction for any given subject is more similar to the subject's own map than to other subjects' maps. This is also shown as a histogram plot, where the extra-diagonal elements of the correlation matrix (subject X vs subject Y) are compared to the diagonal elements (subject X vs subject X). The vertical dashed line corresponds to the median of the correlation coefficients along the diagonal. (C) Correlation matrices and histograms for 6 additional behavioral domains. When normalizing the rows and columns of the correlation matrices (which removes the mean and accounts for higher variability in actual than predicted maps), the diagonal-dominance is even more prominent. In all cases, a Kolmogorov-Smirnov test between the two distributions (self vs other) gives a highly significant difference ($p < 10^{-10}$).

Fig. 3. Capturing qualitative and quantitative inter-individual differences. A. Variations in location, shape and topology are predicted by the model (contrast: LANGUAGE MATH-STORY). B. Peak Z scores were calculated for each hemisphere to examine how well the model can predict the amount of activation for each subject. A lateralization index (difference between right and left peak activation levels) is then calculated for each subject for both predicted and actual data and is shown as red and blue bars, respectively (LANGUAGE task). The model is able to predict individual subjects' lateralization index for both contrasts, including in the case where the majority of the subjects are left-lateralized. Statistical tests: MATH-STORY ($r = 0.47$, $p < 10^{-5}$), STORY-MATH ($r = 0.48$, $p < 10^{-6}$).

Fig. 4. Predictions in atypical subjects. The figure demonstrates the ability of our prediction model to detect inter-subject variability when subjects differ from the group-averaged activation. In each row the group activation for each behavioral domain is shown on the left, and the actual and predicted activations for 2 subjects are shown on the right. In each row we show activations and predictions for one subject that is similar to the group activation (A) and one that differs from it (B) as shown by the black circles. The model can capture the presence of clusters that are not active in the group (rows 1,4,5) as well as the absence of clusters that are active in the group (rows 2,3). Note that subjects A and B are not the same pair of subjects across behavioral domains.

Supplementary stuff