Using Temporal ICA to Selectively Remove Global Noise While Preserving Global Signal in Functional MRI Data

Matthew F. Glasser*[,1,2], Timothy S. Coalson[1], Janine D. Bijsterbosch[3], Samuel J. Harrison[3], Michael P. Harms[4], Alan Anticevic[5], David C. Van Essen[1], and Stephen M. Smith[3].

[1]Department of Neuroscience, Washington University Medical School, Saint Louis, Missouri 63110, USA. [2]St. Luke's Hospital, Saint Louis, Missouri 63017, USA. [3]Centre for Functional MRI of the Brain (FMRIB), Wellcome Centre for Integrative Neuroimaging, Nuffield Department of Clinical Neurosciences, University of Oxford. John Radcliffe Hospital, Headley Way, Oxford, OX3 9DU, UK. [4]Department of Psychiatry, Washington University Medical School, Saint Louis, MO. [5]Department of Psychiatry, Yale University School of Medicine, 300 George Street, New Haven, CT 06511, USA

*Address correspondence to glasserm@wustl.edu

Abstract

Temporal fluctuations in functional Magnetic Resonance Imaging (fMRI) have profitably been used to study brain activity and connectivity for over two decades. Unfortunately, fMRI data also contain structured temporal "noise" from a variety of sources, including subject motion, subject physiology, and the MRI scanner. Recently, methods have been developed to automatically and selectively remove spatially specific structured noise from fMRI data using spatial independent components analysis (ICA) and machine learning classifiers. Spatial ICA is particularly effective at removing spatially specific structured noise from high temporal and spatial resolution fMRI data of the type acquired by the Human Connectome Project and similar studies. However, spatial ICA is unable (mathematically, by design) to separate spatially widespread "global" structured noise—e.g., blood flow modulations from subject respiration—from fMRI data. No methods currently exist to selectively and completely remove global structured noise while retaining the global signal from neural activity. This has left the field in a quandary—to do or not to do global signal regression—given that both choices have substantial downsides. Here we show that temporal ICA can selectively segregate and remove global structured noise while retaining global neural signal in both task-based and resting state fMRI data. We compare the results before and after temporal ICA cleanup to those from global signal regression and show that temporal ICA cleanup removes the biases caused by global physiological noise without inducing the biases of global signal regression. We show that temporal ICA cleanup provides a "best of both worlds" solution to the global signal and noise dilemma and that temporal ICA itself unlocks interesting neurobiological insights from fMRI data.

Introduction

Functional MRI (fMRI) is a remarkably powerful tool for non-invasive mapping of brain function and for estimating the functional interactions between brain areas (i.e., "functional connectivity"). Although its indirectness has been criticized (Farah, 2014), both task-based and resting state fMRI are capable of precisely replicating known fine-scale patterns of functional organization and connectivity in the human visual cortex (Glasser et

al., 2016a; Sereno et al., 1995) when analyzed so that spatial details are preserved by aligning areas across subjects on the cortical surface and avoiding smoothing (Glasser et al., 2016b). Such patterns closely reflect many fine details observed invasively in nonhuman primates (Brewer et al., 2002; Gattass et al., 1988; Van Essen et al., 1984), lending strong support for fMRI as a tool to non-invasively map the brain's functional architecture.

It is also known, however, that a variety of confounding and nuisance signals are present in the fMRI timeseries, particularly those from subject motion (Power et al., 2012; Power et al., 2014; Satterthwaite et al., 2012; Van Dijk et al., 2012; Yan et al., 2013), and subject physiology, such as respiration and heart rate (Birn et al., 2006; Chang et al., 2009; Chang and Glover, 2009; Golestani et al., 2015; Power et al., 2017; Shmueli et al., 2007). If not removed systematically, these signals can lead to false positives, false negatives, and erroneous interpretations of fMRI results. Indeed, how to appropriately and comprehensively remove structured (i.e. non-Gaussian) temporal noise from fMRI data is a longstanding methodological controversy in the functional neuroimaging community (Aguirre et al., 1997; Aguirre et al., 1998; Anderson et al., 2011; Fox et al., 2009; Liu, 2016; Liu et al., 2017; Macey et al., 2004; Murphy et al., 2009; Murphy and Fox, 2017; Power et al., 2014; Power et al., 2015; Saad et al., 2012; Shmueli et al., 2007; Zarahn et al., 1997). Analogous to the approach that has proven valuable in the spatial domain (Glasser et al., 2016b; Glasser et al., 2013), it is desirable to use temporal cleanup methods that selectively remove artifacts while preserving as much of the neurally generated signal of interest as possible.

Significant progress has been made in developing automated methods of selectively removing spatially specific structured temporal noise—i.e., non-random time-varying artifacts that are also spatially nonuniform—using spatial ICA (sICA, e.g., sICA+FIX, FMRIB's ICA component classifier, (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014); see also ICA-AROMA, (Pruim et al., 2015a; Pruim et al., 2015b), and Multi-Echo ICA (Kundu et al., 2012)). Semi-global and global structured noise has presented a more difficult challenge, however. sICA is inherently unable to separate global temporal fluctuations into spatially orthogonal signal and noise components (see Main Supplementary Information Supplementary Introduction), and thus the global fluctuations will remain mixed into all of the sICA component timeseries. Indeed, several studies have used indirect evidence to assert that there is residual structured noise in Human Connectome Project (HCP) resting state fMRI data after sICA+FIX cleanup (Burgess et al., 2016; Power et al., 2017; Siegel et al., 2017); however, this noise has not been selectively separated from the data or characterized as to its spatial and temporal properties. The most common approach for removing global structured noise is to remove the mean (across space) fMRI timecourse from the data either explicitly (using Global Signal Regression, GSR (Power et al., 2017) or implicitly (by including white matter and ventricle voxels that are likely contaminated by the mean grey-matter signal because of voxel size, unconstrained volumetric smoothing, or spatial proximity (Behzadi et al., 2007; Chai et al., 2012; Marx et al., 2013; Muschelli et al., 2014; see Power et al., 2017).

Global signal regression is not selective, and its appropriateness relies on the assumption that global fluctuations are entirely noise. However, GSR will also remove or reduce any global or semi-global neural signal, in particular impacting resting state networks that are spatially widespread or have higher amplitude fluctuations such that they contribute more to the global signal (Glasser et al., 2016b). As a result, GSR and

related approaches are controversial because the resulting connectivity or activation measures may be distorted in complex and network-specific ways (Glasser et al., 2016b; Saad et al., 2012; Yang et al., 2016; Yang et al., 2014). That said, it is clear that non-neuronal physiological processes induce global fluctuations in fMRI timeseries through the BOLD mechanism, particularly those that globally change brain blood flow such as respiration rate and depth (Power et al., 2017). Such artifactual fluctuations will lead to positive biases in functional connectivity (or functional activation if stimulus correlated), which may vary across subjects and groups to create confounds for brain imaging studies (Glasser et al., 2016b; Hayasaka, 2013; Power et al., 2017). These global fluctuations are incompletely removed by current models of physiological regressors based on separate recording of physiological parameters (Power et al., 2017). Thus, the field has lacked an effective and mathematically unbiased solution for removing global and semi-global sources of structured temporal noise while retaining global, semi-global, or high amplitude BOLD neural signal (Glasser et al., 2016b; Power et al., 2017).

Here, we use a temporal ICA-based (tICA) approach to address this problem (see Main Supplementary Information Supplementary Introduction). Building upon recent success in separating spatially specific neural signal and structured noise by decomposing fMRI data into spatially orthogonal independent components (using sICA), we sought to further subdivide the resulting sICA-cleaned data into temporally orthogonal independent components (using tICA). Unlike sICA, which is mathematically blind to global signal and noise, tICA is able to identify and separate global and semi-global temporally independent components (Smith et al., 2012). We explore the relationship between tICA components and both physiological and subject motion parameters in task-based and resting state fMRI. We identify multiple specific components as non-neuronal structured noise, remove these components from the fMRI data, and compare the results of common task-based and resting state fMRI analyses before tICA cleanup, after tICA cleanup, and after global signal regression.

At the outset, it is also worth setting some expectations of what the proposed method aims to do and what is outside of the purview of a data denosing approach. Our objective is to clean the fMRI timeseries of global and semi-global noise from physiological or other sources and also to remove any remaining spatially specific noise that may have been left behind by sICA+FIX so that the timeseries reflects, as accurately as possible, the "true" neural BOLD signal that occurred in each subject's brain during their scan. Our aim is not to remove the neurally driven BOLD effects of undesired subject behavior in the scanner, however. Some subjects may not have actually done the requested task in task-based fMRI (or failed to keep their eyes open and fixated on the cross-hairs for resting state fMRI); they may have moved frequently and as a result have neurally driven BOLD effects arising from motion; and some subjects may have become drowsy or fallen asleep during the long HCP resting state scans, causing profound changes in their functional connectivity (Laumann et al., 2017; Liu et al., 2017; Tagliazucchi and Laufs, 2014; Wong et al., 2016; Wong et al., 2013; Yeo et al., 2015). While it may be appropriate for some studies to set stringent data inclusion/exclusion criteria surrounding such "neural compliance" issues, dealing with these issues outside the scope of our method and this study (see Discussion), though temporal ICA does shed some light upon them. Importantly, we should not expect measures of neural activity or connectivity to be the same during periods of non-compliance as they would have been if the subject had been compliant. A number of data

cleanup metrics in the literature (Burgess et al., 2016; Ciric et al., 2017; Power et al., 2017; Power et al., 2015; Siegel et al., 2017) assume (explicitly or implicitly) that neural activity will be the same during periods of compliance and non-compliance, and so we will not use these metrics in this study (see Supplementary Information: Supplementary Discussion). Finally, we emphasize the particular relevance of our approach to the growing amount of high spatial and temporal resolution multi-band fMRI data being acquired as a part of the Human Connectome Project and related HCP-Style neuroimaging efforts (Glasser et al., 2016b). We focus our efforts on multi-band fMRI data because this approach is quickly becoming standard in the field and the powerful data-driven methods we will use work best on such data.

Subjects and Methods

Subject Population: Data from 449 Human Connectome Project young healthy adults (ages 22 – 35) were used in this study. These data were acquired in accordance with the Washington University Institutional Review Board (Van Essen et al., 2013). Only subjects with complete fMRI acquisitions (resting-state and task) from the S500 HCP data release were included (Glasser et al., 2016a).

Images Acquired: T1-weighted and T2-weighted structural scans were acquired as previously described (Glasser et al., 2013). Resting-state fMRI data were acquired with 2.0mm isotropic resolution, TR=0.72s, and 1200 frames (14.4 min) per run. Two runs with reversed phase encoding directions, RL or LR, with the order counterbalanced across each of two sessions, were acquired (Smith et al., 2013a). Task fMRI data were acquired with identical pulse sequence settings while subjects performed 7 tasks (Barch et al., 2013) with runs lasting between 2 and 5 minutes (176 - 405 frames), and totaling 22.6 min (1884 frames) in session 1 and 24.0 min (1996 frames) in session 2. Each task is comprised of a pair of runs with reversed phase encoding directions, RL and then LR. Spin echo EPI scans were acquired (termed "spin echo field map" scans by HCP) and used to correct the fMRI data for B0-induced geometric distortions and B1-receive-induced image intensity inhomogeneities (Glasser et al., 2016a; Glasser et al., 2013).

Image Preprocessing: HCP spatial image preprocessing has been previously described in detail (Glasser et al., 2016a; Glasser et al., 2013). In brief, it involves minimizing smoothing while doing the following: 1) removing MR-induced image distortions so that each image represents the physical space of the subject; 2) removing the spatial effects of subject motion within and between modalities; and 3) projecting the data to a 2mm average spacing standard CIFTI grayordinates space after a FNIRT-based T1w nonlinear subcortical volume alignment and an MSMAll areal-feature-based cortical alignment, which uses myelin maps, resting state network maps, and resting state visuotopic maps for registration (Robinson et al., 2014). For resting state fMRI, sICA+FIX was run (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014) to identify and remove spatially specific noise components using a machine learning classifier trained on HCP data. Task fMRI data was processed using a modified sICA+FIX pipeline as described below.

sICA+FIX and a Modified Pipeline for task fMRI: The original sICA+FIX pipeline begins by removing linear trends and means from the fMRI timeseries (mean across time, not space as in GSR) and from the 24 movement regressors (Friston et al., 1996). Then Melodic's implementation of FastICA is run (Beckmann et al., 2005; Beckmann and Smith, 2004), and the FIX machine learning algorithm automatically classifies the resulting sICA components as "signal" or "noise" using the FIX machine learning classifier (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). Then the 24 movement regressors are "aggressively" regressed out of the data and component timeseries (i.e., removing all variance spanning the space of the movement regressors) (Smith et al., 2013a). Finally the sICA components classified as noise are "non-aggressively" regressed out of the data (i.e., removing the unique variance associated with the noise sICA components while leaving in any shared variance also associated with the signal components (Smith et al., 2013a). Importantly, sICA benefits from having a large number of time points per subject to enable finer and more accurate splitting of components, which leads to better separation of signal and noise and easier component classification. Unpublished HCP pilot analyses of running sICA+FIX on single task fMRI runs showed that, while it reduced some false positives, it did not produce a marked statistical sensitivity benefit (improvement in Z scores). Importantly, the shorter the task fMRI run, the more there appeared to be a statistical sensitivity penalty (i.e., lower Z scores). As a result, sICA+FIX had not been previously performed for the released HCP task fMRI data. Here the objective was to study task fMRI data in parallel with resting-state fMRI data using temporal ICA. Therefore, despite above noted caveats, it was essential to process the task fMRI runs similarly to the resting-state fMRI data. Thus, as a part of this study we modified the sICA+FIX pipeline to combine across task fMRI runs to enable cleaning despite the individual runs being substantially shorter in length than the resting state fMRI runs.

As before, linear trends and means were removed from each run's timeseries and 24 movement regressors. Then the data and movement regressors were concatenated across runs within a single task fMRI session (of which there were two in the HCP: Working Memory + Gambling + Motor with 1884 timepoints and Language + Social + Relational + Emotion with 1996 timepoints). Melodic sICA was run on the concatenated timeseries, and the FIX classification and cleanup pipelines were then run on the concatenated data. Finally, the timeseries were split back apart, and the corresponding single run mean images were added back. Using the same classifier training as was used for the resting state fMRI data, we found that the performance of the FIX classifier on the concatenated task fMRI data was similar to what we previously had found with single resting state fMRI runs of 1200 timepoints (99.5% classification accuracy vs manual classification, assessed on the 2 concatenated runs in each of 28 subjects (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). Given that both resting state and task fMRI perform well with the same training, it would be reasonable to combine across resting state and task fMRI data for future studies that acquire less fMRI data, such as the HCP Lifespan Development and Aging projects.

Group-level Spatial ICA Dimensionality Reduction: sICA+FIX cleaned, MSMAll aligned (CIFTI grayordinates) resting state fMRI data were combined across subjects after demeaning and variance normalization using MELODIC's Incremental Group-PCA (MIGP), an iterative, distributed Principal Component Analysis (PCA) data reduction algorithm, to achieve a 4500-dimensional group-PCA, as was done in a prior publication on MIGP (Smith et al.,

2014). This PCA series was then variance-normalized a second time to eliminate spatial differences in unstructured noise variance after the PCA process (Glasser et al., 2016a). The unstructured noise variance was estimated in the first round with Melodic's auto-dimensionality estimator and in the second using the Wishart roll off approach described in (Glasser et al., 2016a) and briefly below. Group spatial ICA was run in the standard grayordinates space at a dimensionality of d=137, optimized according to the method we used previously (Glasser et al., 2016a), whereby the optimal dimensionality was chosen as a local minimum in the number of discrete pieces after a find-the-biggest labeling across ICA component spatial maps. In this case, after the minimum at d=137 the number of pieces began to increase monotonically, likely reflecting splitting of spatially autocorrelated unstructured noise, see (Glasser et al., 2016a).

Individual subject component timecourses were created using weighted spatial regression (Glasser et al., 2016a), a variant of dual regression (Filippini et al., 2009) that helps to further compensate for residual misalignments after MSMAll areal-feature-based registration, to individualize the group components to each subject's resting state and task fMRI data. The last stage of weighted regression involves temporally regressing the individual subject component timecourses into both the grayordinate and volume-based individual subject timeseries to produce individual subject component spatial maps for both CIFTI grayordinate and volume spaces. These spatial maps were then averaged across subjects. No smoothing was applied to either set of maps prior to averaging so as maximize the spatial sharpness of the group averages (Coalson et al., 2017) and the group average volume maps were only used for interpretation of the ICA components outside of the CIFTI grayordinates space to better characterize and examine the contamination of grey matter signal from neighboring sources such as veins, CSF, or white matter. The group PCA, ICA, and individual subject spatial regression steps were all performed in MSMAll aligned CIFTI grayordinate space because cortical grey matter has poor alignment of cortical areas across subjects in the volume (Glasser et al., 2016b; Coalson et al., 2017).

Thus the final outputs for each subject of the spatial ICA dimensionality reduction for both resting state and task fMRI were: (1) a timecourse for each of the d=137 group-ICA components, (2) a grayordinate-based spatial map for each component, and (3) a volume-based spatial map for each component. Additionally, group averages of (2) and (3) were generated and (1) was temporally concatenated across subjects into 137 X ~2 Million timepoint matrices (one for rfMRI and one for tfMRI). Equation #1 summarizes the spatial ICA decomposition (either at the group level for the initial sICA computation or after projecting the decomposition to the individuals using weighted regression) with the $\text{Data}_{\text{Nspace X Ntime}}$ representing either a dense timeseries in individuals or a dense PCA Series from MIGP at the group level. $\text{sICA_Maps}_{\text{Nspace X DsICA}}$ are the sICA spatial maps, $\text{sICA_TCS}_{\text{DsICA X Ntime}}$ are the sICA timecourses (TCS). Error is the subspace left over after the dimensionality reduction of the PCA components prior to ICA (e.g., d=137). The equation also applies to the individual volume dense timeseries and volume-based spatial maps (where space is voxels instead of grayordinates) when the volume dense timeseries is substituted for the last temporal regression stage of weighted regression mentioned above (Glasser et al., 2016a):

$$\text{Data}_{\text{Nspace X Ntime}} = \text{sICA_Maps}_{\text{Nspace X DsICA}} * \text{sICA_TCS}_{\text{DsICA X Ntime}} + \text{Error} \quad (1)$$

Temporal ICA: Temporal ICA was performed on the cross-subject concatenated individual subject sICA component timecourses using a method similar to (Smith et al., 2012) with the FAST ICA algorithm (Hyvarinen, 1999) implemented in Matlab. Although the maximum possible temporal ICA dimensionality was 137 (which is the number of concatenated sICA component timecourses), the temporal ICA decompositions were not reproducible at this dimensionality. We used binary search and the ICASSO (Himberg et al., 2004) algorithm to identify the dimensionality yielding the maximum number of clusters of reproducible components (greater than 0.5 in the ICASSO Iq cluster quality measure, range 0 to 1). We found 70 reproducible clusters for task fMRI and 84 for resting state fMRI. ICASSO was then used to find the component cluster centrotypes of (i.e., the component estimate closest to the cluster center), which were used as an initialization to a final temporal ICA run with FastICA to produce the final temporally orthogonal decomposition. The other FastICA parameter settings were: nonlinear function = tanh, estimation method = symm, and there was no further PCA-based dimensionality reduction prior to the 100 ICASSO iterations or the final ICA decomposition.

The result of the temporal ICA was a set of unmixed timecourses (representing underlying temporally independent latent sources) and a mixing matrix (which represents how to combine the tICA timeseries to recover the estimate of the original sICA component timeseries). Equation #2 describes the group level tICA computation (on concatenated sICA component timecourses) or the single subject decomposition, with $\text{sICA_TCS}_{\text{DsICA X Ntime}}$ representing either the concatenated group sICA timeseries or that of a single subject. $\text{Mixing_Matrix}_{\text{DsICA X DtICA}}$ always represents the tICA mixing matrix computed by FastICA at the group level, and $\text{tICA_TCS}_{\text{DtICA X Ntime}}$ represents the unmixed tICA component timecourses at either the group or individual subject level. The single subject tICA component timecourses are produced by deconcatenating the group timecourses.

$$\text{sICA_TCS}_{\text{DsICA X Ntime}} = \text{Mixing_Matrix}_{\text{DsICA X DtICA}} * \text{tICA_TCS}_{\text{DtICA X Ntime}} \quad (2)$$

To understand how to get spatial maps of temporal ICA components it is necessary to write out the full equation that substitutes Equation #2 into Equation #1 and thus includes both forms of ICA: Equation #3. This equation can be rewritten as Equation #4 by introducing a term that represents the spatial maps of the temporal ICA components $\text{tICA_Maps}_{\text{Nspace X DtICA}}$. At the group level $\text{Data}_{\text{Nspace X Ntime}}$ would be an enormous file of concatenated dense timeseries and thus is never be generated. The purpose of MIGP is to provide a highly accurate approximation of the structured portion of the fully concatenated dense timeseries, as the vast majority of the information left out by MIGP is unstructured noise represented by Error—see Main Supplementary Information on the fMRI Variance Partitioning Model.

$$\text{Data}_{\text{Nspace X Ntime}} = \text{sICA_Maps}_{\text{Nspace X DsICA}} * \text{Mixing_Matrix}_{\text{DsICA X DtICA}} * \text{tICA_TCS}_{\text{DtICA X Ntime}} + \text{Error} \quad (3)$$

$$\text{Data}_{\text{Nspace X Ntime}} = \text{tICA_Maps}_{\text{Nspace X DtICA}} * \text{tICA_TCS}_{\text{DtICA X Ntime}} + \text{Error} \quad (4)$$

Thus, the spatial maps of temporal ICA are given by Equation 5, and these can be either the individual subject spatial maps or the mean spatial maps at the group level:

$$tICA_Maps_{Nspace\ X\ DtICA} = sICA_Maps_{Nspace\ X\ DsICA} * Mixing_Matrix_{DsICA\ X\ DtICA} \quad (5)$$

Thus, the group spatial ICA maps in both grayordinate and volume spaces were multiplied by the temporal ICA mixing matrix, producing grayordinate-based and volume-based spatial maps for each temporal ICA component. These components were ordered according to the RMS of the mixing matrix, and the maximum spatial map value was set by convention to be positive. This imposes a specific order and sign convention on the components, making stronger components that explain more variance appear first and weaker components appear last (similar to how FSL's Melodic orders components), though the order is not exactly by percent variance explained. Because the data used for temporal ICA was the 137 X ~2 Million timepoints concatenated sICA component timecourse matrix that has so many more timepoints than spatial points, group tICA is able to outperform single subject temporal ICA (Smith et al., 2012). Because we have functionally aligned the data, have already cleaned the data of subject-specific, spatially-specific temporal artifacts, and are primarily interested in semi-global and global artifacts, we did not expect the group analysis to be substantially harmed by either residual misalignments or large numbers of uniquely subject-specific artifacts. That said, we did find some single-subject-specific artifacts of both spatially specific and global natures; however, these components did not use up the available degrees of freedom, as the number of reproducible components for resting state (d=84) and task (d=70) temporal ICA were substantially less than the d=137 maximum imposed by the spatial ICA dimensionality reduction.

Physiological Noise Modeling: Physiological data were acquired using a respiratory belt and a heart rate monitor (pulse oximeter on a finger) in most runs from most subjects (87%). We used FSL's PNM tool (Brooks et al., 2008) to convert these traces into 14 physiological regressors including 4 cardiac regressors, 4 respiratory regressors, 4 interaction regressors, a heart rate regressor and an RVT regressor (Respiration Volume per Time; (Birn et al., 2006), see also: https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/PNM/UserGuide). These physiological measures were then compared with temporal ICA component timeseries in those subjects that had both measures acquired. Because of the very large number of physiological traces, the HCP was unable to manually review and quality assure the peak detection of each trace for each subject. Therefore, these data represent a useful but imperfect 'proof-of-principle' measure of subject-specific physiology, and we rely on prior work in a smaller study that shows the extent to which physiological regressors can remove artifacts from fMRI data (Power et al., 2017).

In practice, the RVT trace was found to have the strongest relationship to the data (likely because the more spatially specific, higher frequency artifacts related to the other traces had already been removed by sICA+FIX). Indeed this could be predicted from the group average beta maps of the physiological regressors after only detrending and motion regression (i.e., no sICA noise component removal), as RVT has the strongest global relationship with the fMRI data. Because the quality of the RVT data is variable and there is a variable amount of respiratory signal contamination of fMRI data, we chose to focus on the subjects who had 1) good quality RVT data and 2) a substantial contamination of the fMRI timeseries by respiratory signal. Specifically, we computed the correlations between temporal ICA component timeseries and RVT, but limited that computation to the top 10% of resting state runs or concatenated task sessions that had the highest correlation between

RVT and the sICA+FIX-cleaned timeseries. This set of runs and concatenated sessions was found by correlating RVT with each run or concatenated session's 360 parcellated timeseries from the 360 areas of the HCP-MMP1.0 multi-modal parcellation[1] (Glasser et al., 2016a) and then taking the mean correlation value across parcels. These parcellated timeseries were generated by averaging the dense timeseries within each area.

"Motion" and Other Artifact Modeling: It might seem relatively straightforward to model motion in the fMRI timeseries: Simply realign the data (as a standard preprocessing step) and then use what the realignment algorithm had to do as a measure of motion. Indeed, several movement QC measures have been proposed (e.g., FD (Power et al., 2012), absolute or relative RMS (Jenkinson, 1999) based on motion traces and have proven very useful in traditional fMRI data (with low resolution and slow TR). When the imaging TR is shortened, however, apparent MR image motion can arise from more than just true physical head motion, and the realignment algorithm will still correct for this 'apparent' motion. Respiration (i.e., air entering the lungs) causes changes in the magnetic field that can lead to shifts in the reconstructed image position (as gradients in the magnetic field are used to encode image position (Raj et al., 2001). With imaging TRs shorter than the respiration interval, these field changes will be reflected in the image realignment parameters. This effect was variably present within the main HCP cohort, but importantly it means that movement-trace-based measures such as frame displacement (FD) do not represent veridical measures of movement across all subjects in the HCP data used in this study because they include both real movements and phantom movements related to respiration induced by magnetic field fluctuations.

Another approach to identifying movement is the use of "DVARS" (Burgess et al., 2016; Power et al., 2012; Smyser et al., 2010), which equals the root-mean-square (RMS) of the frame-to-frame image intensity derivatives. DVARS is attractive insofar as what we care most about from an image denoising perspective are sources of image intensity changes, not whether variations in the magnetic field are causing spatial shifts in the image. Such image intensity changes often show up as "spikes" in the raw timeseries DVARS trace. Conversely, however, motion is not the only source of DVARS spikes. These might also be caused by a variety of MR physics etiologies, including radio-frequency interference, coil instability, or other scanner instabilities.

We are particularly interested in those artifacts that potentially influence the data even after sICA+FIX cleanup. After sICA+FIX cleanup, some DVARS spikes disappear entirely, returning the DVARS to its baseline level, whereas others become "dips" instead of spikes. These dips occur because sICA+FIX is removing substantial variance from their corresponding timepoints, above and beyond that which is needed to return the spike to the baseline DVARS value. Indeed this is an inevitable and expected outcome when fully modeling a frame-specific artifact, as not just the artifact, but some or even all of the residuals are also removed. As an illustration in the limit, a simplified regressor in a 10-timepoint timeseries of 0000100000 would remove all of the variance at timepoint 5 and would be essentially equivalent to scrubbing timepoint 5 from the data and replacing it

---

[1]Throughout the manuscript, the "HCP-MMP1.0 multi-modal cortical parcellation" refers specifically to the one derived using the maximum probability map (MPM) in the 210 validation subjects (210V subject set): https://balsa.wustl.edu/file/show/3VLx

with the mean image of the other frames. Indeed, if all of the structured signal is removed from the data (see Main Supplementary Information Supplementary Methods Section on fMRI Variance Modeling), it becomes clear that these DVARS dips occur as a result of removing some of the unstructured noise from these timepoints. Supplementary Figure 1 shows example FD spikes and DVARS dips from subjects with and without respiratory field induced 'apparent' motion and with and without substantial numbers of DVARS dips. This figure shows that FD spikes and DVARS dips may be poorly correlated (Panels 1-3) in the setting of respiration-related magnetic-field-induced 'apparent' motion; however, they are correlated in the absence of this phenomenon (Panels 4-7). Panels 2 and 3 illustrate the similarity between the DVARS of the unstructured noise timeseries and the cleaned timeseries (see Main Supplementary Information Supplementary Methods Section on fMRI Variance Modeling). These DVARS dips reflect time segments in which the sICA+FIX cleanup solution begins to approach the scrubbing solution; however, unlike scrubbing, sICA+FIX does not make an all-or-nothing binary decision based on a chosen threshold (of FD or DVARS), rather it behaves like a weighted ("softer") form of scrubbing, while at the same time handling subtler effects that would be below the scrubbing threshold.

The time periods of the DVARS dips (or if there are any residual spikes) are places where it is important to pay particular attention to ensure that no residual structured noise remains. These DVARS dips will largely be due to subject head motion (as shown in Supplementary Figure 1 rows 4-7), however, they may result from other kinds of punctate signal change as described above. We detect DVARS spikes/dips using a threshold of +/-25 above or below the DVARS baseline of the unstructured noise timeseries after sICA+FIX cleanup (on data scaled to grand mean=10,000), which was manually chosen to provide a balance between oversensitivity (detecting random fluctuations in the baseline) and undersensitivity (missing obvious dips or spikes). We used the unstructured noise timeseries to focus on the fact that unstructured noise variance is being removed by sICA+FIX at these locations; however, Supplementary Figure 1 rows 2 and 3 show virtually identical DVARS traces for the cleaned timeseries and the corresponding unstructured noise if all structured signal is removed.

Examining the Effects of Sleep on Resting State fMRI: In addition to motion, sleep is another subject behavior that potentially has both neural and artifactual correlates. During sleep, respiratory patterns may change, which may lead to more artifactual physiologically driven BOLD fluctuations. In addition, sleeping or drowsy subjects may also exhibit different amounts of motion. Also, arousal state is known to increase the amount of global signal in the brain (Laumann et al., 2017; Liu et al., 2017; Tagliazucchi and Laufs, 2014; Wong et al., 2016; Wong et al., 2013; Yeo et al., 2015), though it is not known to what extent this increase is due physiological noise or true neural effects. The HCP unfortunately was unable to implement eye tracking of its young-adult cohort; however, the personnel operating the scanner were instructed to document when subjects were obviously sleeping. These acquisition logs were extracted from the internal HCP database, and any subject who was ever noted to be sleeping was flagged in the current analysis as a sleepy subject. This allowed explicit comparisons between subjects who had been noted to be sleeping and those who had not. As with the physiology measures discussed above, this an imperfect but highly useful proof of concept metric.

Classification of Temporal ICA Components as Signal or Noise: All temporal ICA components were classified by hand using multiple sources of information, as summarized for each component in the Supplementary Information Sections TC, RC, and TCr. The tICA Spatial Maps (Eq. (5)), both on the surface and in the volume, were the most important source of information, as components could usually be clearly determined as "Neural" based on similarity to known resting state networks or known task activation patterns (in the task fMRI data), or they contained boundaries between positive and negative regions that matched known areal boundaries (Glasser et al., 2016a). Indeed, as has been found with spatial ICA, temporal ICA signal components are usually visually distinct from structured noise components.

Prior work demonstrated that global signal often appears as a pan-grey matter 'stripes' (see section on *Greyplots* below) that are attenuated within white matter and CSF that is more distant from grey matter (Power et al., 2017). Additionally, the global signal after sICA+FIX is known to be a grey matter specific signal (Glasser et al., 2016b). Therefore, we hypothesized that at least one component should have a globally positive spatial map across grey matter in both the resting-state and task fMRI. We also looked for components with significant white matter signal, venous signal, patterns that reflected vascular territories of the brain rather than functional networks, or other spatial patterns not compatible with neural signal.

We computed several quantitative component-wise measures to provide additional support for component classification as signal or noise and to further explore the identities of the components. One measure was the components' direct temporal correlation with RVT. We also focused on component amplitudes, i.e., the standard deviations of the component timeseries, in particular timeseries epochs or between different groups of subjects. For example, we looked for components that had particularly high amplitudes during periods of DVARS dips compared to the non-DVARS dips periods. We also looked at the variability of component amplitudes across subjects, the differences of component amplitudes with subjects who had been noted to be sleeping vs those who had not been noted to be sleeping, and components that were prominent in only one subject, run, or concatenated task session (i.e., where there was a large difference between highest component amplitude and second highest amplitude across runs or subjects). For task fMRI we also looked for component amplitudes that increased during particular tasks relative to the other tasks in order to associate components with specific tasks.

Temporal ICA-based Cleanup: sICA+FIX cleaned timeseries were further cleaned by removing the temporal-ICA components that were classified as noise (i.e., non-neural). All group concatenated tICA component timecourses were split according to single subject concatenated task sessions and resting state runs and were temporally regressed into the dense task or resting state fMRI timeseries data of each concatenated task fMRI session and each resting state fMRI run to compute spatial beta maps. Then the tICA timecourses classified as noise and their associated beta maps were matrix multiplied to determine the portion of the dense timeseries that was best explained by the noise timecourses, and this noise dense timeseries was subtracted from the sICA+FIX cleaned timeseries data, producing the sICA+FIX + tICA cleaned dense timeseries data (see Equation #6 of Main Supplementary Information Section on fMRI Variance Partitioning Model). Note that the inclusion of all tICA timecourses (i.e., both "signal" and "noise") in the multiple regression

for computing the betas makes this a "non-aggressive" cleanup similar to FIX, leaving any variance associated with the tICA signal components intact (in practice, however, the issue of components having non-orthogonal variance will be a much smaller effect for tICA than sICA, because tICA component timecourses are defined to be temporally orthogonal at the group level and this property will largely persist at the individual run or concatenated task session level).

Generation of Greyplots: Greyplots, which display the timeseries intensities in grey scale with some compressed form of space on the y axis and time on the x axis, have proven to be a useful method of visualizing the spatio-temporal structure of fMRI data and in particular have been used to highlight global fluctuations in timeseries data (Power, 2017; Power et al., 2014; Power et al., 2017). The greyplots used in this study were generated for the data after each cleanup step by within-parcel averaging of the timeseries using the HCP-MMP1.0 multi-modal parcellation (Glasser et al., 2016a). This enables structured patterns in the data to be more easily seen (as unstructured noise is averaged out), but avoids obscuring features by spatial or temporal smoothing of the data. Because simply displaying the 360 parcellated timeseries will bias the resulting greyplot image towards the smaller areas and away from the larger ones by giving smaller areas more space than their size would dictate (relative to an unparcellated greyplot), the smallest area (by surface area in mm) was given one row of the greyplot and larger areas were given proportionally more rows based on surface area. Additionally, because semi-global signals will appear more global if randomly mixed along the spatial (i.e., y) axis of the grey plot, cortical areas were clustered according to the group-average full correlation resting state functional connectome (see below) computed after tICA cleanup, so that areas with more similar timeseries are placed closer together. This also makes it easier to see the neurobiological structure in the data as like rows are averaged with like. All plots have these transformations applied to them identically and then Matlab's image downsampling produces the final rasterized image.

Mean Grey Timecourse Regression: The global signal (across grey, white, and CSF), after sICA+FIX cleanup, is known to be a grey matter specific signal (see Supplementary Figure 5 in Glasser et al 2016 (Glasser et al., 2016b). Furthermore, it is highly correlated with the mean grey signal (r=0.98+/-0.01 for the n=449 resting state subjects used in this study, see also (Power et al., 2014). Additionally the mean signal across parcels (Mean Parcel Timeseries) is very highly correlated with the MGT (r=0.99+/-0.006). We used Mean Grey Timecourse Regression (MGTR)—the mean signal across grey matter within the CIFTI grayordinates standard space—or the MPT—the mean signal across the parcels—for the parcellated connectome analysis instead of the mean across the whole-brain mask (as has traditionally been done for global signal regression). For the group dense connectome and gradient analyses we regress the Mean PCA Series (MPS) out of the MIGP PCA series, which is a close approximation to MGTR on the full concatenated across subjects dense timeseries (Smith et al., 2014). All other analyses use grayordinate-wise regression of the MGT. We made these choices largely for methodological convenience, as the different global timeseries are very similar. We use the specific term MGTR in the methods and results for this study, but use the more general and well-known term Global Signal Regression (GSR) in the introduction and discussion when speaking about the technique in general.

Generation of Task fMRI Statistical Maps: After each cleanup methodology (including no cleanup, sICA+FIX, MGTR, sICA+FIX + MGTR, and sICA+FIX + tICA), task fMRI dense timeseries data were processed as in Glasser et al 2016 (Glasser et al., 2016a). Briefly, single subject statistical maps were computed natively in CIFTI grayordinate space on each run using FSL's surface-compatible FILM GLM tool (Woolrich et al., 2001), were combined across runs within subjects using fixed effects and FSL's FLAMEO tool, and then were combined across subjects using mixed effects and FSL's FLAME1 tool (Woolrich et al., 2004). Beta maps were corrected for intensity bias field computed from a combination of spin echo and gradient echo images (Glasser et al., 2016a) and generated on the full 449-subject dataset (intensity bias field correction scales beta maps so that a given value has the same biological meaning across the whole brain).

Cluster mass (the sum of above threshold z values) was computed using a Z=+/-5 threshold (this happens to be roughly the two tailed Bonferroni corrected significance level across the 91282 grayordinate space). Because cluster mass thresholds and thereby results are dependent on the number of subjects used (because Z is dependent on n, as standard error decreases with increasing n), we continued to use the 28-subject registration optimization dataset from (Glasser et al., 2016a) for cluster mass computation. This dataset is a subset of the 449 subjects that were used in both studies and was originally chosen because it has examples of high and low SNR and both typical and atypical areal topologies.

Wishart roll off reconstructed data were generated as described below in the next section (on data concatenated across all task fMRI runs within each subject, at which point the Wishart roll off was run, and then the reconstructed concatenated timeseries were split back into their constituent runs and the task analysis performed). Parcellated task fMRI data were generated by first parcellating the dense timeseries (by averaging within each cortical area of the HCP-MMP1.0 multi-modal cortical parcellation) and then running the data through the same task analysis procedure described above (without the surface-based smoothing of the temporal autocorrelation estimates) and as described in Glasser et al 2016 (Glasser et al., 2016a). Hence, in the parcellated analyses, the within parcel averaging was applied prior to fitting the task GLM or computation of statistics, which leads to gains in statistical sensitivity (Glasser et al., 2016a).

Generation of Functional Connectomes: After a specific cleanup approach (sICA+FIX, sICA+FIX + MGTR, and sICA+FIX + tICA), we generated dense functional connectomes from resting state fMRI data using the approach described in Glasser et al 2016 (Glasser et al., 2016a). We did not compare conditions with and without sICA+FIX (or for example MGTR only) because it is already well established that sICA+FIX is strongly beneficial for HCP-Style resting state fMRI data (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014; Smith et al., 2013a; Smith et al., 2013b). Briefly, data from each resting state fMRI run were demeaned and variance normalized according to the standard deviation of the unstructured noise timeseries. Then the data from the 210P group of subjects (Glasser et al., 2016a)) were (i) run through the MELODIC's Incremental Group-PCA (MIGP) iterative, distributed PCA data reduction algorithm (Smith et al., 2014) (ii) variance normalized a second time according to the amount of unstructured noise found by the Wishart roll off approach; (iii) PCA was rerun; and then (iv) a final Wishart roll off was done to smoothly downweight unstructured noise without spatially or temporally smoothing the neural

signal (Glasser et al., 2016a; Glasser et al., 2016b). The Wishart roll off approach determines the portion of the PCA eigenvalues of the data that are above a Wishart distribution of unstructured noise with similar spatial and temporal autocorrelation properties and then subtracts this fitted unstructured noise distribution from the eigenvalues to downweight the effects of unstructured noise relative to the structured signal in the data (Glasser et al., 2016a; Glasser et al., 2016b). Dense functional connectivity matrices were computed from this reweighted PCA series, and functional connectivity gradient maps were computed from the result for each cleanup approach (Glasser et al., 2016a). Parcellated connectomes in Figure 11 were generated by averaging the dense PCA series within the parcels of the HCP-MMP1.0 multi-modal cortical parcellation and computing the full correlation values. Parcellated connectomes in Figure 12 were generated in individual subjects by computing the full covariance values, Fisher Z-transformed fully correlation values, or Fisher Z-transformed partial correlation values using FSLNets and then averaging across subjects. The full correlation temporal ICA cleaned connectome was clustered using hierarchical clustering implemented in FSLNets (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLNets) to order the cortical areas according to similarity of connectivity and enable grouping the areas into clusters with similar connectivity.

Data sharing. The data and annotations used in each brain image figure are stored as Connectome Workbench scenes and will be uploaded to the BALSA neuroimaging results database (http://balsa.wustl.edu) upon final publication. Scene-specific identifiers in each figure legend will allow easy previewing of scenes in BALSA and immediate download of individual scenes, complete scene files, or individual data files (Glasser et al., 2016b; Van Essen et al., 2017).

Results

We will discuss task fMRI data first because subjects' underlying neural activity is explicitly manipulated by the task, and thus we have a well defined hypothesis about what at least part of our subjects' neural BOLD activation should look like, allowing objective tests of how well that hypothesis is matched by the data after different cleanup approaches. In addition, subjects who are performing a cognitive task will be less likely to fall asleep, and the scanner technicians were trained to stop the scan if the task performance became unacceptable. We will examine how task-induced neural BOLD activation manifests itself in temporal ICA, revealing important properties of this relatively unexplored method in fMRI data. We then compare and contrast task fMRI data with resting state fMRI data where we do not have the benefit of a prior hypothesis about our subjects' neural BOLD activation and where subject behavior is also likely less well controlled. In this way, insights gained from the task-based analyses can directly inform our interpretation of the more challenging resting state analyses. Within the task fMRI section, we first explore the temporal ICA components and their properties, then show the effect of temporal ICA cleanup and MGTR on grey plots, and finally discuss the effects of MGTR, sICA+FIX, sICA+FIX + tICA, and sICA+FIX + MGTR on the task GLM contrast maps themselves. Within the resting state fMRI section, we like-wise first explore the temporal ICA components and their properties and then discuss data cleanup.

Exploration of Temporal ICA Components in Task fMRI Data

We identified 70 reproducible temporal ICA components in the task fMRI data (reproducibility as assessed by ICASSO) and ordered them approximately according to temporal variance explained (See Supplementary Task Components (TC)). The strongest of these components (6.7% of the tICA explained variance) was globally positive in grey matter, though with less intensity in regions with lower GRE signal intensity (i.e., gradient echo fMRI dropout regions), near zero in the white matter, and negative within the ventricles (Figure 1, Supplementary Figure TC1). This component also had the highest temporal correlation with the respiratory measure, RVT (r=0.30), of all 70 components (Panel A, Figure 2). Further, the component exhibited high cross-subject variability in amplitude (Panel B, Figure 2) and has higher amplitude during DVARS dips (Panel C, Figure 2; see also Supplementary Figure 1). The grey matter specific spatial pattern and correlation with RVT of this component suggests that it likely largely reflects global changes in grey matter blood flow from physiological sources, such as those arising from variations in breathing depth and rate, end tidal CO2, and/or heart rate (Birn et al., 2006; Chang et al., 2009; Chang and Glover, 2009; Golestani et al., 2015; Power, 2017). Though it is correlated with DVARS dips, this may be because motion may be more likely to co-occur with changes in respiration (Power, 2017). This component may partly or completely explain the global artifact that has been shown to persist after sICA+FIX in HCP resting state fMRI data (Burgess et al., 2016; Power, 2017; Siegel et al., 2017). Indeed its spatial pattern reflects what many practitioners of global signal regression believe that they are removing from their data (though see below Figure 12, and (Glasser et al., 2016b) Supplementary Figures 5 and 6 for what is actually being removed by GSR).

One initially puzzling feature of this component is the anti-correlation between the ventricles and the grey matter. This kind of anti-correlation is also seen in other global or semi-global components that are categorized as noise, signal, or even are specifically task modulated. We believe this anti-correlation is related to changes in global or semi-global blood flow in the brain due either to physiological or neural effects. When there is more blood flow (i.e., increase in T2* image intensity in grey matter), the brain tissue expands and this pushes CSF with bright signal on T2* images out of the ventricles. This would lead to a decrease in ventricular signal intensity, particularly around the ventricular margins. Conversely, when blood flow drops, the CSF volume is restored and the ventricular signal intensity around the margins would increase (Bright et al., 2014; Thomas et al., 2013). Additionally, flowing CSF will also have a lower T2* image intensity because of dephasing in the centers of CSF spaces (Chen et al., 2015; Yildiz et al., 2017). The relationship between T2* BOLD activity, blood flow, and CSF volume has been shown in previously cat visual cortex in response to visual stimulation (Jin and Kim, 2010). Thus, we do not believe that anti-correlation in the CSF by itself is an indicator of a component being either noise (i.e., of physiological origin) or signal (i.e., of neural origin), but rather is a function of how widespread the blood flow reflected in a given component is.

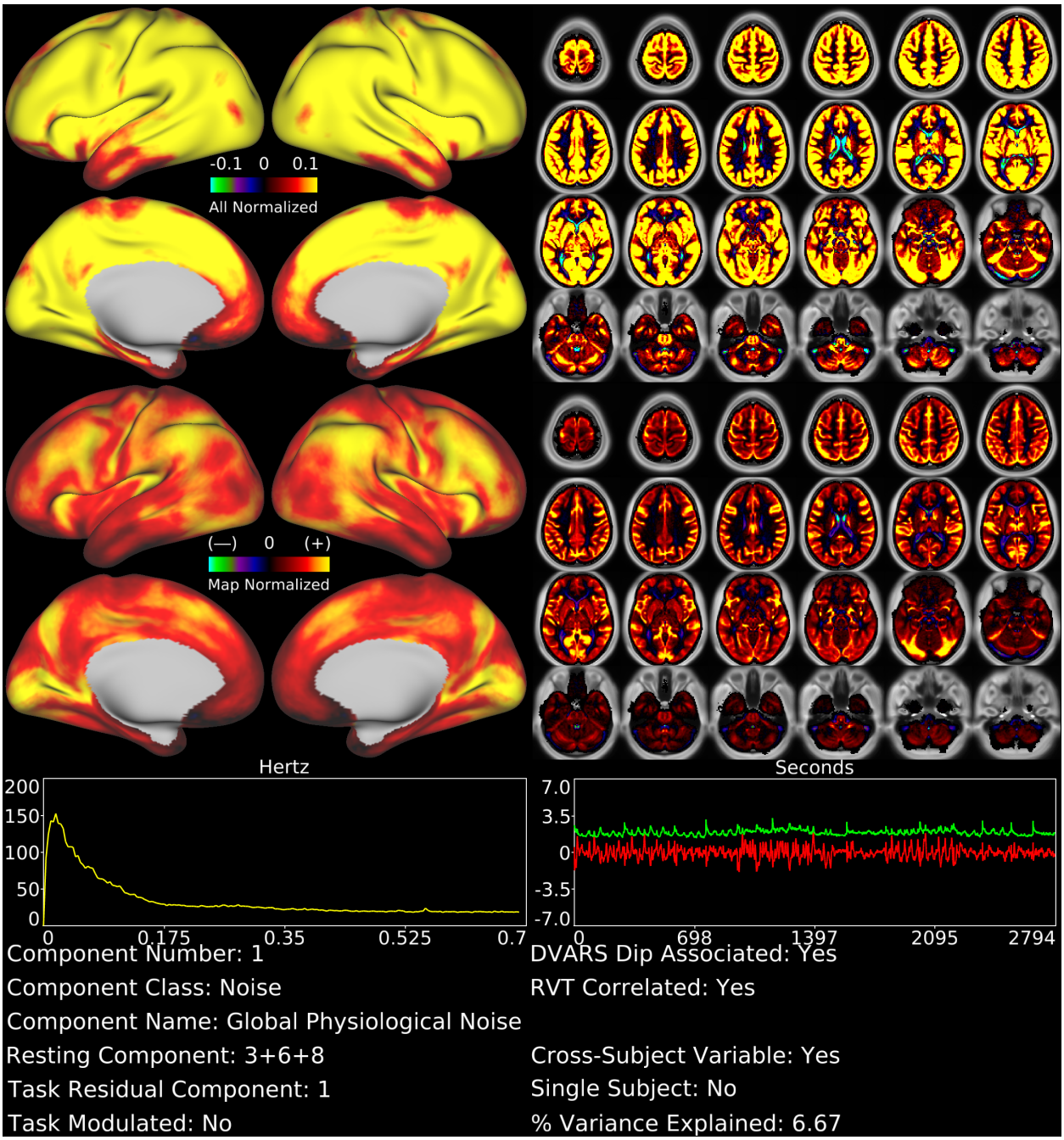


Figure 1 shows the first task fMRI tICA component in a standardized view (see also Supplementary Figures TC 1-70). The top row of data has the color scale normalized and held constant across all 70 components (in percent BOLD), whereas the second row has the color scale set independently for each component. Hence the first view allows components to be compared with each other on the same scale, whereas the second view highlights patterns in the map of each specific component. The left chart in the third row indicates the power spectra of the component averaged across subjects, and the right chart indicates the average timeseries (red) and average absolute value of the timeseries (green). Both of these should show evidence of the task stimulus for task-modulated components because of consistent task timing across subjects. Additional information about the component is provided on the bottom row (see Figures 2 and 3 for thresholds that determine

Yes/No status). Supplementary Figure 2 shows how the task fMRI runs were concatenated. Percent variance explained is the percent of the total variance explained by the tICA model.

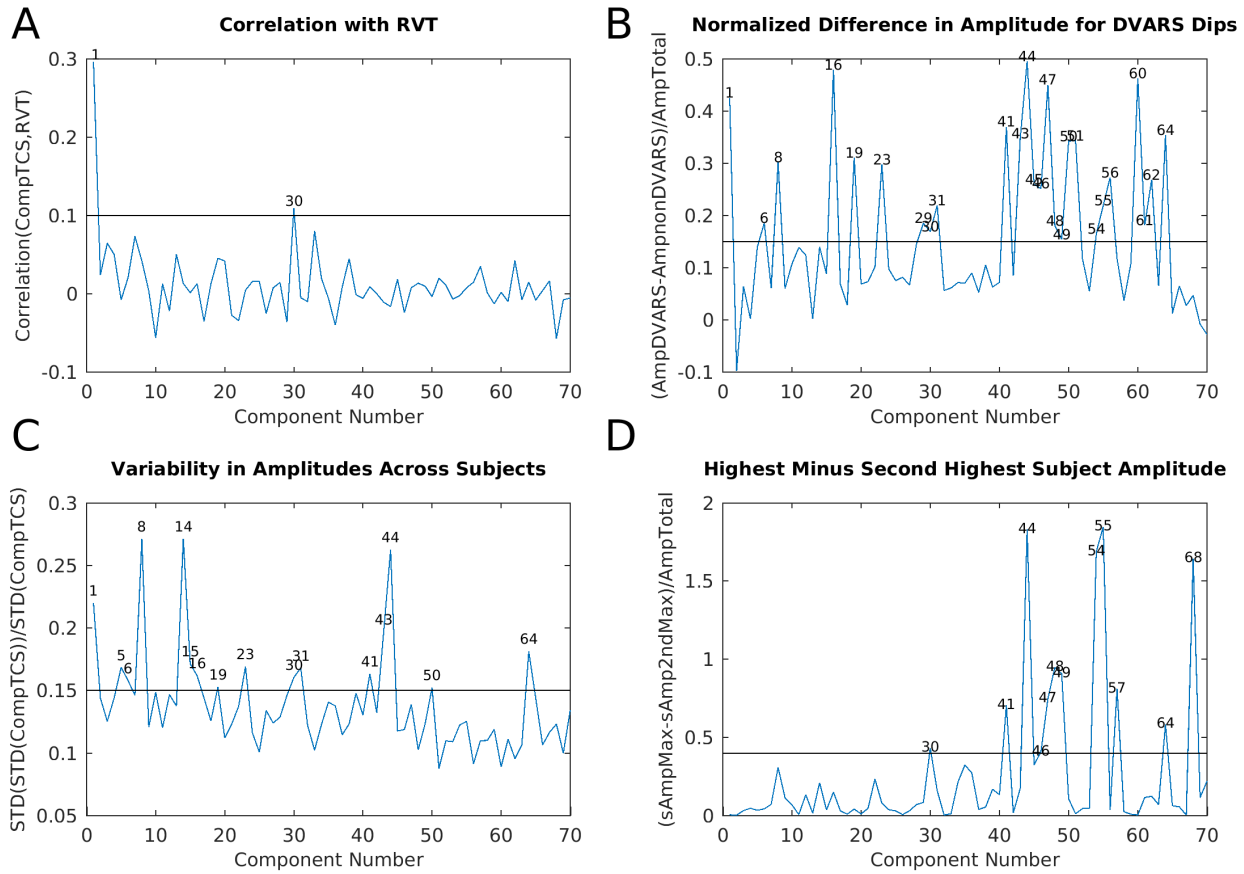


Figure 2 shows four plots that were helpful during task fMRI component classification into signal and noise. Panel A is the correlation between the component timeseries and RVT for the 70 concatenated sessions that had the top 10% mean correlation between RVT and their parcellated fMRI data. This selects for subjects having both good quality RVT traces and substantial respiratory contamination of their data (see methods). The line is at r=0.1. Panel B is the difference in component amplitude (standard deviation of the component timeseries) between frames with DVARS dips and those without DVARS dips normalized by the component amplitude across all frames (see methods), so as to highlight those components that have stronger temporal fluctuations during DVARS dips. The line is at 0.15. Panel C shows the variability of component amplitudes across subjects normalized by the overall amplitude of each component. The line is at 0.15. Panel D shows the difference between the maximum subject's component amplitude and the next highest subject's component amplitude normalized by the overall amplitude of each component. This measure highlights those components that are particularly strong in a single subject. The line is at 0.4. The discriminatory thresholds in this figure and in Figure 6 and Supplementary Figure 2 were chosen empirically to divide the components into useful groups with similar properties. Those components above each threshold are numbered on each graph.

Together with the global noise component, we identified a total of 25 components having properties consistent with structured noise (Supplementary Figures TC1, 8, 30, 38, 41-52, 54-57, 59-60, 62, 64, and 68), which together accounted for 26.9% of the tICA explained

variance (notably this is in strong contrast to sICA+FIX for the tfMRI data where 89% of structured variance was noise). The list of noise components includes those showing similar patterns to the main global physiological component (Supplementary Figures TC30, 38, and 52), substantial extensions into white matter (Supplementary Figures TC8, 30, 42, 50-52, 60, and 64), inclusion of veins (Supplementary Figures TC8 and 62), similarity to movement regressor beta maps (Supplementary Figure TC51), high values around brain margins (Supplementary Figures TC38, 44, 46, 47, 48, and 52), similarity to known image reconstruction artifacts (Supplementary Figures TC41, 43, 45, 49, 54, 55, 56, and 59) or receive coil instabilities (Supplementary Figures TC50, 60, and 64), appearance of banding patterns (Supplementary Figures TC 56), amplitudes that are correlated with DVARS dips (Supplementary Figures TC8, 30, 41, 43-51, 54-56, 60, 62, and 64), and prominence only in one subject or one run (Figure 2, Supplementary Figures TC30, 41, 44, 46-49, 54, 55, 57, 64, and 68). While sICA+FIX classification performance exceeds 99% accuracy on HCP-Style data (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014) some spatially specific noise does slip through (either due to the rare misclassifications or due to being in the null space of the sICA dimensionality reduction), and this structured noise is likely the source of some of these components. As shown below, this residual spatially specific structured noise is removed along with the global structured noise by tICA cleanup.

The other 45 temporal ICA components had properties consistent with neural signal (Supplementary Figures TC2-7, 9-29, 31-37, 39, 40, 53, 58, 61, 63, 65-67, 69-70) and accounted for the other 73.1% of the tICA explained variance. The upper panel in Figure 3 shows difference in amplitude for each task relative to all other tasks, with each of the 7 tasks in a different color, revealing the components that are specifically modulated by particular tasks. The amplitudes of 18 of the signal components were clearly modulated by a specific task (being stronger in one or more tasks than the other tasks, Supplementary Figures TC2-5, 7, 9, 11, 13, 18, 19, 21, 24, 28, 31, 33, 37, 58). In striking contrast, none of the 25 noise components' amplitudes were clearly modulated by a specific task. Many of these 18 task-modulated components look very similar to their corresponding task fMRI GLM contrast beta maps (Supplementary Figures 3-13), which is to be expected if temporal ICA is sensibly decomposing the task fMRI data into meaningful temporally orthogonal components.

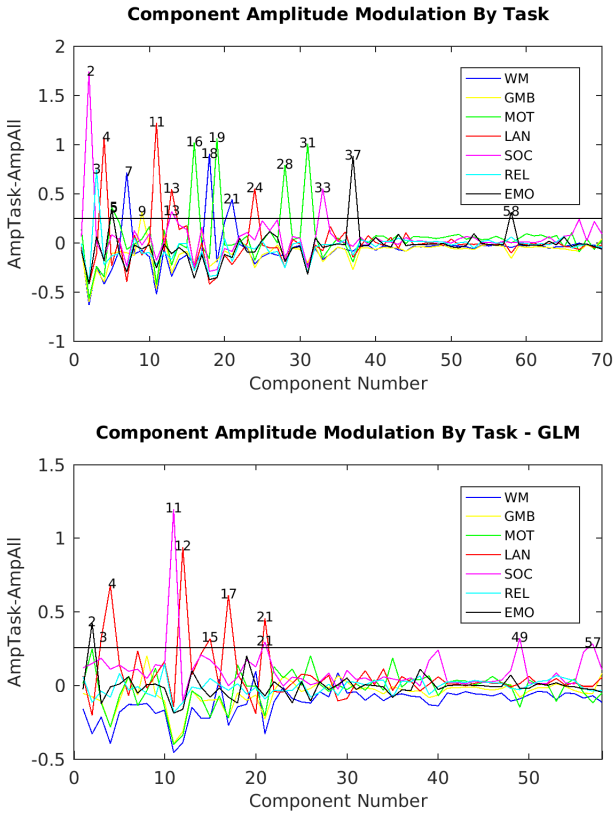


Figure 3 shows the component amplitudes modulated by task. The measure is the amplitude of a component (standard deviation over time) during a given task versus all other tasks. Task abbreviations are WM=Working Memory, GMB=Gambling, MOT=Motor, LAN=Language, SOC=Social, REL=Relational, and EMO=Emotion. The top panel shows the original task component amplitudes (70) while the bottom panel shows the component amplitudes after regressing out the task GLM (58). The line is at 0.25 in both cases.

A supplementary analysis was performed on the task fMRI data by rerunning weighted regression and temporal ICA on the residuals of regressing out the task fMRI GLM design matrix from each task run's dense timeseries. 58 reproducible temporal ICA components were found (See Supplementary Task residual Components (TCr)). 38 were classified as signal, accounting for 71.7% of the tICA explained variance, and 20 were classified as noise, accounting for 28.3% of the tICA explained variance, with TCr1, the main global noise component, accounting for 7.9% of the tICA explained variance (Supplementary Figure 14, Supplementary Figures TCr11-58). Eleven of the signal components' amplitudes remained modulated by specific tasks (Figure 3, lower panel), suggesting that regressing out the HCP's task GLM design matrix does not completely eliminate the effect of task on the underlying spontaneous fluctuations. Additionally, some of these task-modulated components are not present during resting-state (see below).

Five of the 45 signal components (Supplementary Figures TC6, 16, 19, 29, and 31) in the main task fMRI temporal ICA analysis represent topographically organized sensori-motor networks for the head, right hand/upper extremity, left hand/upper extremity, eyes/neck/trunk, and the feet/lower extremities. These components all include nodes in the primary sensori-motor cortex (M1 and S1), the supplementary sensory cortex (SII), the insular cortex, the supplementary and cingulate motor cortices, superior and inferior cerebellar motor areas, the thalamus, and the striatum (in addition the head sensori-motor component also appears to include brainstem cranial nerve nuclei). Notably, the hemispherically lateralized hand networks show correct hemispheric specificity (contralateral to the body part they control, for all nodes except the cerebellum, which has ipsilateral nodes). All but the left hemisphere right hand sensori-motor network (TC6) and the eye/neck/trunk network (TC29) are specifically modulated by the MOTOR task. That the left hemisphere right hand motor component is not specifically modulated by only the motor task is expected because the right hand was used in the other tasks for button box pressing and thus was likely heavily engaged in all of them. The motor task also did not explicitly ask the subject to move their eyes, neck, or trunk. There is one additional predominantly cortical network (Supplementary Figure TC23) that spans all of S1/M1, SII, M2 and auditory cortex that is not specifically modulated by task.

All six of these sensori-motor networks have higher amplitude during DVARS dips, a property shared with only one other signal component (Supplementary Figure TC61), but with many of the noise components. Thus, these sensori-motor networks share two interesting properties: 1) They are modulated by a MOTOR task and 2) They have higher amplitude during DVARS dips, which are usually the result of head motion (see methods). Indeed, the MOTOR task has the highest rate of DVARS dips per frame of all the tasks at 0.029 dips per frame vs the average across other tasks of 0.016 dips per frame (resting

state has 0.013 dips per frame). We can thus infer that during periods of DVARS dips, subjects are using their sensori-motor networks more than usual. Though there will be the usual period of hemodynamic delay, neural activity in the sensori-motor networks may precede movement (e.g. the subject trying to stay still but having a strong urge to move), and movements are often clustered together meaning that BOLD effects of prior movement may still be present (this was explicitly true in the blocked motor task design). These results lead to the conclusion that we should not assume that the neural signal of subjects during DVARS dips is the same as during non-dip periods. The impact of these findings on the methods that we chose to use to measure data cleanup quality for this study is discussed in the supplementary discussion, and the impact on some study designs is discussed in the main discussion.

Several other interesting properties are present within the 45 signal components of the task fMRI temporal ICA analysis. Many networks that are reminiscent of canonical resting state networks are both present and less modulated by specific tasks. This includes components representing the default mode network (Supplementary Figures TC5, 12, 25-27, 32, 40, 53, and 61), the cingulo-opercular network (Supplementary Figure TC10), the language network (Supplementary Figure TC22), and the visual system (Supplementary Figure TC34). On the other hand, some signal networks do not clearly match canonical resting state networks, e.g., Supplementary Figures TC35, 36, and 39 (Cerebellum), 63 (left vs right network), and 66 (extra-striate visual network).

Interestingly, four signal networks show visuotopic organization with respect to polar angle, with positive vs. negative separations across the horizontal (Supplementary Figures TC69 and 70) and vertical meridians (Supplementary Figures TC65 and 67). These networks also have some visuotopic organization relative to eccentricity (foveal vs peripheral), as do three other signal components (Supplementary Figures TC15, 17, and 58). We suspect that these temporal ICA components underlie the ability to extract visuotopy from fMRI data that does not contain an explicit visuotopic task (Figure 8 in the Supplementary Methods and Figures 3, 4, 5, and 6 of the Supplementary Neuroanatomical Results, both from Glasser et al 2016 (Glasser et al., 2016a)

Finally, one component (Supplementary Figure TC 14) highlights specific interactions between V1 and the LGN that may account for the ability to sharply delineate area V1 using functional connectivity in our prior parcellation (Figure 2 in the Supplementary Neuroanatomical Results in Glasser et al 2016 (Glasser et al., 2016a)), as the V1 boundary is not present in any of the other components. We suspect that this component is related to whether a subject's eyes are open or closed, as identifying V1 as a separate parcel using automated winner-take-all approaches depends on whether a subjects' eyes are open or closed (Laumann et al., 2015). Interestingly, this component is variable across subjects, perhaps reflecting differences in subject eyes open vs eyes closed behavior in the scanner despite the presence of a task. Other signal components that showed high cross subject variability included the left hand sensori-motor component (TC31), the pan-motor and auditory component (TC23), and the strongest default mode component (TC5).

We also correlated the component amplitudes (temporal standard deviation of each component in each concatenated task session—in this case an entire tfMRI session on day 1 or day 2 for each subject). For each component, there were therefore 2 * N_subjects amplitudes, and those amplitudes were cross-correlated across tICA components to generate the correlation of component amplitudes. We then performed hierarchical clustering on the resulting correlation matrix (Supplementary Figure 15). In general, signal components cluster together with other signal components and noise components cluster with other noise components. Not surprisingly, components that are strongly modulated by task and whose tasks occurred on a particular day cluster together and are anti-correlated with the components whose tasks occurred on the other day, likely because they have strong amplitudes on one day and near zero on the other. Between these are spontaneous components that are generally not specifically task modulated. For the residuals after regressing out the tfMRI GLM designs, again signal and noise components tended to cluster together (Supplementary Figure 16). Components with residual task modulations tend to cluster together and spontaneous components generally form a separate cluster.

Finally, we computed the betas of each tICA component on the mean grey timecourse global signal after sICA+FIX cleanup to see which components make the greatest contribution to the global signal. Not surprisingly, this shows that the global components indeed are the biggest contributors (Supplementary Figure 17). Additionally, we looked at the effect of tICA clean up on the mean grey signal component betas.

Effects of Temporal ICA Cleanup and MGTR on Task fMRI Greyplots

Figure 4 shows mean grey signal traces and parcellated greyplots of two concatenated task sessions of two subjects who had particularly high global noise, after sICA+FIX under three cleanup conditions: after sICA+FIX only, after sICA+FIX plus temporal ICA cleanup, and after sICA+FIX plus MGTR. Regressing out the components identified as noise removes the obvious bands in the greyplots (and the corresponding substantial mean grey signal fluctuations) that have been previously associated with respiratory changes (Power et al., 2017). Importantly, regressing out the noise temporal ICA components is not the same as removing the mean grey signal, as the mean grey timeseries retains some fluctuations after temporal ICA cleanup, unlike (by definition) with MGTR. In task fMRI data, the mean grey timecourse variance averaged across all subjects (MGTRVar, the amount of variance removed by regressing out the mean grey timecourse in grand mean 10,000 scaled data) before sICA+FIX cleanup is 2161, after sICA+FIX it is 482, and after temporal ICA cleanup it is 235. Thus, in task fMRI data, the neural global signal variance is 11% of the original global timecourse variance and 49% of the global timecourse variance after sICA+FIX. Across subjects, the global timecourse variance is substantially reduced by tICA cleanup (Supplementary Figure 17).

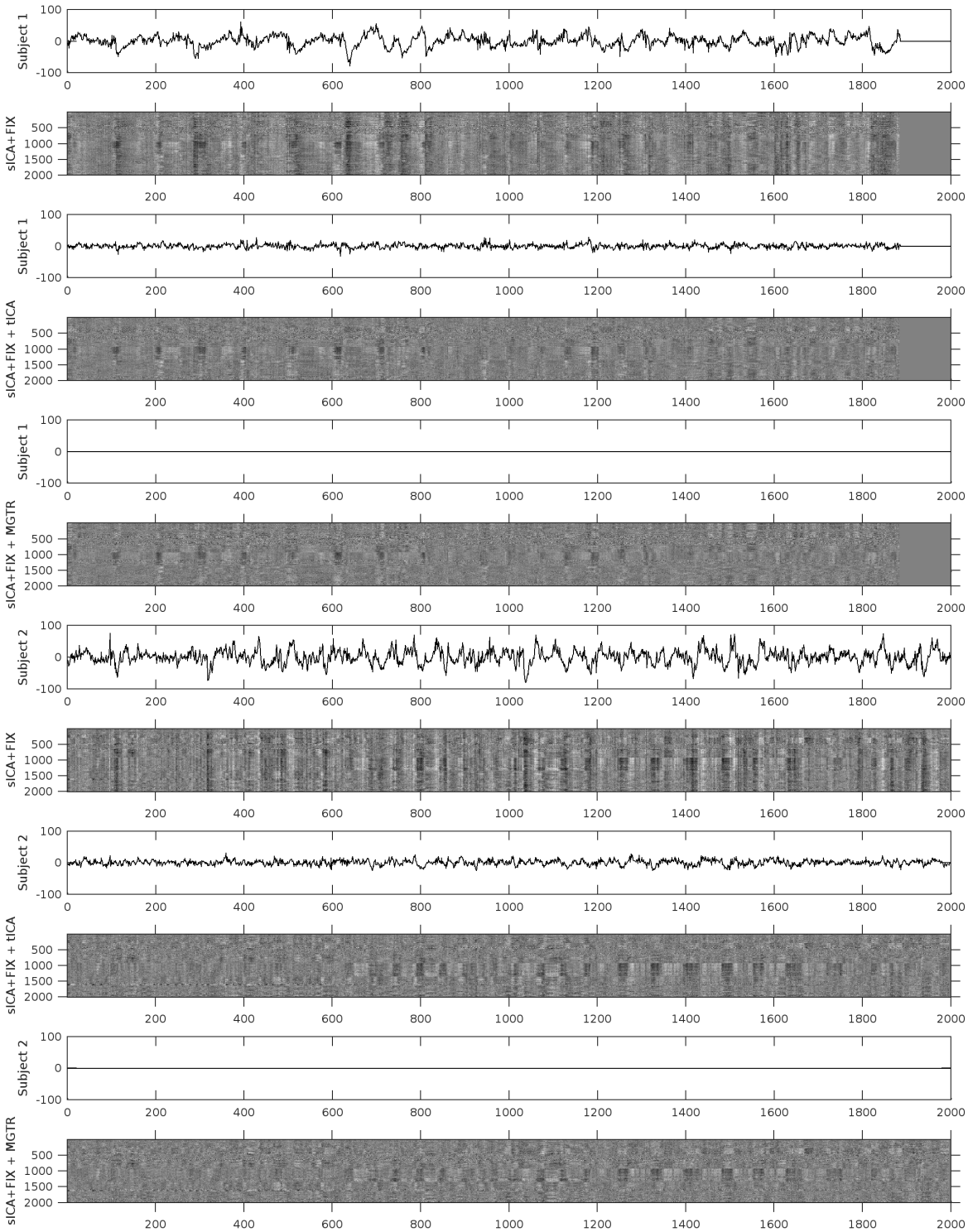


Figure 4 shows the mean grey signals and parcellated grey plots of two concatenated task fMRI timeseries from two subjects after sICA+FIX (Rows 1-2, 7-8), sICA+FIX + tICA (Rows 3-4, 9-10), and sICA+FIX + MGTR (Rows 5-6, 11-12). Note that concatenated session 1 is shorter than concatenated session 2 and so the first 3 rows have zero padding. The data were parcellated, then displayed according to parcel mean surface area (with 1 row for the smallest parcel and proportionally larger numbers of rows for larger parcels such that there are more than 360 rows).

Additionally the data are ordered by hierarchical clustering of the group full correlation parcellated connectome so that parcels with more similar timeseries across the group are closer together (see the bottom row of Figure 12 which shows the "cognitive/task negative vs non-cognitive/positive split that forms the main discontinuity about half way down the y-axis in the grey plots). tICA cleanup removes the vertical "stripes" (signal deviations of the same sign across the whole brain) from the grey plots that were present after sICA+FIX and their corresponding MGT fluctuations without removing the entire MGT as occurs with MGTR. The greyplots after tICA cleanup and MGTR look very similar, though close inspection reveals evidence of subtle biases induced by MGTR (e.g., timepoints with strong dips in non-cognitive/positive locations have this dip reduced, and at the same timepoints cognitive/task negative locations have their the intensity increased slightly relative to tICA). The scale ranges from -2% to +2% BOLD.

Effects of sICA+FIX Cleanup on Task fMRI GLM Contrast Maps

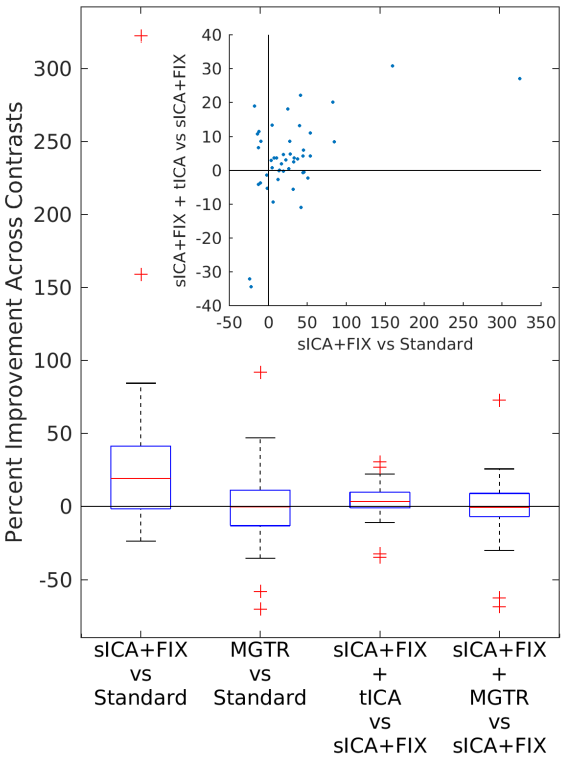
We found a clear statistical advantage of cleaning task fMRI data with sICA+FIX when the data were concatenated across multiple tasks so as to provide a sufficient number of timepoints for sICA+FIX (see Methods) with 62/84 contrasts increasing in cluster mass vs standard MSMAll-aligned CIFTI analysis (Figure 5) and an overall improvement of 10.3% (Supplementary Figure 18).
Cluster mass did not increase for all task contrasts (Figure 5), however, and in several cases it appears that stimulus-correlated noise was responsible for biases or false positives in the group effect size maps. For example, Supplementary Figure 19 shows a neurobiologically implausible false-positive deactivation in orbitofrontal cortex during tongue movement that is removed by sICA+FIX cleanup, and Supplementary Figure 20 shows where a negative bias in the Language STORY contrast is removed (see the strongly negative CSF), revealing an activation that would otherwise have been near zero. The false positive tongue movement feature is likely caused by stimulus correlation head motion given its localization to the orbitofrontal cortex, a site of strong susceptibility vs movement interaction. The language story contrast bias is likely caused by inadequate T1 steady state during the only baseline period that happens to be located at the very beginning of the task, which affects bright CSF signal most. Supplementary Figure 18 shows the percent change in overall task fMRI mixed effects cluster mass across all contrasts after 10 different analyses relative to the standard, MSMAll-aligned CIFTI analysis, showing substantial cluster mass improvements for methods that reduce unstructured temporal noise such as the Wishart roll off and parcellation.

Effects of MGTR on Task fMRI GLM Contrast Maps

In contrast to sICA+FIX applied to task fMRI data, mean grey timecourse regression actually reduced task statistics by 13.1% overall (Supplementary Figure 18). This reduction occurred frequently across contrasts with half (42/84) losing statistical sensitivity after MGTR vs no cleanup and a majority (45/84) losing statistical sensitivity from the use of MGTR after sICA+FIX (Figure 5 left hand panel) compared to sICA+FIX alone. It also consistently shifted the mean of the activation beta maps across all tasks towards zero, which would be highly improbable from a neurobiological perspective (Figure 5 right hand panel). Thus, MGTR evidently removes a portion of the expected neural signal of fMRI data. These properties of global signal regression (or similarly mean stabilization) have been

reported previously for task fMRI data (Aguirre et al., 1997; Aguirre et al., 1998; Macey et al., 2004; Zarahn et al., 1997), but they did not prevent the subsequent widespread use of global signal regression as a cleanup technique for resting state fMRI data (perhaps because the effects were perceived as modest).

That said, 26% of the global neural signal variance (i.e., the mean grey timecourse) after sICA+FIX + tICA cleanup is task related and 55% of the overall neural signal variance (i.e., the variance of the BOLD portion of the dense timeseries across time and averaged across space, see Supplementary Information Section on fMRI Variance Partitioning Model) after sICA+FIX + tICA cleanup is task related. These measures are computed by comparing MGT variance and overall BOLD variance before and after regressing out the task fMRI GLM design matrix. The proportion of overall neural signal and global neural signal explained by the task GLM does vary across tasks, however, with the Working Memory task being highest in terms of overall neural signal explained at 70% overall / 31% global, then the Motor task at 67%/34%, the Gambling task at 53%/36%, the Relational task at 53%/36%, the Social task at 51%/35%, the Emotion task at 46%/13%, and the Language task at 41%/10%. The overall proportion of the neural signal that is explained by the task GLM will vary based on how well the task drives the BOLD fluctuations (to the exclusion of or over and above the spontaneous fluctuations) and how well the task GLM model explains the task induced variance. The global neural variance will depend on how much synchronous activity the task generates across the brain and the amplitude of this activity (particularly in the early non-cognitive/task positive areas that generate most of the global neural signal). The portion of the overall neural signal variance that is global in task fMRI data is 10% (i.e., the portion removed by MGTR after sICA+FIX and tICA cleanup).



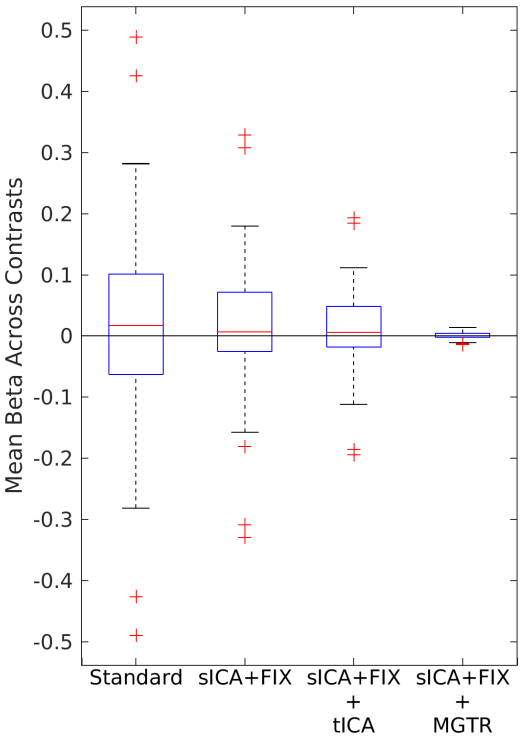


Figure 5 left hand panel shows the percent change in cluster mass for the tfMRI contrasts as a box and whisker plot for sICA+FIX vs Standard, MGTR vs Standard, sICA+FIX + tICA vs sICA+FIX, and sICA+FIX + MGTR vs sICA+FIX. The red line is the median, the edges of the box are at 25th and 75th percentiles, the whiskers are at the data point closest to +/- 2.7 standard deviations, and the outliers are the data points beyond these thresholds. The data are from the 28-subject registration optimization dataset, and the gambling vs reward contrast and its inverse are excluded because they do not appear to contain neural signal (Glasser et al., 2016a). 62/84 (74%) contrasts improve in statistical sensitivity after sICA+FIX vs Standard Processing, 42/84 (50%) improve after MGTR vs standard processing, 54/84 (64%) improve after sICA+FIX + tICA vs sICA+FIX alone, and 39/84 (46%) after sICA+FIX + MGTR vs sICA+FIX alone. The inset shows a scatter plot of the percent improvement from sICA+FIX over standard processing vs the improvement of sICA+FIX + tICA over sICA+FIX processing. The right hand panel shows the spatial means across the contrast beta maps for standard processing, sICA+FIX, sICA+FIX + tICA, and sICA+FIX + MGTR. Unlike standard processing, sICA+FIX, and sICA+FIX + tICA, nearly all of the sICA+FIX + MGTR contrast beta map means are very close to zero because MGTR results in mean centering of each timepoint's spatial map.

Effects of Temporal ICA Cleanup on Task fMRI GLM Contrast Maps

Unlike with MGTR, removal of the noise temporal ICA components does not decrease statistical sensitivity across a majority of the contrasts, rather it increases sensitivity slightly for most contrasts (54/84, Figure 5). Indeed, 59/84 contrasts are higher for sICA+FIX + tICA than after sICA+FIX + MGTR, indicating that MGTR is removing global neural signal in addition to removing global noise. Some contrasts, such as the MOTOR CUE contrast, that appeared to have a substantial global positive bias before tICA cleanup have a reduced bias after tICA cleanup, which decreases their cluster mass (Supplementary Figure 21). As a result, the overall cluster mass is reduced 1.5% from sICA+FIX for sICA+FIX + tICA, giving an overall 8.6% increase vs standard grayordinate-wise analysis (down from 10.3% for sICA+FIX alone). Notably, sICA+FIX + tICA has a cluster mass that is 9.4% higher than sICA+FIX + MGTR, however. That said, cluster mass increases markedly for parcellated analysis after tICA, likely because the additional variance being removed is a substantially larger portion of the remaining nuisance variance in a parcellated analysis than in a grayordinate-wise analysis (see Supplementary Figure 18).

Thus, temporal ICA cleanup generally produces small statistical sensitivity gains in task fMRI analysis by removing both uncorrelated nuisance variance and confound variance that is randomly correlated across subjects, leading to reduced cross-subject variability in activation level after tICA cleanup. Temporal ICA cleanup also removes biases that are consistent across subjects, such as those that could arise from stimulus correlated respiration (e.g., inspiring just before one is cued to move a body part). The effects of tICA cleanup are very different from the effects of MGTR, where in addition to removing global noise, neural signal is also removed, leading to statistical sensitivity decreases (which are larger in magnitude than the increases from tICA) and biasing the means of the contrast beta maps towards zero (Figure 5 and Supplementary Figure 18). Thus, temporal ICA cleanup produces modest benefits for most task fMRI data, particularly in parcellated data and when stimulus correlated global physiological noise is present. Such situations are analogous to what occurs in correlational resting state analyses to be discussed in the following section, where global physiological noise will always maximally impact

univariate analyses because there is no task design matrix with which the fMRI timeseries can be compared. As a result, any additive global or semi-global noise will positively bias correlations and these biases may differ across subjects or groups.

Exploration of Temporal ICA Components in Resting State fMRI Data

We found 84 reproducible (as assessed by ICASSO) temporal ICA components in the resting state fMRI data (See Supplementary Resting state Components (RC)). Three components were globally positive across the cerebral cortex and two of those three were globally positive throughout the grey matter (the other was largely negative in the cerebellum; Supplementary Figures RC3, 6, and 8). Together the global components accounted for 8.5% of the tICA explained variance. Like the single global component in the task fMRI data, CSF was negative (see above discussion on anti-correlations of CSF in global or semi-global components), and white matter was near zero. Interestingly, the two globally positive components (RC6 and 8) had some left/right asymmetries and were associated with the phase encoding direction of the data (Supplementary Figure 22). As with the task fMRI data, these global components were particularly correlated with RVT (Panel A, Figure 6) and were higher in amplitude during DVARS dips (Panel B/C, Figure 6).

Together with these three global noise components, we identified a total of 32 components with properties suggestive/indicative of noise, which together with the global components accounted for 28.4% of the tICA explained variance (notably this is in strong contrast to sICA+FIX for the rfMRI data where 93% of structured variance is noise). These include components with substantial white matter extensions (Supplemental Figures RC19, 25, 28, 44, 47, 50, and 70), inclusion of veins (Supplemental Figures RC19, and 28), similarity to movement regressor beta maps (Supplemental Figure RC50), high values around brain margins (Supplemental Figure RC57), similarity to known receive coil instabilities (Supplemental Figure RC47, 72, and 73) or image reconstruction artifacts (Supplemental Figures RC46, 53, and 64), amplitudes higher during DVARS dips (Figure 6, Supplementary Figures RC19, 28, 34, 38, 47, 50, 57, 70, and 73), and/or prominence in only one subject or run (Panel D, Figure 6, Supplemental Figures RC47, 51-53, 58-64, 66-68, 71-73, 78, and 80). Many of the noise components were similar to those found in the task fMRI data (Supplementary Figures TC/RC 1/3+6+8, 8/19, 8/28, 42/44, 43/46, 50/47, 51/50, 44/57, and 64/70) showing that much of the structured noise is reproducible across resting state and task. As with the task fMRI data, sICA+FIX's classification accuracy (99%) is very good, but not perfect and some noise does slip through (additionally, as before, some weakly structured noise may exist in the unstructured noise subspace in the sICA that is only detectable at the group level). In fact, some of the noise found in an initial temporal ICA analysis was traced back to misclassification of a few sICA+FIX components in a few subjects and was fixed using manual reclassification of these components prior to the HCP 1200 subject release.

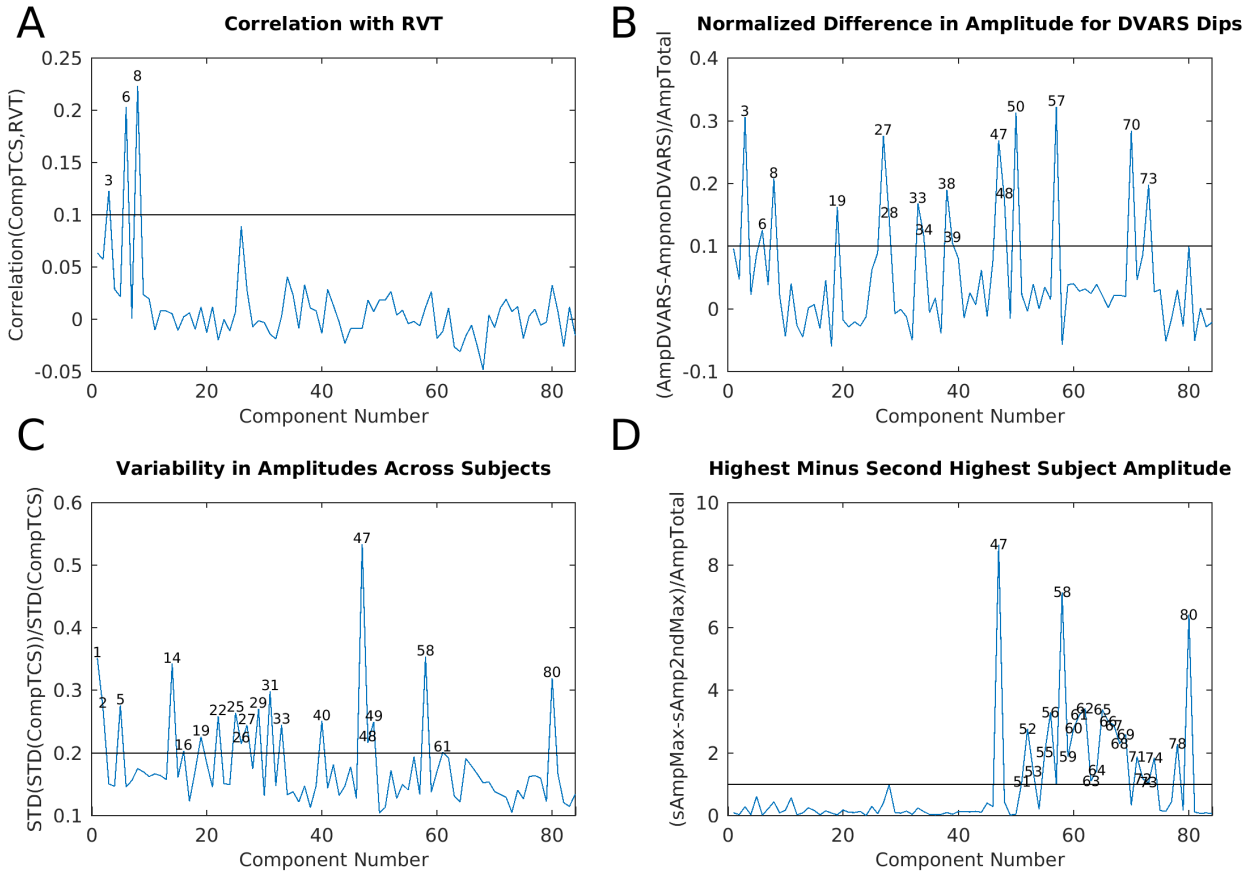


Figure 6 shows four plots that were helpful in classifying the resting state fMRI tICA components into signal and noise. Panel A shows the correlation between the component timeseries and RVT for the 157 runs with the top 10% mean correlation between RVT and their timeseries as in Figure 2. When the 3 global components timeseries (RC3, 6, and 8) are added together, the correlation with RVT is 0.28, very similar to the correlation of the single task global component (0.30) with RVT. The line is at r=0.1. Panel B is the difference in component amplitude (standard deviation of the component timeseries) between frames with DVARS dips and those without DVARS dips normalized by the standard deviation of each component across all frames as in Figure 2. The line is at 0.1. Panel C shows the variability of component amplitudes across subjects normalized by the standard deviation of each component as in Figure 2. The line is at 0.2. Panel D shows the difference between the maximum subject's component amplitude and the next highest subject's component amplitude normalized by the standard deviation of each component as in Figure 2. The line is at 1.0.

The other 52 temporal ICA components were consistent with neural signal (Supplementary Figures RC1, 2, 4, 5, 7, 9-18, 20-24, 26, 27, 29-33, 35-37, 39-43, 45, 48, 49, 54-56, 65, 69, 74-77, 79, and 81-84), and together accounted for 71.6% of the tICA explained variance. 20 of these components had visually corresponding patterns to those in the task fMRI analysis (Supplementary Figures TC/RC 14/2, 23/5, -9/9, 5/10, 4/12, 10/13, 26/15, 17/16, 22+36/20, 16/27, 6/33, 13/37, 29/39, 31/40, 34/41, 19/48, 65/76, 70/81, -67/82, 69/84, negation meaning that the sign of the component was reversed, however the signs of ICA components are arbitrary). Particularly notable correspondences were the five somatotopically organized sensori-motor networks, which retained the property of being

the only signal components associated with DVARS dips (though the left hand component was not as high as the others, Supplementary Figures RC27, 33, 39, 40, 48). The pan-sensori-motor component that was seen in the task data was also found in resting state (Supplementary Figures TC23/RC5); however, it was much stronger, explaining 2.87% of the variance vs 1.57% in task data.

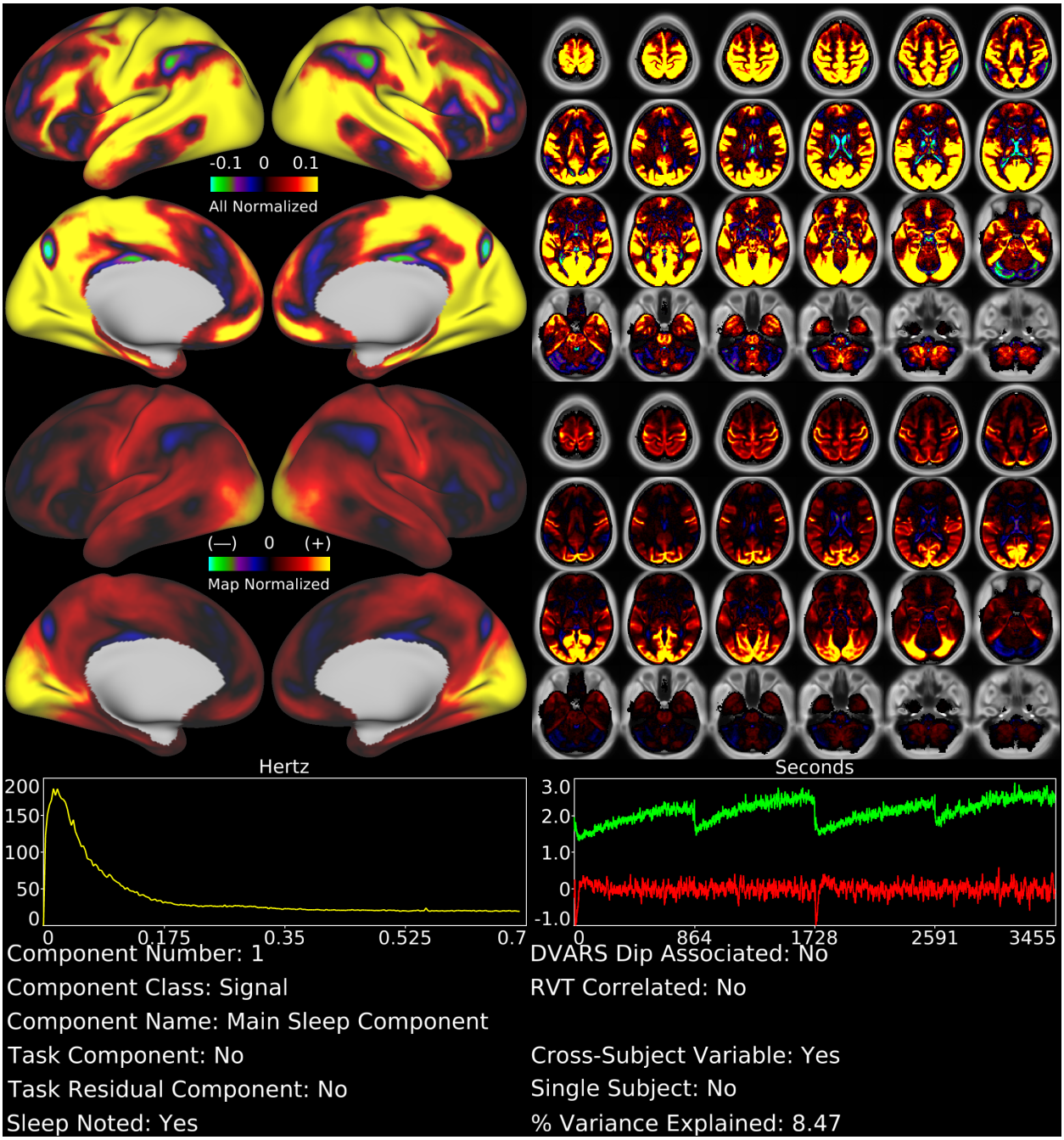


Figure 7 shows the first resting state fMRI tICA component in a standardized view (see Supplementary Figures RC 1-84). The top row of data has the color scale normalized and held constant across all 84 components (in percent BOLD), whereas the second row has the color scale set independently for each component. The first view allows components to be compared with each other on the same scale, whereas the second view highlights spatial variation in the map of

each specific component. The left chart in the third row indicates the power spectra of the component averaged across subjects and the right chart indicates the average timeseries (red) and average absolute value of the timeseries (green). The increasing trend of the absolute average timeseries across the run (which is like an average instantaneous component amplitude) is likely indicative of subjects becoming progressively drowsy or falling asleep. The timeseries are concatenated in the order that the runs were typically acquired (REST1_RL, REST1_LR, REST2_LR, REST2_RL) to enable visualization of trends across the two sessions. Additional data on the components is provided in the bottom row. Percent variance explained is the percent of the total variance explained by the tICA model.

Visual components with task/rest homologues included the four visuotopic components with a polar angle relationship (Supplementary Figures TC/RC 65/76, 70/81, 67/82, and 69/84) and two with an eccentricity relationship (Supplementary Figures TC/RC 17/16 and 34/41). Five additional visuotopically organized components were also found (Supplementary Figures RC29, 42, 49, 77, and 83). The component associated with eyes open vs closed in the task fMRI data (1.95% of explained variance, Supplementary Figure TC14; (Laumann et al., 2015), was the second strongest component in the resting state fMRI data (3.64% of explained variance, Supplementary Figure RC2), and was variable across subjects, suggesting that many subjects actually closed their eyes for part of the resting state scans despite instructions to keep them open and stare at the cross-hairs.

Aside from the somatotopically organized sensori-motor components that were modulated by task fMRI and the primary default mode component that is slightly modulated by the MOTOR and EMOTION tasks, no other resting state fMRI components closely matched task components that were modulated by task, though there were some similarities (e.g., the fronto-parietal network Supplementary Figure TC/RC 4/12, a subsidiary default mode network TC/RC -9/9 (inverted), and the memory retrieval (POS2/RSC/IPS) network TC/RC 13/37). A number of the resting state signal components have correspondence with task fMRI components that are not modulated by task, including the default mode network (Supplementary Figures TC/RC 26/15), the cingulo-opercular network (Supplementary Figures TC/RC 10/13), and the language network (Supplementary Figures TC/RC 22+36/20). Additional canonical resting state networks found only at rest included further subdivisions of the default mode network (Supplementary Figures RC7, 11, 43), the fronto-parietal control network (Supplementary Figures RC17, 21), a component with the default mode network and fronto-parietal control network anti-correlated (Supplementary Figure RC23), the dorsal attention network (Supplementary Figure RC18), the parieto-occipital network (Laumann et al., 2015), Supplementary Figure RC32), the auditory network (Supplementary Figure RC36), and the language network (Supplementary Figure RC45). Other non-canonical signal components included (Supplementary Figures RC4 24, 26, 30, 31, 35, 54, 74, 75, and 79).

In addition to the signal components described above, an additional set of signal components was generally only present at rest. The strongest component of the resting state data (Figure 7, Supplementary Figure TC1) accounted for 8.5% of the tICA-explained variance and has visual, sensori-motor, and auditory regions with positive correlation and multi-modal cognitive regions generally with negative or near zero correlation. This component bears a strong resemblance to the not purely global portion of the overall pattern seen in the mean grey signal of sICA+FIX cleaned resting state fMRI data (See

Figure S6 of Glasser et al 2016 (Glasser et al., 2016b). It is also notably variable across subjects (Figure 6). Being the strongest component and only present at rest, this component presumably is linked to a behavior that is present in the resting state but not the task state. Drowsiness or sleep is a good candidate, as this component shows similarities to prior studies of the pattern of correlation during sleep vs wake (Tagliazucchi and Laufs, 2014). Also, increased amplitude of this semi-global neural component would contribute to the higher amplitude of global signal that is seen during sleep or decreased arousal that has been previously described (Laumann et al., 2017; Liu et al., 2017; Wong et al., 2016; Wong et al., 2013; Yeo et al., 2015).

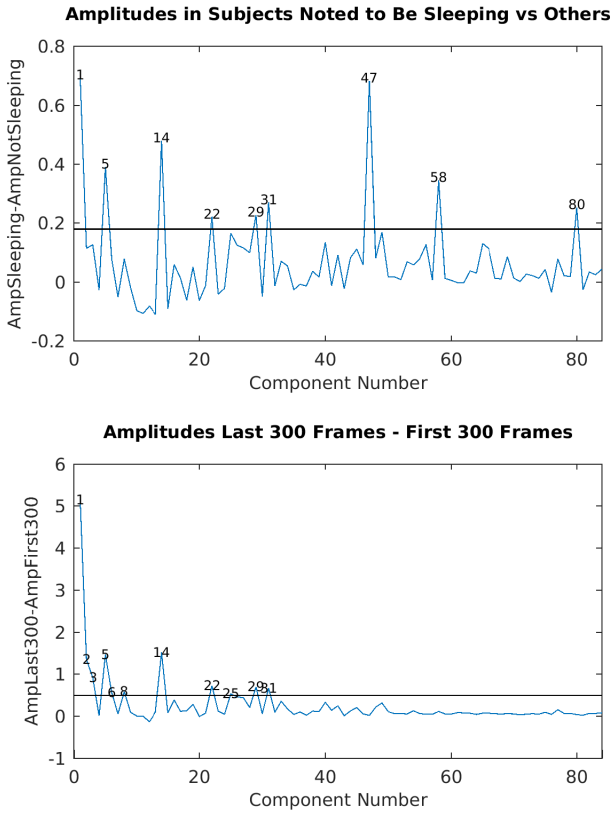


Figure 8 shows two useful sleep related measures. The top row is the difference in component amplitude in subjects that were noted to be sleeping vs those who were not. The bottom row shows the difference in component amplitude between the last 300 frames and the first 300 frames of the 1200 frame runs, as subjects will presumably be more likely to be asleep at the end of a 14.4 minute run than at the beginning. Components 47, 58, and 80 are "single subject" components that are unlikely to be specifically sleep related. The 3 global physiological noise components (RC3, 6, and 8) are all more likely to be higher at the end of a run than at the beginning, likely indicating greater respiratory variability at the end of runs when subjects are more likely to be drowsy. This is also true of noise component 25. Otherwise, these two measures agree on the components most likely to be related to sleep.

Though sleep was not explicitly monitored for in HCP subjects through eye tracking or EEG, technicians scanning the subjects were instructed to note when they noticed subjects were sleeping during resting state scans. When we compared the amplitude of the components for the 70 subjects who were noted to be sleeping vs the 379 who were not noted to be sleeping, component RC1 had the strongest association (Figure 8) and additionally had much stronger amplitude in the last 300 frames of the resting state run than the first 300 frames (Figure 8). Additionally in Figure 7, the green line showing the average absolute timeseries value is analogous to an average instantaneous amplitude that likely shows the effects of subjects on average becoming progressively drowsier or falling asleep over the course of each 14.4 minute-long resting state run. Five other components were associated both with sleep and with being stronger in the last 300 frames relative to the first 300 (Figure 8, Supplemental Figures RC5, 14, 22, 29, and 31) and had the same property of high cross-subject variability (Figure 6). Most of these components were not present in task fMRI data either, the pan-sensori-motor component being the only exception (Supplemental Figures RC5). There

were also 4 single subject components that appeared similar to RC1 or its spatial map negated (Supplemental Figures RC 55, 56, 65, and 69).

The sleep components in aggregate accounted for 19.3% of the tICA explained variance. In resting state fMRI data, after removing the global noise components, the next two biggest contributors to the global signal are the semi-global sleep component and the pan-sensori-motor component (Supplementary Figure 23). Because of these components, resting state fMRI data overall has more global signal than does task fMRI data: MGTRVar=581 (mean variance removed by MGTR after sICA+FIX + tICA cleanup in grand mean 10000 scaled data) vs MGTRVar=235 for the task data. If sleep-related components are also removed, however, this difference disappears (MGTRVar=238 vs MGTRVar=235). Indeed there is no indication from the data that the global neural BOLD signal differs in magnitude between task fMRI and awake resting state fMRI once the data have been appropriately cleaned using spatial and temporal ICA (i.e., when subjects in the resting state paradigm are not sleeping).

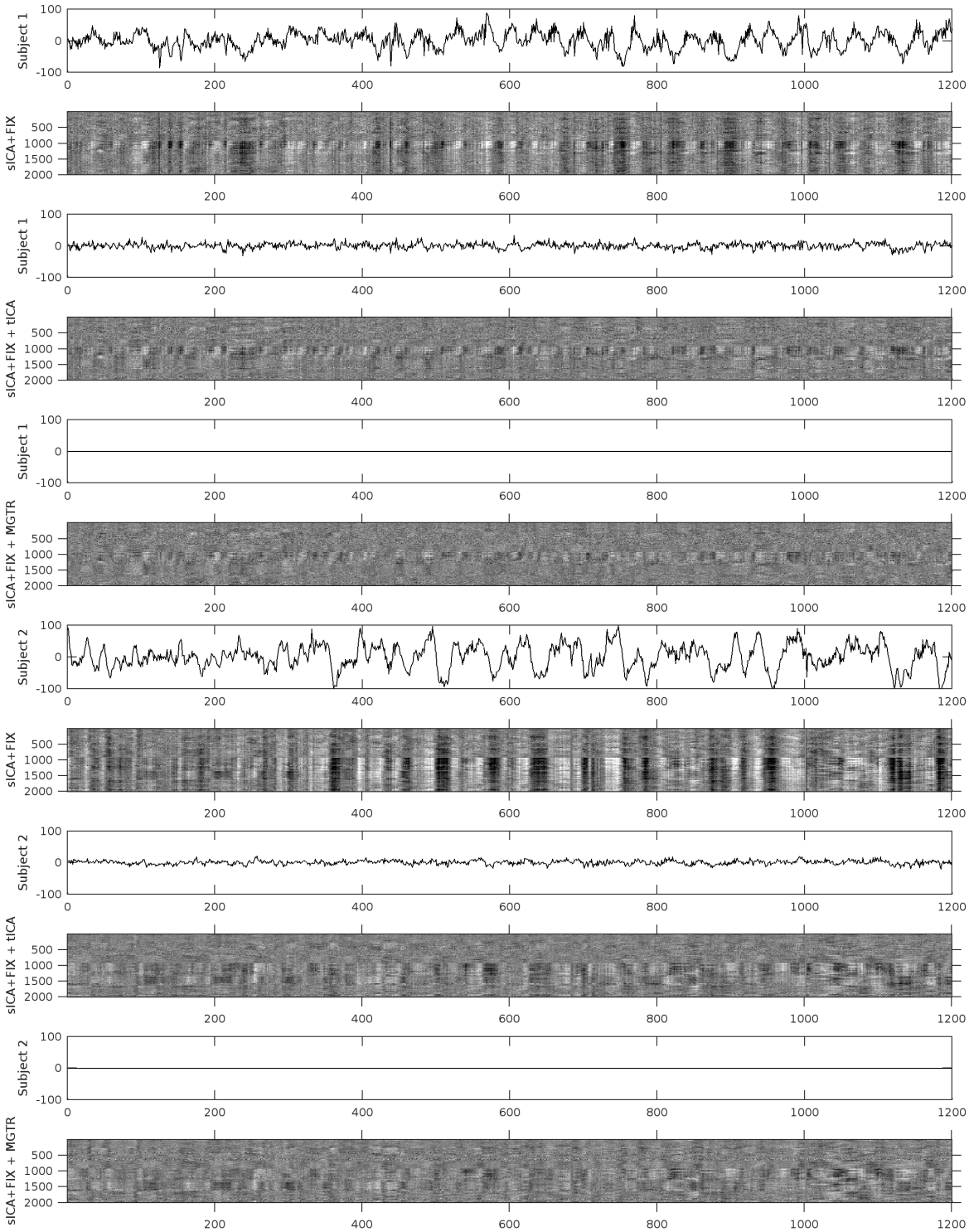
As with task fMRI, we clustered resting state fMRI components by run-wise amplitude correlations along a subjects X runs dimension (Supplementary Figure 24). Again we found that the clusters were dominated by a signal group and a noise group. Additionally, we found two groups of resting state signal components. One group tended to be correlated with RC1, and perhaps are networks that are generally more active in a drowsy or asleep state, whereas the other group was not as correlated with RC1 and may reflect networks that are more active in awake subjects. Additionally, we correlated the grayordinate spatial maps of the resting state fMRI, task fMRI, and task fMRI residual component maps (Supplementary Figure 25) and used this information when assigning component matches in the RC, TC, and TCr supplementary materials and in the text above.

Cleanup of Resting State fMRI Data Using Temporal ICA

Figure 9 shows greyplots and mean grey signal of two runs of two subjects having particularly high global noise after sICA+FIX under three cleanup conditions: sICA+FIX alone, sICA+FIX plus temporal ICA cleanup, and sICA+FIX plus MGTR. As with task fMRI, regressing out the components identified as noise removes the obvious bands in the grey plots (and substantial mean grey signal fluctuations at these times) that have previously been associated with respiratory changes (Power et al., 2017). Importantly, as with task fMRI, regressing out the noise temporal ICA components is not the same as removing the mean grey signal, as the mean grey signal after tICA cleanup retains some fluctuations. In resting state fMRI data, the mean grey timecourse variance (i.e. the amount of variance removed by regressing out the mean grey timecourse in grand mean 10,000 scaled data) before sICA+FIX cleanup is MGTRVar=9975; after sICA+FIX it is MGTRVar=1114; and after temporal ICA cleanup it is 581 (but see above point about increased sleep-related global variance). Thus, the neural global signal variance is 5.8% of the original global signal variance and is 52% of the global signal variance after sICA+FIX in resting state data. (For comparison, Liu et al found ~7% of the global signal variance remained after removing drift, motion, physiology, and WM/CSF (Liu et al., 2017). The neural global signal variance is 13.3% of the overall neural signal variance (see above for task fMRI, Supplemental Information fMRI Variance Partitioning Model), though this is increased somewhat because of sleeping subjects (7.4% after removal of sleep components). In contrast to task fMRI,

there is no task design with which to assess the extent of neural signal removed using temporal ICA cleanup or MGTR. This lack of a standard makes assessing the effect of neural signal removal more challenging. It is the primary reason that we present the resting state data after the task fMRI data, even though the majority of the recent debate regarding fMRI denoising has been centered on resting-state fMRI because of the greater impact that residual global structured noise has on resting state data and analyses.

Figure 9 shows the mean grey signals and parcellated grey plots for two resting state fMRI runs from two subjects after sICA+FIX (Rows 1-2, 7-8), sICA+FIX + tICA (Rows 3-4, 9-10), and sICA+FIX + MGTR (Rows 5-6, 11-12). The data are parcellated, then displayed according to mean parcel surface area (with 1 row for the smallest parcel and proportionally larger numbers of rows for larger parcels such that there are more than 360 rows). Additionally the data are ordered by hierarchical clustering of the group full correlation functional connectome after sICA+FIX + tICA so that parcels with more similar timeseries are closer together. tICA cleanup removes the vertical "stripes" from the grey plots after sICA+FIX and their corresponding MGT fluctuations without removing the entire MGT as occurs with MGTR. The greyplots after tICA cleanup and MGTR look very similar, although the tICA cleanup plots retain more semi-global structure, and there is subtle evidence of bias in the MGTR results similar to that found in the task fMRI data in Figure 9.


Subject 1
sICA+FIX
Subject 1
sICA+FIX + tICA
Subject 1
sICA+FIX + MGTR
Subject 2
sICA+FIX
Subject 2
sICA+FIX + tICA
Subject 2
sICA+FIX + MGTR
100
0
-100
500
1000
1500
2000
0
200
400
600
800
1000
1200

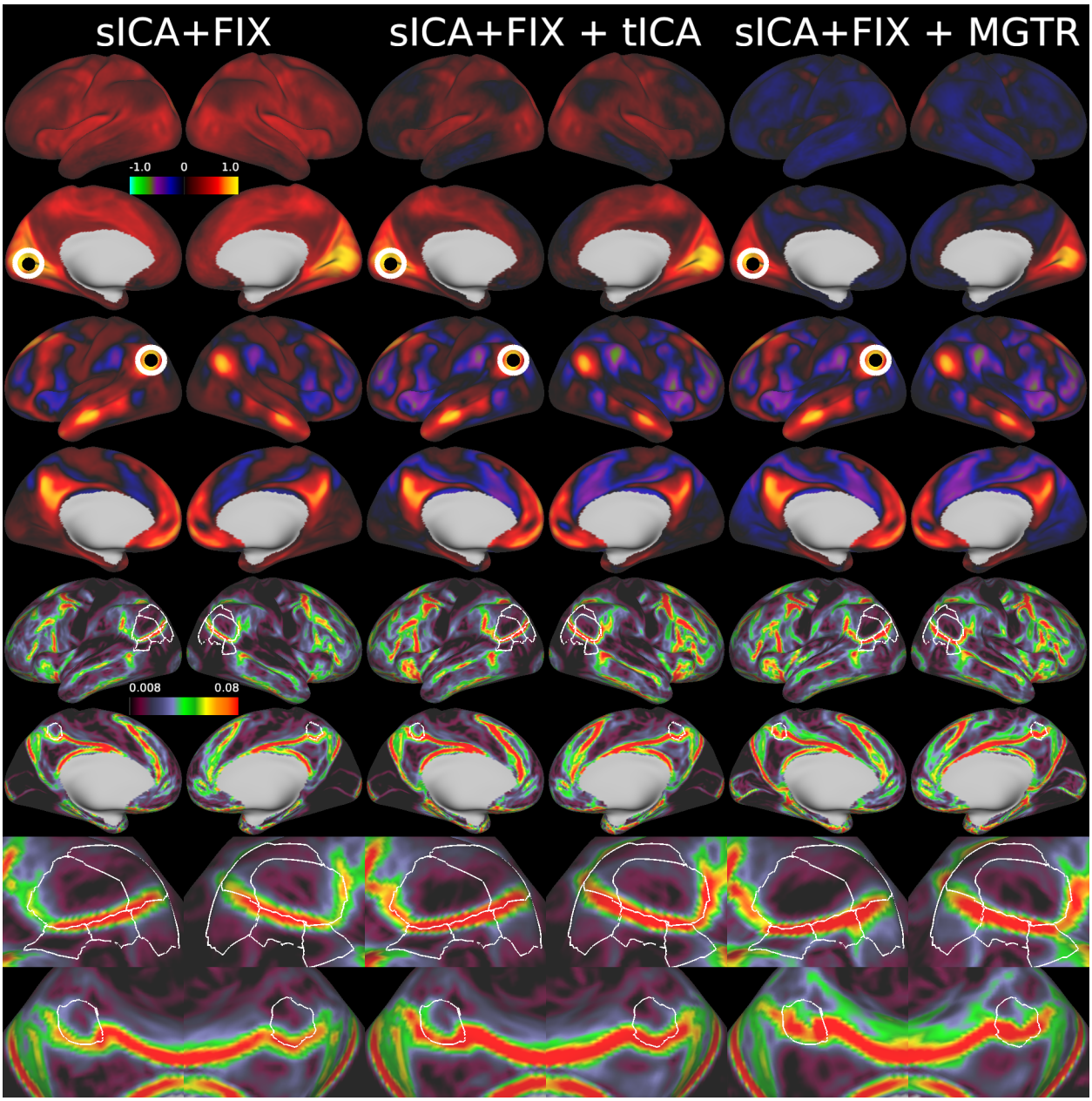


Figure 10 shows the results of group grayordinate-wise functional connectivity and functional connectivity gradient maps after sICA+FIX, sICA+FIX + tICA, and sICA+FIX + MGTR. The top row shows a seed in left hemisphere area V1 (circled) that has generally positive correlation with other grayordinates after sICA+FIX, positive to zero correlation after sICA+FIX + tICA, and widespread induced negative correlation after sICA+FIX + MGTR. The second row shows a seed in left hemisphere Area PGi (circled), which shows much more similar maps in an already anti-correlated network that is not as affected by MGTR. The third row shows the mean gradients after each kind of processing, with the gradients after sICA+FIX and sICA+FIX + tICA matching well (with an increase in gradient strength after tICA cleanup), whereas the gradients after MGTR shift in several regions highlighted by the outlined cortical areas (in white). The fourth row zooms in on several regions showing gradient shift after MGTR.

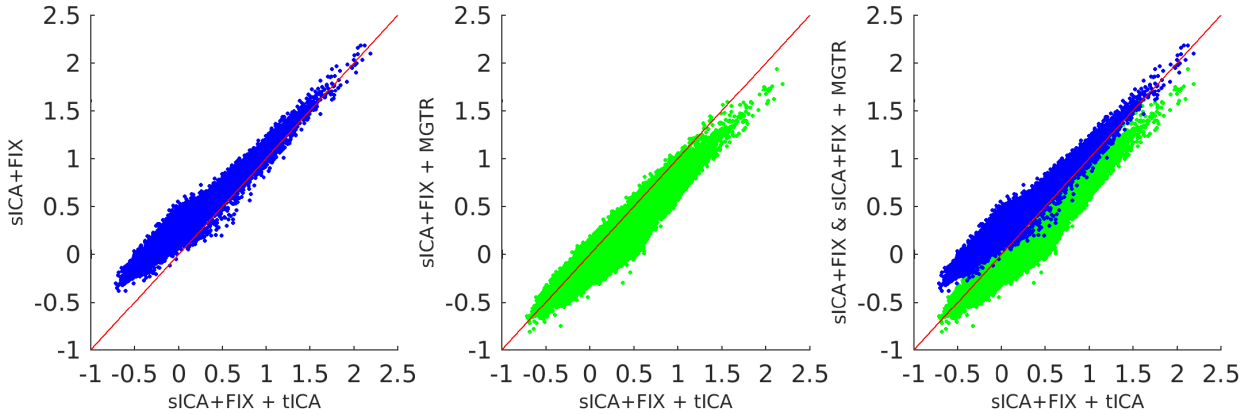
Figure 10 compares the three types of cleanup for resting state fMRI data. The top row shows group dense functional connectivity maps for a seed in area V1 after sICA+FIX (left column), after sICA+FIX + tICA cleanup (middle column), and after sICA+FIX + MGTR (right column). For this and other visual cortex seeds, the three methods differ substantially, with temporal ICA appearing to remove a global positive bias in functional connectivity and MGTR appearing to induce a negative bias in functional connectivity (anti-correlations that were not present in either the original or the tICA cleaned data). This mathematical effect of MGTR has long been known (Murphy et al., 2009; Saad et al., 2012) and is one of the arguments against MGTR, as it is neurobiologically implausible for the "true" functional connectome to have zero mean correlation. On the other hand, the large global positive bias in the correlations without temporal ICA cleanup shows why others have argued for the necessity of using something akin to MGTR (Fox et al., 2009; Hayasaka, 2013; Power et al., 2014). The second row shows dense functional connectivity maps for a seed in a cognitive/task-negative network with strong anti-correlation with other regions, including the 'task-positive' network (Figure 10). Here, the three methods yield more similar maps, with temporal ICA cleanup and MGTR being particularly similar. Thus, not all networks are equivalently biased by MGTR (though for this seed a global bias still appears to be removed by temporal ICA relative to no additional cleanup beyond sICA+FIX).

The bottom two rows of Figure 10 shows mean functional connectivity gradients computed for each method (see Methods), plus areal boundaries for six selected cortical areas (white contours, 5 lateral, 1 medial) in regions where methodological differences in gradient ridge location are particularly prominent. Both temporal ICA cleanup and MGTR increase the gradient magnitudes due to the removal of global noise. Unfortunately, MGTR also introduces a substantial shift in functional connectivity gradients relative to both no global noise cleanup and the temporal ICA-based cleanup solution. (This is one reason why MGTR was not used for the HCP's multi-modal cortical parcellation (Glasser et al., 2016a), as a global positive bias will not affect areal boundary locations, only gradient strength, whereas a network-specific negative bias will lead to biased areal boundaries). Indeed this altering of the gradients is apparent quantitatively as well, as the spatial correlation between the sICA+FIX and sICA+FIX + tICA gradients is r=0.87, whereas the correlation between sICA+FIX and sICA+FIX + MGTR is r=0.69, and between sICA+FIX + tICA and sICA+FIX + MGTR it is r=0.77. Such shifting and altering of gradients is not consistent with pure global artifact removal by MGTR (which would not affect some networks more than others) but is consistent with removal of network-specific neural signal as demonstrated above for the task fMRI data. Temporal ICA cleanup, on the other hand, removes the same global noise as MGTR without introducing a gradient shift because it does not remove semi-global neural signal.

Indeed, when one parcellates the data and examines the functional connectivity between all pairs of cortical areas, it is apparent that the temporal ICA solution lies in between no global noise cleanup and MGTR (Figures 11 and 12). On the one hand, if no additional cleanup is done after sICA+FIX, there is a positive bias in the data (Figure 11 left hand panel and Figure 12 Panel 1), and this bias may be correlated with other subject traits like motion, arousal, or body mass index (Siegel et al., 2017). On the other hand, doing MGTR removes both the global physiological noise and some of the neural signal in a network-specific way (Figure 12 compare Panel 2 vs 3 and 4 vs 5) as has been argued previously (Glasser et al., 2016b; Saad et al., 2012). In particular, neural signal is mainly

removed from both non-cognitive and task-positive regions, whereas there is less impact on other cognitive, task-negative regions (Figure 12 Panel 6). We believe that the temporal ICA cleanup approach offers a principled "best of both worlds" solution to the problem of separating global structured noise from global neural signal (see Discussion). The effect of tICA cleanup (Figure 12 Panel 4) is more spatially homogeneous than the effect of MGTR (Panel 5), and being spatially homogeneous is arguably the desired behavior when cleaning up the global artifacts that have prompted the use of MGTR. Partial correlation is an alternative approach for establishing functional connectivity relationships between cortical areas that does not suffer from the positive bias that exists with full correlation after applying only sICA+FIX. We have previously recommended this approach (including in cases where the tICA cleanup approach has not been applied, (Glasser et al., 2016b), in part because partial correlation estimation is much less sensitive to global confounds than is full correlation. Indeed, consistent with this, the benefits from tICA cleanup with partial correlation are much more modest than for full correlation (Figure 12 Panels 7-9).

Figure 11 shows the entire full correlation parcellated connectome computed by parcellating the MIGP PCA series that was the input to the analyses in Figure 10 to show an overall summary of the trends in this data. It has sICA+FIX + tICA (x-axis) plotted vs sICA+FIX (y-axis; blue) showing a positive bias in sICA+FIX on the left, and sICA+FIX + tICA plotted vs sICA+FIX + MGTR (green) showing a negative bias in sICA+FIX + MGTR in the middle. Both are plotted on the right, showing how the sICA+FIX + tICA solution falls between the two extremes of no MGTR and MGTR after sICA+FIX.



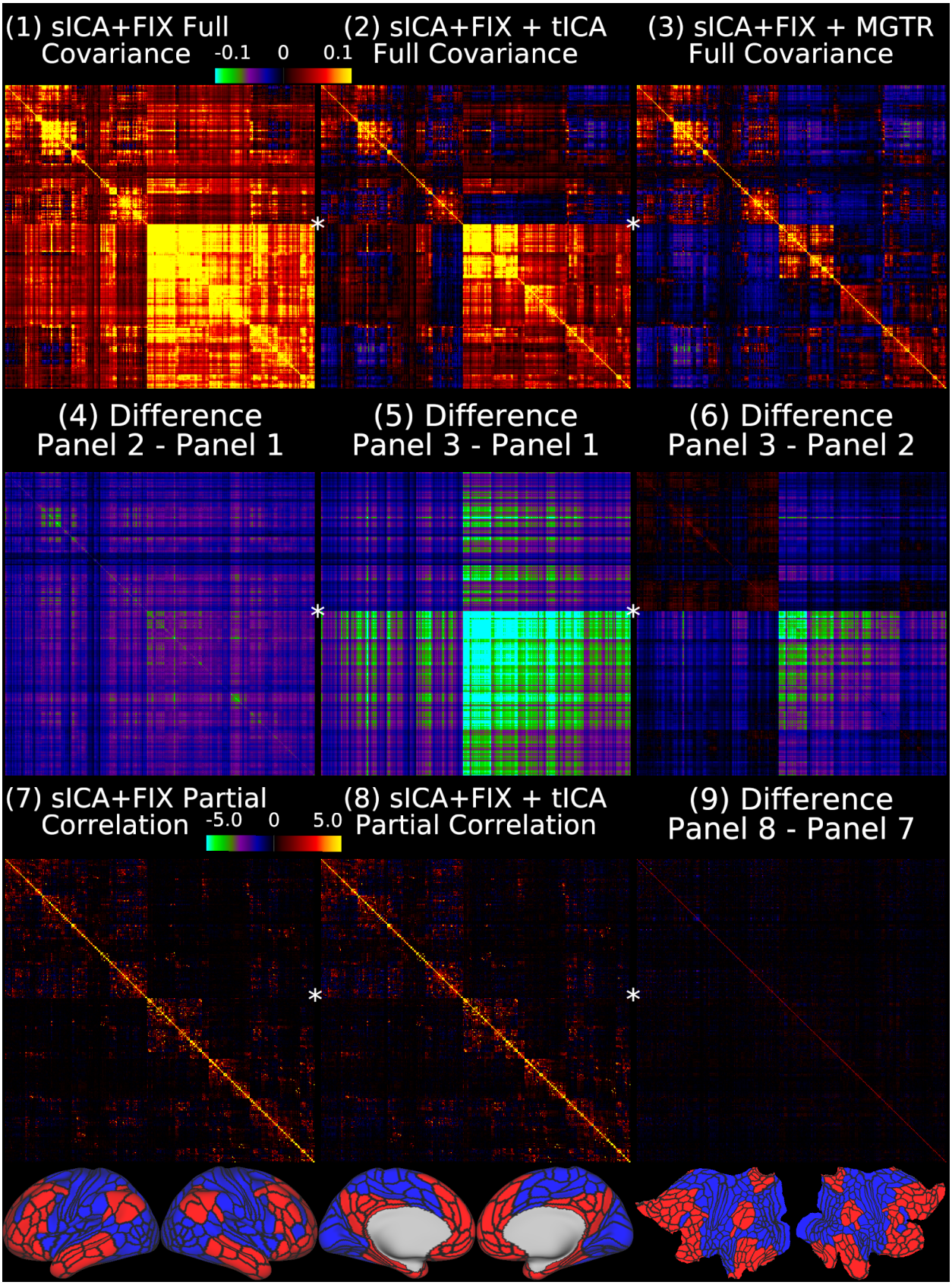

(1) sICA+FIX Full Covariance
-0.1
0
0.1
(2) sICA+FIX + tICA Full Covariance
(3) sICA+FIX + MGTR Full Covariance
(4) Difference Panel 2 - Panel 1
(5) Difference Panel 3 - Panel 1
(6) Difference Panel 3 - Panel 2
(7) sICA+FIX Partial Correlation
-5.0
0
5.0
(8) sICA+FIX + tICA Partial Correlation
(9) Difference Panel 8 - Panel 7

Figure 12 shows the group average full covariance matrices after sICA+FIX, sICA+FIX + tICA, and sICA+FIX + MGTR in Panels 1-3. We use covariance here because, like variances, covariances are additive and represent the absolute amount of variance shared by any two pairs of ROIs (scaled from -0.1 to 0.1 in percent BOLD for Panels 1-6). As in other figures, a global positive bias is removed by tICA cleanup (Panel 4 shows the difference between Panel 2 and Panel 1), but MGTR also removes additional signal in the bottom right quadrant of the matrix relative to the upper left quadrant with the off diagonal quadrants in between (Panel 5 shows the difference between Panel 3 and Panel 1). Indeed, the difference between the tICA cleanup and MGTR (Panel 6 shows the difference between Panel 3 and Panel 2) is highly network specific, including small increases in cognitive/task-negative regions (bottom row, red) and large decreases in non-cognitive regions (bottom row, blue), with connections between these two groups of regions showing smaller decreases. Panels 7 and 8 show that mean across subjects partial correlation regularized with ridge regression (rho=0.23, which was optimal in matching the individual matrices to the group matrix, scaled Z=+/-5)) is much less affected by tICA cleanup, as it already controls for global artifacts (Glasser et al., 2016b). Thus Panel 9 (difference between Panel 8 and Panel 7) does not reveal substantial differences. The 360 cortical areas are ordered according to the same hierarchical clustering as the grey plots, and the first split, into cognitive/task negative (red) and non-cognitive/task positive (blue) regions, is shown in the bottom row and noted by a star on the netmats, with red areas above the star and blue ones below it. Note that it would be inappropriate to use partial correlation after MGTR, as any dataset that has zero global signal is rank deficient, because each parcel's timeseries equals the negated sum of all other parcels' timeseries.

Discussion

Here we have shown that it is possible to separate semi-global and global noise from global neural signal using temporal ICA and then to selectively remove the noise. Based on its spatial pattern (specific to grey matter but global) and temporal characteristics (BOLD power spectrum, correlated with RVT), much of this global noise is likely due to global fluctuations in blood flow from changes in respiratory rate and depth, leading to changes in pCO_2, or from other physiological sources like heart rate variations. This global noise is left behind by spatial-ICA-based cleanup approaches like sICA+FIX. Selectively removing this noise without removing the global neural signal addresses a major issue in functional MRI denoising by allowing for a best of both worlds solution between not performing global signal regression (and thus tolerating a positive bias from global noise in univariate connectivity measures that may differ across subjects and groups), or performing global signal regression (and thus inducing network-specific negative biases in connectivity measures). Therefore, we enable the generation of univariate connectivity measures that are largely free of global physiological noise and are not biased by global signal regression. That said, multivariate approaches such as partial correlation, which we have previously recommended as a way around the global noise dilemma (Glasser et al., 2016b), is substantially less affected by global artifacts (amongst other advantages).

We first considered task fMRI data, where we have a prediction of what a portion of the subjects' BOLD signal should look like during the task. In the task fMRI data, we identified a single global component that is strongly correlated with the respiratory physiology parameter RVT and represents a major source of global noise (e.g., correlation of r=0.68 with the mean grey timecourse after sICA+FIX). We then selectively removed this component, together with other noise components, and compared the results of task fMRI

statistics before and after this processing and to those after global signal regression. Because it had not previously been done, we also compared the effects of sICA+FIX cleanup vs no cleanup on task fMRI data. We found that sICA+FIX and temporal ICA cleanup generally increased statistical sensitivity and reduced biases in the task activation maps, whereas global signal regression generally decreased statistical sensitivity and biased the GLM contrast maps towards a mean of zero. Importantly, our tICA cleanup approach also removed the "stripes" associated with global artifacts in greyplots (Figures 4 and 9), something that physiological regressors have only incompletely achieved to date and that has otherwise required global signal regression to achieve (Power et al., 2017). We then applied the same temporal ICA approach to resting state fMRI data, finding 3 global components, the sum of whose timecourses correlated well with the mean grey signal after sICA+FIX (r=0.71), and demonstrated the removal of a global positive bias in connectivity without inducing a network-specific negative bias in connectivity. Additionally, we showed that whereas functional connectivity gradients shift after global signal regression (consistent with removing neural signal in a network-specific way), they do not shift with temporal ICA-based global noise cleanup, consistent with removing only a global artifact. Indeed, the effect of tICA cleanup (Figure 12 Panel 4) is arguably more like the spatially homogeneous effect that practitioners of global signal regression aspire to apply to their data, whereas the effect of global signal regression (Figure 12 Panel 5) is actually quite spatially inhomogeneous and network specific.

In addition to these successful denoising results, the temporal ICA - run on millions of task fMRI and resting state fMRI timepoints - generates some interesting neurobiological findings. We identified a number of neural signal components that are modulated by tasks in task fMRI data, many of which are very similar to the task fMRI contrast beta maps. Additionally, we identified components present in both task and resting state fMRI data that are not modulated by task. We found a set of five somatotopically organized sensori-motor networks that are modulated by the MOTOR task, but are also correlated with DVARS dips, which are more robust indicators of subject head movement in the high spatial and temporal resolution fMRI data that we used in this study than are common measures based on movement parameter traces such as FD. This finding is notable because it explains how some data cleanup benchmarks that do not take into account the expected neural changes of different subject behaviors (such as movement or sleep) may suggest an apparent benefit for approaches that remove neural signal (e.g., scrubbing or global signal regression, see Main Supplementary Information: Supplementary Discussion). There may also be correlations between noncompliant behavior in the scanner and other behavioral variables collected outside the scanner. We also identify a set of components present only in the resting state that are associated with subjects noted to be sleeping and, in particular, a semi-global neural component that may underlie the higher global signal in sleep or drowsy states (Laumann et al., 2017; Liu et al., 2017; Tagliazucchi and Laufs, 2014; Wong et al., 2016; Wong et al., 2013; Yeo et al., 2015). In general, a number of components are shared between task and resting state fMRI data, but there are clear differences as well—above and beyond the task-modulated components. That said, we found no evidence that task and resting state fMRI have differential amounts of global neural signal once sleep is taken into account. Therefore, using first spatial and then temporal ICA allows for selective structured noise removal from both kinds of fMRI data. We believe that the converging evidence from task and resting state fMRI presented in this study support temporal ICA

cleanup as an attractive "best of both worlds" approach that avoids both the pitfalls of failing to clean global structured noise from the data, and of removing global neural signal from the data with overly aggressive methods.

The method described in this study is designed to work with high spatial and temporal resolution fMRI data at the group level (including a large number of subjects) after precise functional alignment across subjects. Many studies (often by necessity) acquire relatively small amounts of resting state fMRI data in small numbers of subjects at low spatial and temporal resolutions. It is already known that subject-wise methods like sICA+FIX do not perform as well on such datasets because of the small number of timepoints (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). Temporal ICA will be even more affected by these issues (Smith et al., 2012), so the methods presented in this study are intended for use in HCP-Style neuroimaging acquisitions (i.e., better than 2.6mm isotropic spatial resolution, better than 1s temporal resolution, and large total numbers of timepoints (#timepoints per subject * #subjects), e.g., something greater than the 20,000 used in Smith et al 2012 (Smith et al., 2012). It is currently unknown how many would be needed, however. Thus, these methods are particularly applicable to the existing HCP data, the ongoing HCP Lifespan (Development and Aging) projects, Connectomes Related to Human Disease ("Disease Connectome") data, HCP-Style studies of individual investigators and the UK Biobank (Miller et al., 2016). In general, it will be particularly important in the future to assess tICA denoising performance as a function of the number of subjects included in a study and how many timepoints there are for each subject. A corollary approach that also warrants careful examination is whether it will be profitable to perform tICA on task and resting-state data combined, thereby potentially increasing the number of data points available for analysis in many projects. Additionally, it may be possible to use the temporal ICA decomposition from a larger group of subjects to correct other individual subjects that were not a part of the original group; however, this problem is left for future work.

The approach presented here has several limitations. First, it currently relies on manual classification of temporal ICA components at the group level before cleanup can proceed. However, this is likely not a long-term impediment, given the success of machine learning classifiers at classifying spatial ICA components. Not all investigators may agree with the specific component classifications used here, and further understanding of temporal ICA may lead to reclassifications in the future. Indeed, the global signal/global noise problem is now one of component classification, rather than to do or not do global signal regression. The study is not easily generalizable outside of HCP-Style datasets because of the data quality and quantity requirements of temporal ICA, which are even more demanding than those for spatial ICA (which itself works less well on more traditional data (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). That said, we have shown that the method works in two large datasets, a task fMRI dataset and a resting state fMRI dataset from the HCP.

Finally, we reiterate that our approach is intended only to clean data of physiological, movement, and MR physics related artifacts, not to alter the BOLD neural signal present in subjects' scans. Subjects who do not follow instructions and fail to do the assigned task, move, or fall asleep will have differing neural BOLD activity and connectivity patterns than those who are compliant. As a result we do not recommend simply regressing out the somatosensory-motor networks or the sleep components, as these

represent real neural signals. That said, investigators might rightly wish to exclude portions of scans or subjects who engage in off-task behaviors. The occurrence of off-task behaviors is generally more prevalent in resting state fMRI data than task fMRI data, as we have shown here. It may also be more prevalent in some groups of subjects than others, which is an important confound for resting state fMRI studies, particularly those comparing two or more groups of subjects. Other paradigms, such as naturalistic movies may do a better job of maintaining subject "neural compliance" with the study design. Nevertheless, it will still be important to selectively remove spatially specific and global structured noise from such fMRI data with spatial and temporal ICA.

Acknowledgements

Supported in part by the Human Connectome Project, WU-Minn-Ox Consortium (1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; the McDonnell Center for Systems Neuroscience at Washington University; and NIH F30 MH097312 (M.F.G.), RO1 MH-60974 (D.C.V.E.). Funding to SS, JB, SH gratefully acknowledged via Wellcome Trust strategic award 098369/Z/12/Z. The authors would also like to thank Andreas Bartsch for helpful references related to negative CSF correlations in global and semi-global components.

References

Aguirre, G.K., E. Zarahn, and M. D'Esposito. 1997. Empirical analyses of BOLD fMRI statistics. II. Spatially smoothed data collected under null-hypothesis and experimental conditions. *NeuroImage*. 5:199-212.

Aguirre, G.K., E. Zarahn, and M. D'Esposito. 1998. The inferential impact of global signal covariates in functional neuroimaging analyses. *NeuroImage*. 8:302-306.

Anderson, J.S., T.J. Druzgal, M. Lopez-Larson, E.K. Jeong, K. Desai, and D. Yurgelun-Todd. 2011. Network anticorrelations, global regression, and phase-shifted soft tissue correction. *Human brain mapping*. 32:919-934.

Barch, D.M., 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, and D.C. Van Essen. 2013. Function in the human connectome: task-fMRI and individual differences in behavior. *NeuroImage*. 80:169-189.

Beckmann, C.F., M. DeLuca, J.T. Devlin, and S.M. Smith. 2005. Investigations into resting-state connectivity using independent component analysis. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 360:1001-1013.

Beckmann, C.F., and S.M. Smith. 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE transactions on medical imaging*. 23:137-152.

Behzadi, Y., K. Restom, J. Liau, and T.T. Liu. 2007. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*. 37:90-101.

Birn, R.M., J.B. Diamond, M.A. Smith, and P.A. Bandettini. 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. *NeuroImage*. 31:1536-1548.

Brewer, A.A., W.A. Press, N.K. Logothetis, and B.A. Wandell. 2002. Visual areas in macaque cortex measured using functional magnetic resonance imaging. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 22:10416-10426.

Bright, M.G., M. Bianciardi, J.A. de Zwart, K. Murphy, and J.H. Duyn. 2014. Early anti-correlated BOLD signal changes of physiologic origin. *NeuroImage*. 87:287-296.

Brooks, J.C., C.F. Beckmann, K.L. Miller, R.G. Wise, C.A. Porro, I. Tracey, and M. Jenkinson. 2008. Physiological noise modelling for spinal functional magnetic resonance imaging studies. *NeuroImage*. 39:680-692.

Burgess, G.C., S. Kandala, D. Nolan, T.O. Laumann, J.D. Power, B. Adeyemo, M.P. Harms, S.E. Petersen, and D.M. Barch. 2016. Evaluation of Denoising Strategies to Address Motion-Correlated Artifacts in Resting-State Functional Magnetic Resonance Imaging Data from the Human Connectome Project. *Brain connectivity*. 6:669-680.

Chai, X.J., A.N. Castanon, D. Ongur, and S. Whitfield-Gabrieli. 2012. Anticorrelations in resting state networks without global signal regression. *NeuroImage*. 59:1420-1428.

Chang, C., J.P. Cunningham, and G.H. Glover. 2009. Influence of heart rate on the BOLD signal: the cardiac response function. *NeuroImage*. 44:857-869.

Chang, C., and G.H. Glover. 2009. Relationship between respiration, end-tidal CO2, and BOLD signals in resting-state fMRI. *NeuroImage*. 47:1381-1393.

Chen, L., A. Beckett, A. Verma, and D.A. Feinberg. 2015. Dynamics of respiratory and cardiac CSF motion revealed with real-time simultaneous multi-slice EPI velocity phase contrast imaging. *NeuroImage*. 122:281-287.

Ciric, R., D.H. Wolf, J.D. Power, D.R. Roalf, G.L. Baum, K. Ruparel, R.T. Shinohara, M.A. Elliott, S.B. Eickhoff, C. Davatzikos, R.C. Gur, R.E. Gur, D.S. Bassett, and T.D. Satterthwaite. 2017. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *NeuroImage*. 154:174-187.

Coalson, T.S., D.C. Van Essen, and M.F. Glasser. 2017. Lost in Space: The Impact of Traditional Neuroimaging Methods on the Spatial Localization of Cortical Areas. *(in preparation)*.

Farah, M. 2014. Brain images, babies, and bathwater: Critiquing critiques of functional neuroimaging. *The Hastings Center Report*. 44.

Filippini, N., B.J. MacIntosh, M.G. Hough, G.M. Goodwin, G.B. Frisoni, S.M. Smith, P.M. Matthews, C.F. Beckmann, and C.E. Mackay. 2009. Distinct patterns of brain activity in young carriers of the APOE-epsilon4 allele. *Proceedings of the National Academy of Sciences of the United States of America*. 106:7209-7214.

Fox, M.D., D. Zhang, A.Z. Snyder, and M.E. Raichle. 2009. The global signal and observed anticorrelated resting state brain networks. *Journal of neurophysiology*. 101:3270-3283.

Friston, K.J., S. Williams, R. Howard, R.S. Frackowiak, and R. Turner. 1996. Movement-related effects in fMRI time-series. *Magnetic resonance in medicine*. 35:346-355.

Gattass, R., A.P. Sousa, and C.G. Gross. 1988. Visuotopic organization and extent of V3 and V4 of the macaque. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 8:1831-1845.

Glasser, M.F., T.S. Coalson, E.C. Robinson, C.D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C.F. Beckmann, M. Jenkinson, S.M. Smith, and D.C. Van Essen. 2016a. A multi-modal parcellation of human cerebral cortex. *Nature*. 536:171-178.

Glasser, M.F., S.M. Smith, D.S. Marcus, J.L. Andersson, E.J. Auerbach, T.E. Behrens, T.S. Coalson, M.P. Harms, M. Jenkinson, S. Moeller, E.C. Robinson, S.N. Sotiropoulos, J. Xu, E. Yacoub, K. Ugurbil, and D.C. Van Essen. 2016b. The Human Connectome Project's neuroimaging approach. *Nature neuroscience*. 19:1175-1187.

Glasser, M.F., S.N. Sotiropoulos, J.A. Wilson, T.S. Coalson, B. Fischl, J.L. Andersson, J. Xu, S. Jbabdi, M. Webster, J.R. Polimeni, D.C. Van Essen, and M. Jenkinson. 2013. The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*. 80:105-124.

Golestani, A.M., C. Chang, J.B. Kwinta, Y.B. Khatamian, and J. Jean Chen. 2015. Mapping the end-tidal CO2 response function in the resting-state BOLD fMRI signal: spatial specificity, test-retest reliability and effect of fMRI sampling rate. *NeuroImage*. 104:266-277.

Griffanti, L., G. Salimi-Khorshidi, C.F. Beckmann, E.J. Auerbach, G. Douaud, C.E. Sexton, E. Zsoldos, K.P. Ebmeier, N. Filippini, C.E. Mackay, S. Moeller, J. Xu, E. Yacoub, G. Baselli, K. Ugurbil, K.L. Miller, and S.M. Smith. 2014. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage*. 95:232-247.

Hayasaka, S. 2013. Functional connectivity networks with and without global signal correction. *Frontiers in human neuroscience*. 7:880.

Himberg, J., A. Hyvarinen, and F. Esposito. 2004. Validating the independent components of neuroimaging time series via clustering and visualization. *NeuroImage*. 22:1214-1222.

Hyvarinen, A. 1999. Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks,*. 10:626-634.

Jenkinson, M. 1999. Measuring Transformation Error by RMS Deviation. *fMRIB technical report TR99MJ1*.

Jin, T., and S.G. Kim. 2010. Change of the cerebrospinal fluid volume during brain activation investigated by T(1rho)-weighted fMRI. *NeuroImage*. 51:1378-1383.

Kundu, P., S.J. Inati, J.W. Evans, W.M. Luh, and P.A. Bandettini. 2012. Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *NeuroImage*. 60:1759-1770.

Laumann, T.O., E.M. Gordon, B. Adeyemo, A.Z. Snyder, S.J. Joo, M.Y. Chen, A.W. Gilmore, K.B. McDermott, S.M. Nelson, N.U. Dosenbach, B.L. Schlaggar, J.A. Mumford, R.A. Poldrack, and S.E. Petersen. 2015. Functional System and Areal Organization of a Highly Sampled Individual Human Brain. *Neuron*. 87:657-670.

Laumann, T.O., A.Z. Snyder, A. Mitra, E.M. Gordon, C. Gratton, B. Adeyemo, A.W. Gilmore, S.M. Nelson, J.J. Berg, D.J. Greene, J.E. McCarthy, E. Tagliazucchi, H. Laufs, B.L. Schlaggar, N.U.F. Dosenbach, and S.E. Petersen. 2017. On the Stability of BOLD fMRI Correlations. *Cereb Cortex*. 27:4719-4732.

Liu, T.T. 2016. Noise contributions to the fMRI signal: An overview. *NeuroImage*. 143:141-151.

Liu, T.T., A. Nalci, and M. Falahpour. 2017. The global signal in fMRI: Nuisance or Information? *NeuroImage*. 150:213-229.

Macey, P.M., K.E. Macey, R. Kumar, and R.M. Harper. 2004. A method for removal of global effects from fMRI time series. *NeuroImage*. 22:360-366.

Marx, M., K.B. Pauly, and C. Chang. 2013. A novel approach for global noise reduction in resting-state fMRI: APPLECOR. *NeuroImage*. 64:19-31.

Miller, K.L., F. Alfaro-Almagro, N.K. Bangerter, D.L. Thomas, E. Yacoub, J. Xu, A.J. Bartsch, S. Jbabdi, S.N. Sotiropoulos, J.L. Andersson, L. Griffanti, G. Douaud, T.W. Okell, P. Weale, I. Dragonu, S. Garratt, S. Hudson, R. Collins, M. Jenkinson, P.M. Matthews, and S.M. Smith. 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature neuroscience*. 19:1523-1536.

Murphy, K., R.M. Birn, D.A. Handwerker, T.B. Jones, and P.A. Bandettini. 2009. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *NeuroImage*. 44:893-905.

Murphy, K., and M.D. Fox. 2017. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *NeuroImage*. 154:169-173.

Muschelli, J., M.B. Nebel, B.S. Caffo, A.D. Barber, J.J. Pekar, and S.H. Mostofsky. 2014. Reduction of motion-related artifacts in resting state fMRI using aCompCor. *NeuroImage*. 96:22-35.

Power, J.D. 2017. A simple but useful way to assess fMRI scan qualities. *NeuroImage*. 154:150-158.

Power, J.D., K.A. Barnes, A.Z. Snyder, B.L. Schlaggar, and S.E. Petersen. 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*. 59:2142-2154.

Power, J.D., A. Mitra, T.O. Laumann, A.Z. Snyder, B.L. Schlaggar, and S.E. Petersen. 2014. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*. 84:320-341.

Power, J.D., M. Plitt, T.O. Laumann, and A. Martin. 2017. Sources and implications of whole-brain fMRI signals in humans. *NeuroImage*. 146:609-625.

Power, J.D., B.L. Schlaggar, and S.E. Petersen. 2015. Recent progress and outstanding issues in motion correction in resting state fMRI. *NeuroImage*. 105:536-551.

Pruim, R.H., M. Mennes, J.K. Buitelaar, and C.F. Beckmann. 2015a. Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. *NeuroImage*. 112:278-287.

Pruim, R.H., M. Mennes, D. van Rooij, A. Llera, J.K. Buitelaar, and C.F. Beckmann. 2015b. ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage*. 112:267-277.

Raj, D., A.W. Anderson, and J.C. Gore. 2001. Respiratory effects in human functional magnetic resonance imaging due to bulk susceptibility changes. *Physics in medicine and biology*. 46:3331-3340.

Robinson, E.C., S. Jbabdi, M.F. Glasser, J. Andersson, G.C. Burgess, M.P. Harms, S.M. Smith, D.C. Van Essen, and M. Jenkinson. 2014. MSM: a new flexible framework for Multimodal Surface Matching. *NeuroImage*. 100:414-426.

Saad, Z.S., S.J. Gotts, K. Murphy, G. Chen, H.J. Jo, A. Martin, and R.W. Cox. 2012. Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. *Brain connectivity*. 2:25-32.

Salimi-Khorshidi, G., G. Douaud, C.F. Beckmann, M.F. Glasser, L. Griffanti, and S.M. Smith. 2014. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage*. 90:449-468.

Satterthwaite, T.D., K. Ruparel, J. Loughead, M.A. Elliott, R.T. Gerraty, M.E. Calkins, H. Hakonarson, R.C. Gur, R.E. Gur, and D.H. Wolf. 2012. Being right is its own reward: load and performance related ventral striatum activation to correct responses during a working memory task in youth. *NeuroImage*. 61:723-729.

Sereno, M.I., A.M. Dale, J.B. Reppas, K.K. Kwong, J.W. Belliveau, T.J. Brady, B.R. Rosen, and R.B. Tootell. 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*. 268:889-893.

Shmueli, K., P. van Gelderen, J.A. de Zwart, S.G. Horovitz, M. Fukunaga, J.M. Jansma, and J.H. Duyn. 2007. Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI BOLD signal. *NeuroImage*. 38:306-320.

Siegel, J.S., A. Mitra, T.O. Laumann, B.A. Seitzman, M. Raichle, M. Corbetta, and A.Z. Snyder. 2017. Data Quality Influences Observed Links Between Functional Connectivity and Behavior. *Cereb Cortex*. 27:4492-4502.

Smith, S.M., C.F. Beckmann, J. Andersson, E.J. Auerbach, J. Bijsterbosch, G. Douaud, E. Duff, D.A. Feinberg, L. Griffanti, M.P. Harms, M. Kelly, T. Laumann, K.L. Miller, S. Moeller, S. Petersen, J. Power, G. Salimi-Khorshidi, A.Z. Snyder, A.T. Vu, M.W. Woolrich, J. Xu, E. Yacoub, K. Ugurbil, D.C. Van Essen, and M.F. Glasser. 2013a. Resting-state fMRI in the Human Connectome Project. *NeuroImage*. 80:144-168.

Smith, S.M., A. Hyvarinen, G. Varoquaux, K.L. Miller, and C.F. Beckmann. 2014. Group-PCA for very large fMRI datasets. *NeuroImage*. 101:738-749.

Smith, S.M., K.L. Miller, S. Moeller, J. Xu, E.J. Auerbach, M.W. Woolrich, C.F. Beckmann, M. Jenkinson, J. Andersson, M.F. Glasser, D.C. Van Essen, D.A. Feinberg, E.S. Yacoub, and K. Ugurbil. 2012. Temporally-independent functional modes of spontaneous brain activity. *Proceedings of the National Academy of Sciences of the United States of America*. 109:3131-3136.

Smith, S.M., D. Vidaurre, C.F. Beckmann, M.F. Glasser, M. Jenkinson, K.L. Miller, T.E. Nichols, E.C. Robinson, G. Salimi-Khorshidi, M.W. Woolrich, D.M. Barch, K. Ugurbil, and D.C. Van Essen. 2013b. Functional connectomics from resting-state fMRI. *Trends in cognitive sciences*. 17:666-682.

Smyser, C.D., T.E. Inder, J.S. Shimony, J.E. Hill, A.J. Degnan, A.Z. Snyder, and J.J. Neil. 2010. Longitudinal analysis of neural network development in preterm infants. *Cereb Cortex*. 20:2852-2862.

Tagliazucchi, E., and H. Laufs. 2014. Decoding wakefulness levels from typical fMRI resting-state data reveals reliable drifts between wakefulness and sleep. *Neuron*. 82:695-708.

Thomas, B.P., P. Liu, S. Aslan, K.S. King, M.J. van Osch, and H. Lu. 2013. Physiologic underpinnings of negative BOLD cerebrovascular reactivity in brain ventricles. *NeuroImage*. 83:505-512.

Van Dijk, K.R., M.R. Sabuncu, and R.L. Buckner. 2012. The influence of head motion on intrinsic functional connectivity MRI. *NeuroImage*. 59:431-438.

Van Essen, D.C., W.T. Newsome, and J.H.R. Maunsell. 1984. The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies and individual variability. *Vision Res.* 24:429-448.

Van Essen, D.C., J. Smith, M.F. Glasser, J. Elam, C.J. Donahue, D.L. Dierker, E.K. Reid, T. Coalson, and J. Harwell. 2017. The Brain Analysis Library of Spatial maps and Atlases (BALSA) database. *NeuroImage*. 144:270-274.

Van Essen, D.C., S.M. Smith, D.M. Barch, T.E. Behrens, E. Yacoub, and K. Ugurbil. 2013. The WU-Minn Human Connectome Project: an overview. *NeuroImage*. 80:62-79.

Wong, C.W., P.N. DeYoung, and T.T. Liu. 2016. Differences in the resting-state fMRI global signal amplitude between the eyes open and eyes closed states are related to changes in EEG vigilance. *NeuroImage*. 124:24-31.

Wong, C.W., V. Olafsson, O. Tal, and T.T. Liu. 2013. The amplitude of the resting-state fMRI global signal is related to EEG vigilance measures. *NeuroImage*. 83:983-990.

Woolrich, M.W., T.E. Behrens, C.F. Beckmann, M. Jenkinson, and S.M. Smith. 2004. Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*. 21:1732-1747.

Woolrich, M.W., B.D. Ripley, M. Brady, and S.M. Smith. 2001. Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*. 14:1370-1386.

Yan, C.G., B. Cheung, C. Kelly, S. Colcombe, R.C. Craddock, A. Di Martino, Q. Li, X.N. Zuo, F.X. Castellanos, and M.P. Milham. 2013. A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *NeuroImage*. 76:183-201.

Yang, G.J., J.D. Murray, M. Glasser, G.D. Pearlson, J.H. Krystal, C. Schleifer, G. Repovs, and A. Anticevic. 2016. Altered Global Signal Topography in Schizophrenia. *Cereb Cortex*.

Yang, G.J., J.D. Murray, G. Repovs, M.W. Cole, A. Savic, M.F. Glasser, C. Pittenger, J.H. Krystal, X.J. Wang, G.D. Pearlson, D.C. Glahn, and A. Anticevic. 2014. Altered global brain signal in schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*. 111:7438-7443.

Yeo, B.T., J. Tandi, and M.W. Chee. 2015. Functional connectivity during rested wakefulness predicts vulnerability to sleep deprivation. *NeuroImage*. 111:147-158.

Yildiz, S., S. Thyagaraj, N. Jin, X. Zhong, S. Heidari Pahlavian, B.A. Martin, F. Loth, J. Oshinski, and K.G. Sabra. 2017. Quantifying the influence of respiration and cardiac pulsations on cerebrospinal fluid dynamics using real-time phase-contrast MRI. *Journal of magnetic resonance imaging : JMRI*. 46:431-439.

Zarahn, E., G.K. Aguirre, and M. D'Esposito. 1997. Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. *NeuroImage*. 5:179-197.