Task-evoked activity quenches neural correlations and variability across cortical areas

Authors: Takuya Ito[1,2,*], Scott L. Brincat[3], Markus Siegel[4,5,6], Ravi D. Mill[1], Biyu J. He[7], Earl K. Miller[3], Horacio G. Rotstein[1,8,9], Michael W. Cole[1]

[1]Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ 07102
[2]Behavioral and Neural Sciences PhD Program, Rutgers University, Newark, NJ 07102
[3]The Picower Institute for Learning and Memory, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139
[4]Centre for Integrative Neuroscience & MEG Center, University of Tübingen, 72076 Tübingen, Germany
[5]Hertie Institute for Clinical Brain Research, University of Tübingen, 72076 Tübingen, Germany
[6]MEG Center, University of Tübingen, 72076 Tübingen, Germany
[7]Neuroscience Institute, Langone Medical Center, New York University, New York, NY 10016
[8]Federated Department of Biological Sciences, Rutgers University, Newark, NJ 07102
[9]Institute for Brain and Neuroscience Research, New Jersey Institute of Technology, Newark, NJ 07102

Corresponding Author:
Takuya Ito
Center for Molecular and Behavioral Neuroscience, Rutgers University
197 University Avenue, Newark, NJ 07102
taku.ito1@gmail.com

Conflict of Interest: None

Acknowledgements: We thank Drew B. Headley, Bart Krekelberg, Olaf Sporns, and Mauricio Delgado for helpful feedback on earlier versions of the manuscript. The authors acknowledge support by the US National Institutes of Health under awards R01 AG055556 and R01 MH109520, by the European Research Council (ERC StG335880), and the Centre for Integrative Neuroscience (DFG, EXC 307). Some of the data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.

Summary

Many large-scale functional connectivity studies have emphasized the importance of communication through increased inter-region correlations during task states. In contrast, local circuit studies have demonstrated that task states primarily reduce correlations among pairs of neurons, likely enhancing their information coding by suppressing shared spontaneous activity. Here we sought to adjudicate between these conflicting perspectives, assessing whether co-active brain regions during task states tend to increase or decrease their correlations. We found that variability and correlations decrease across a variety of cortical regions in two highly distinct data sets: non-human primate spiking data and human functional magnetic resonance imaging data. Moreover, this observed variability and correlation reduction was accompanied by an overall increase in dimensionality (reflecting less information redundancy) during task states, suggesting that decreased correlations increased information coding capacity. We further found in both spiking and neural mass computational models that task-evoked activity increased the stability around a stable attractor, globally quenching neural variability and correlations. Together, our results provide an integrative mechanistic account that encompasses measures of large-scale neural activity, variability, and correlations during resting and task states.

Keywords: functional connectivity, neural correlations, neural variability, dynamical systems

Introduction

Measures of neural correlations and variability are widely used in neuroscience to characterize neural processes. During task states, neural variability has consistently been shown to be reduced during tasks across human functional magnetic resonance imaging (fMRI) (He, 2011, 2013), local neural populations (Churchland et al., 2010; Hennequin et al., 2018; Jacobs et al., 2018), and both spiking (Hennequin et al., 2018; Litwin-Kumar and Doiron, 2012) and mean-field rate models (Deco and Hugues, 2012; Ponce-alvarez et al., 2015). Despite this convergence in the neural variability literature, there are disparities in the use and interpretation of neural correlations. In the human fMRI literature, neural correlations are often estimated by measuring the correlation of blood oxygenated level-dependent (BOLD) signals and is commonly referred to as functional connectivity (FC) (Biswal et al., 1995). In the animal spiking literature, neural correlations have been measured by computing the correlation between the spike rate of two or more neurons and is commonly referred to as the spike count correlation (or noise correlation) (Cohen and Kohn, 2011). Yet despite the use of different terms, the target statistical inference behind these two techniques is consistent: to characterize the interaction among neural units.

In the human fMRI literature, studies have identified large-scale functional brain networks through clustering sets of correlated brain regions using resting-state activity (Ji et al., 2018; Power et al., 2011; Yeo et al., 2011). During task states, the FC structure has been demonstrated to dynamically reconfigure (Cole et al., 2014; Gonzalez-Castillo and Bandettini, 2017; Krienen et al., 2014). Though it has been suggested that correlation increases and decreases respectively facilitate and inhibit inter-region communication (Gonzalez-Castillo and Bandettini, 2017; Tomasi et al., 2014), the mechanistic bases of these FC changes remain unclear.

Studies in the local circuit literature using electrophysiological recordings in animals have characterized the correlation structure between neuron spikes across a range of task demands. These studies have found that the spike count correlation between neuron spikes generally decrease during task states, particularly for neurons that are responsive to the task (Cohen and Maunsell, 2009; Ecker et al., 2010; Pinto et al., 2019; Ruff and Cohen, 2014). Though the reduction in spike count correlations during task states has been linked to better information coding by the local neuronal population by suppressing shared spontaneous activity and reducing noise (Averbeck et al., 2006; Cohen and Kohn, 2011), the mechanism underlying these reduced correlations remains unclear.

These differing perspectives of the role of neural correlations in (large-scale) human imaging and (local circuit) animal neurophysiology studies appear to be incompatible. Moreover, it is unclear whether observations at the local circuit level would be consistent with observations made across large cortical areas. To further complicate this issue, we recently found that task activations can inflate task functional connectivity estimates in human fMRI data (Cole et al., 2019), suggesting some previous neuroimaging studies may have erroneously reported correlation increases due to inaccurate removal of the mean-evoked response (a standard

procedure in the spiking literature). Here we sought to adjudicate the differing perspectives in the neural correlation literature while simultaneously confirming and extending previous findings on task-state neural variability reduction.

We report multiple sources of empirical and theoretical evidence demonstrating that task-evoked activity quenches neural correlations and variability across cortical areas. First, we characterize task-evoked neural variability and correlations in empirical data using two highly distinct data sets: multi-site and multi-unit NHP electrophysiology and whole-brain human fMRI. This allowed us to test whether there were consistent large-scale variability and correlation changes during task states independent of data acquisition technique. Moreover, this allowed us to take advantage of the more direct neural recording with NHP electrophysiology along with the more comprehensive coverage of human fMRI (in addition to testing for translation of findings to humans). Next, to provide a mechanistic account capable of explaining our empirical findings, we used both spiking and neural mass models to parsimoniously explain variability and correlation suppression across mean-field cortical areas. This led to theoretical insight using dynamical systems analysis, demonstrating that task-evoked activity strengthens the system's attractor dynamics around a stable fixed point in neural mass models, quenching neural correlations and variability. The combination of simultaneously recorded multi-unit recordings from six cortical sites, whole-brain fMRI obtained from seven different cognitive tasks, and dynamical systems modeling and analysis provide a comprehensive account of task-related correlation and variability changes spanning species and data acquisition techniques.

Results

We first show empirically that task-evoked activity suppresses neural correlations and variability across large cortical areas in two highly distinct neural data sets: NHP spiking and human fMRI data (Figure 1). This confirms previous findings showing quenched neural variability during task states in both NHPs and humans (Churchland et al., 2010; He, 2011, 2013; Hennequin et al., 2018), while going beyond those previous studies to report globally quenched inter-area task-state neural correlations. In particular, we focused on neural variability and correlation changes across large cortical areas in our electrophysiology data set (rather than between pairs of neurons) given our focus on large-scale neural interactions, and to facilitate a comparison between different correlation approaches (FC in fMRI data and spike count correlation in electrophysiology data). Moreover, we limited our inferences to neural interactions between large cortical areas, since the complexities of spike count correlations between pairs of neurons with differential receptive fields cannot be accounted for at the mean-field (Ruff and Cohen, 2014, 2016). Following our empirical results, we provide a mechanistic framework using computational simulations and detailed dynamical systems analyses to explain the quenching of neural variability and correlations during task-evoked states.

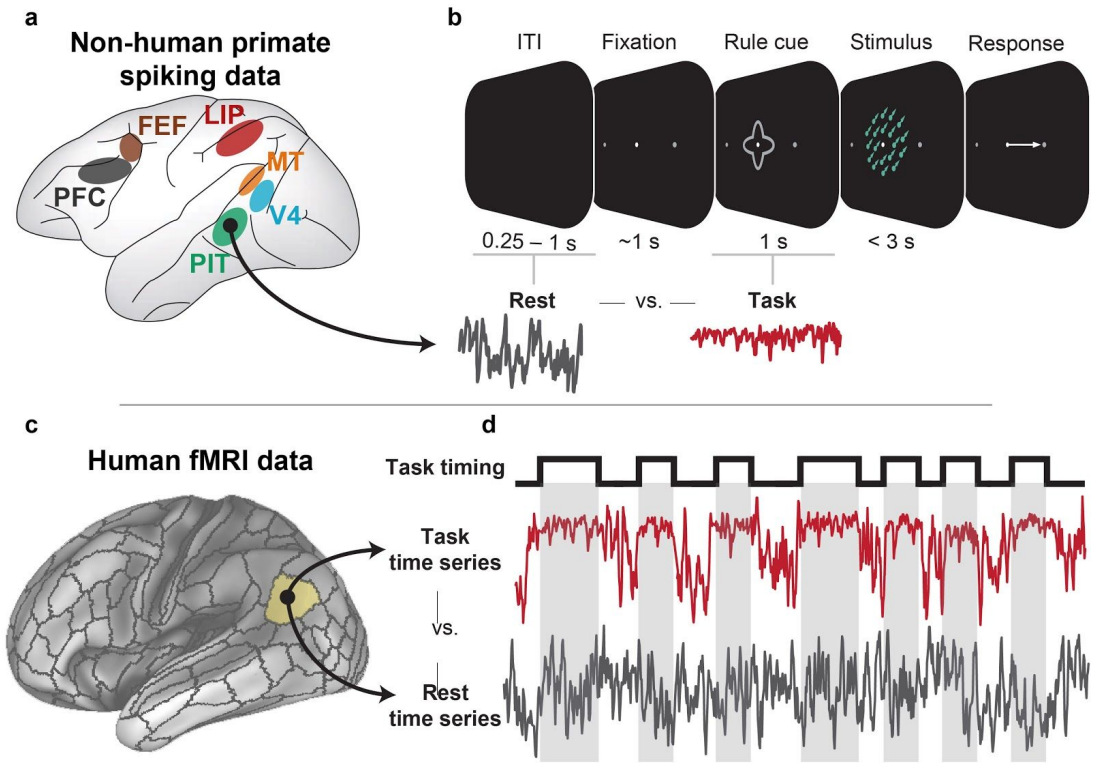


Figure 1. Testing the hypothesis that task-evoked neural variability and correlations are quenched across cortical areas in NHP spiking and human fMRI data sets. We used two highly distinct data sets to test the hypothesis that task-evoked activity globally quenches neural variability and correlations to suppress background spontaneous activity/noise. This contrasts with the alternate hypothesis, namely that task-evoked activity increases variability and correlation to facilitate inter-region communication. **a,b)** Using multi-unit spiking data collected simultaneously from six different cortical areas (Siegel et al., 2015), we compared the spiking variability and spike count correlations between task-state (i.e., following task cue onset) and rest-state spiking activity. We defined rest state as the inter-trial interval (ITI) directly preceding the trial. This was performed using the spike rate averaged within each cortical area, allowing us to target the activity of large neural populations. **c,d)** Using human fMRI data obtained from the Human Connectome Project (Barch et al., 2013), we compared the neural variability and correlations (i.e., FC) of the BOLD signal during task block intervals to equivalent resting-state intervals. We used seven highly distinct cognitive tasks. Time series and task timings are illustrative, and do not reflect actual data.

Task onset reduces neural variability and correlations across spiking populations in NHPs

We estimated the spiking variability and spike count correlations of cortical populations in NHPs following task cue onset (task periods) and during the inter-trial intervals (ITI) (rest periods). We found that across trials, global spiking variability and spike count correlations (r_{sc}) decreased during task as compared to rest (exploratory subject, variance diff $=-0.005$, $t_{308}=-12.92$, $p<0.0001$; r_{sc} diff = -0.04, $t_{308}=-4.29$, $p<0.0001$ Figure 2C,D; replication subject, variance diff $=-0.02$, $t_{830}=-15.36$, $p<0.0001$; r_{sc} diff = -0.03; $t_{830}=-7.44$; $p<0.0001$

Supplementary Figure 1C,D). These global reductions were also observed using spike count covariance, rather than correlations (Figure 2H). In addition, we demonstrated that variability and correlation decreased within trial (across time points within a trial), demonstrating that task state quenching also occurs on a moment-to-moment basis, rather than only on a slower trial-to-trial timescale (Supplementary Figure 2). We also measured the spiking variability for each cortical area separately, finding that 5/6 cortical areas reduced their spiking variability during task states in the exploratory subject (all areas except for V4, FDR-corrected $p < 0.0001$). In the replication subject, all cortical areas, including V4, reduced their spiking variability (FDR-corrected $p < 0.01$). Similarly, we found that during task states, the spike count correlation significantly decreased for a majority of cortical areas (FDR-corrected $p < 0.05$; Figure 2D-G).

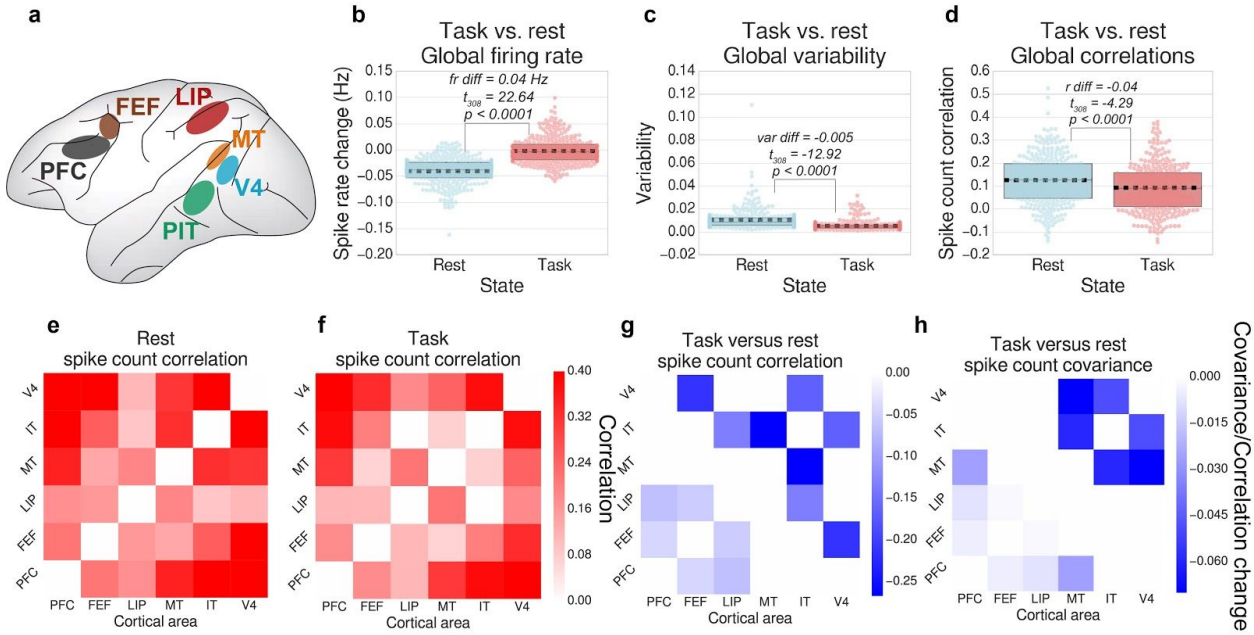


Figure 2. Neural variability and correlations decrease during task states relative to rest in spiking data. Results for the replication subject are reported in Supplementary Figure 1. **a)** We obtained multi-unit spike recordings from six different cortical areas during a motion-color categorization task. **b)** We calculated the average spike rate across all recordings during the rest period (ITI) and task period (task cue), across trials. **c)** We calculated the cross-trial spiking variance for each region during task and rest states, and then averaged across all regions. **d)** We calculated the average cross-trial neural correlation for task and rest states between all pairs of recorded brain regions. (Spike rates were averaged within each cortical area.) **e-g)** For each pair of brain regions, we visualize the correlation matrices between each recording site for the averaged rest period, task period, and the differences between task versus rest state spike count correlations. **h)** We also observed no increases in covariance (non-normalized correlation) (Cole et al., 2016a; Duff et al., 2018; Siegel et al., 2012). For panels **e-h**, plots were thresholded and tested for multiple comparisons using an FDR-corrected p<0.05 threshold. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, grey line indicates the median, and the distribution is visualized using a swarm plot.

Though we observed some task-state correlation increases in the replication subject (Supplementary Figure 1G), we were unable to identify any spike count correlation increases in the exploratory subject (Figure 2G). Importantly, when calculating the spike count covariance

(rather than correlation), we did not find significant covariance increases for either subject (Figure 2H, Supplementary Figure 1H). This suggests that the observed correlation increases in the replication subject are not due to increases in shared variance, but are rather primarily driven by decreases in unshared variance (i.e., private neural noise) (Cole et al., 2016a; Duff et al., 2018; Siegel et al., 2012). These results are consistent with the results observed in the fMRI data (reported below), demonstrating that the variability and correlations across cortical areas largely decrease during task states.

To ensure that correlation and variability decreases were associated with increases in activity (rather than just the task period), we estimated the mean spike rate across all regions during the task cue interval and the preceding ITI. Indeed, we found that the mean spike rate during task states was significantly greater than the mean spike rate during rest (exploratory subject, task vs rest firing rate difference $= 0.04$ Hz, $t_{308} = 22.64$, $p < 0.0001$, replication subject, rate difference $= 0.05$ Hz, $t_{830} = 32.46$, $p < 0.0001$). These findings suggest that task states increase neural activity while quenching spiking variability and spike count correlations across large cortical areas.

Task-state variability is globally quenched across a wide battery of tasks in human fMRI data

Consistent with the spiking literature, previous work in the fMRI literature has demonstrated that increased activity associated with task-evoked states quenches neural variability (Churchland et al., 2010; He, 2011, 2013). We extend those findings to evaluate variability quenching across seven additional cognitive tasks in humans using data from the Human Connectome Project (HCP) (Van Essen et al., 2013). We calculated the variability (estimated using time series variance) during task blocks, averaged across tasks and across regions. Consistent with our hypothesis, we found that the global variability during task blocks was significantly lower than the variability during equivalent periods of resting-state activity (exploratory cohort variance difference = -0.019, $t_{175} = -23.89$, $p < 0.0001$; replication cohort variance difference = -0.019, $t_{175} = -20.72$, $p < 0.0001$; Figure 3A). These findings suggest that task states are associated with task-evoked variability reductions, consistent with our hypothesis and previous empirical accounts (He, 2013).

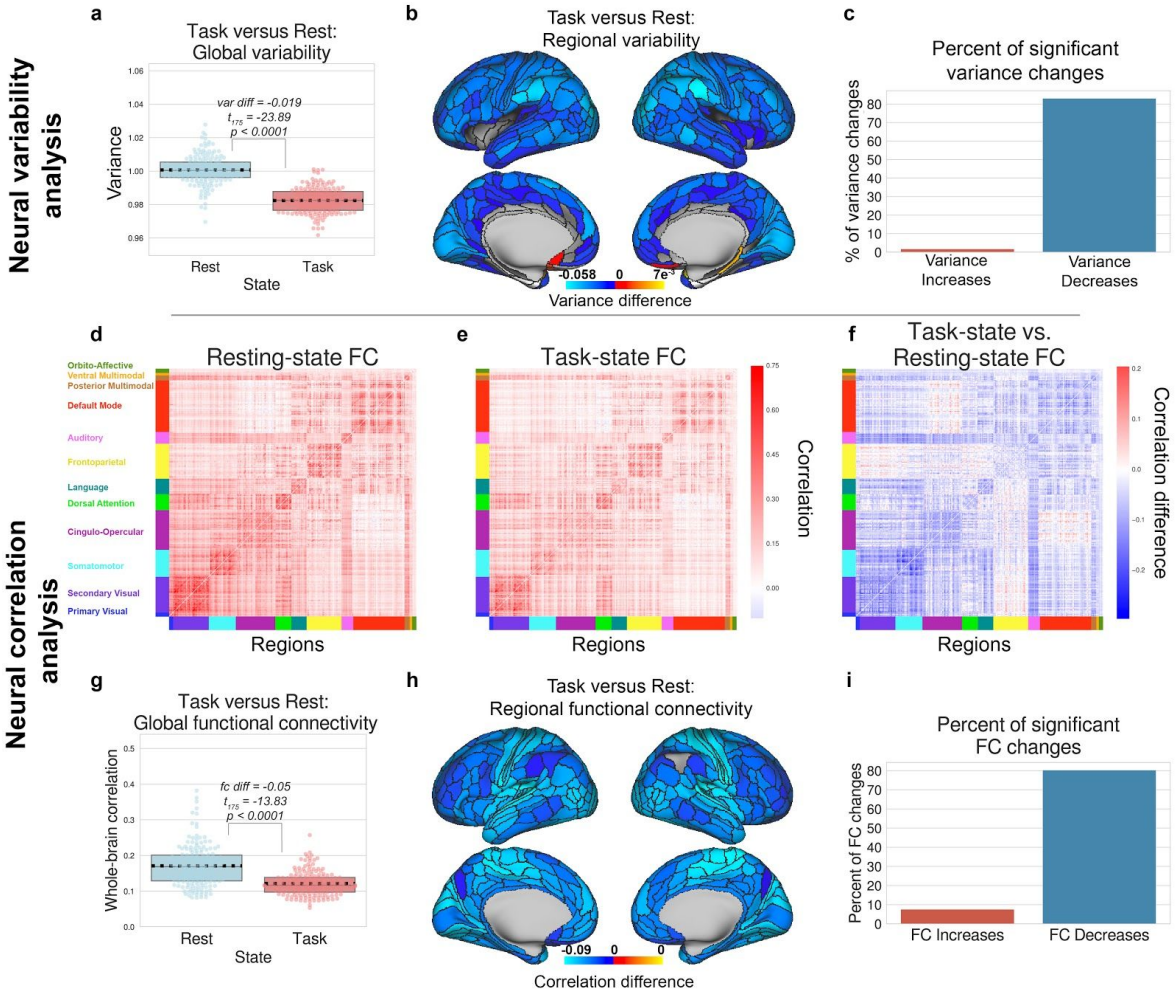


Figure 3. Neural variability and correlations decrease during task states in human fMRI data. Figures for the replication cohort are in Supplementary Figure 4. Figures for each task separately are shown in Supplementary Figures 8 and 9. **a)** We first compared the global variability during task and rest states, which is averaged across all brain regions, and then **b)** computed the task- versus rest-state variability for each brain region. **c)** The percentage of brain regions with significant increases/decreases of neural variability during task. **d)** We next compared the correlation matrices for resting state blocks with **(e)** task state blocks, and **(f)** computed the task- versus rest-state correlation matrix difference. **g)** We found that the average FC between all pairs of brain regions is significantly reduced during task state. **h)** We found that the average correlation for each brain region, decreased for each brain region during task state. **i)** We calculated the percentage of correlation increases/decreases for all pairs of brain regions. For panels **b-f, h,** and **i**, plots were tested for multiple comparisons using an FDR-corrected p<0.05 threshold. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, grey line indicates the median, and the distribution is visualized using a swarm plot.

To better understand how global this phenomenon was, we plotted the change in variability from rest to task for each brain region separately. We found that almost all brain regions significantly reduced their variability from rest to task, suggesting that variability reduction occurs across most brain regions (cortical maps are thresholded using an FDR-corrected threshold of $p < 0.05$; Figure 3B,C). This finding extends the work of a previous

study in human fMRI data during a finger tapping task (He, 2011; Ponce-alvarez et al., 2015), suggesting that task-induced variability reduction is a general phenomenon consistent across most cortical regions, and across a wide variety of cognitive tasks.

Lastly, we evaluated whether variability quenching occurred during task blocks relative to inter-block intervals (rather than comparing task runs to resting-state runs). Since we z-normalized each task run with unit variance, we could evaluate the degree to which variability was quenched during task blocks relative to inter-block intervals by computing the average variance during task blocks relative to 1. (Note that z-normalization of the task time series was performed after removing the mean-evoked task response via a task GLM, such that reduced variability was not an artifact of preprocessing/z-normalizing the time series.) Indeed, we found that the variance during task blocks was reduced relative to the inter-block intervals (exploratory cohort variance - 1 = -0.019, $t_{175} = -36.58$, $p < 0.0001$; replication cohort variance difference = -0.018, $t_{175} = -33.01$, $p < 0.0001$). Our findings demonstrate that task-evoked periods quench neural variability relative to both resting-state activity and inter-block intervals.

Task-state FC is globally quenched across a wide battery of tasks in human fMRI data

Despite multiple studies describing task-evoked FC changes (Cole et al., 2014; Gonzalez-Castillo and Bandettini, 2017; Krienen et al., 2014), the precise mechanisms of how FC can change remain unclear. Our current findings illustrate that mean-field spike count correlations decrease during task-evoked states, consistent with previous literature that focused on local circuits (Churchland et al., 2010; Cohen and Kohn, 2011). Compounding the complexity of interpreting this contrast, we recently found that task-evoked activations can inflate task FC estimates due to the hemodynamic lag of the BOLD signal (Cole et al., 2019). Consistent with the spiking literature's perspective on spike count correlations, and the theoretical evidence suggesting that the correlation of ongoing spontaneous activity should be suppressed during task to facilitate information coding (Averbeck et al., 2006), we hypothesized that when properly accounting for the mean-evoked response (as is commonly done in the spiking literature), FC would also be globally reduced during task states.

We first calculated the mean FC across all pairwise correlations across all cortical regions for both task and rest states (Figure 3D-F). We found that during task states, the global FC was significantly reduced relative to resting-state fMRI (exploratory cohort FC diff = -0.05, $t_{175} = -13.83$, $p < 0.0001$; replication cohort FC diff = -0.046, $t_{175} = -14.00$, $p < 0.0001$; Figure 3G). Recent studies have suggested that the use of correlation provides an ambiguous description of how shared variability (relative to unshared variability) change between brain areas (Cole et al., 2016a; Duff et al., 2018). Thus, to generalize these results, we also calculated FC using covariance rather than correlation using non-normalized BOLD data, finding that covariance also globally decreases (covariance diff = -192.96, $t_{351} = -27.30$, $p < 0.0001$ Supplementary Figure 5). Task-evoked global FC was also reduced in each of the 7 HCP tasks separately (all tasks FDR-corrected $p < 0.0001$; Supplementary Figure 9A). To identify exactly how global this phenomenon was, we plotted the average task versus rest FC change for each brain region (Figure 3H,I). We found that nearly all cortical regions significantly reduced their

correlation with the rest of cortex during task states. To ensure that correlation differences between rest and task states were not associated with in-scanner head motion, we calculated the average number of motion spikes during rest and task scans using a relative root mean squared displacement threshold of 0.25mm (Ciric et al., 2017). For both the exploratory and replication cohorts, we found no significant differences in the percentage of motion spikes between rest and task states (exploratory set, average task=0.91% of frames, average rest=0.81% of frames, $t_{175} = 1.08$, $p = 0.28$; replication set, average task=0.009% of frames, rest=0.008% of frames, $t_{175} = 1.53$, $p = 0.12$).

While we primarily observed global decreases in FC, a small portion of connections increased their FC during task states (exploratory cohort, 7.59% of all connections; replication cohort, 9.07% of all connections; FDR-corrected *p*<0.05) (Figure 3I, Supplementary Figure 4I). However, FC increases were typically limited to cross-network correlations between networks with different functions, where baseline resting-state FC is already quite low (e.g., cingulo-opercular network with the default mode network, or the frontoparietal network with the visual network) (Figure 3D-F).

Task-evoked activity is negatively correlated with neural variability and correlations in human fMRI data

We hypothesized that decreases in neural variability and correlations are related to task-evoked activity. To provide evidence for this hypothesis, we computed the mean task-evoked activity (averaging across all regions). We found that the global activity was significantly greater than baseline across different task states (exploratory cohort, $t_{175} = 6.46$, $p < 0.0001$; replication cohort, $t_{175} = 12.63$, $p < 0.0001$), demonstrating that decreases in neural variability and FC were accompanied by global increases in evoked activity.

Previous work has shown that regions that are more active during task tend to have greater variability reductions (He, 2013). We sought to replicate this effect in the current data set, while extending those results to demonstrate that more task-active regions also tend to reduce their FC during task states. We first correlated regional task-evoked activity with task-evoked variability reduction (task variance minus rest variance) across regions at the group-level. We found that regions with more task-evoked activity (averaged across tasks) exhibited greater variability reductions during task states, confirming previous findings in a finger tapping task (exploratory cohort rho=-0.32, *p*<0.0001; replication cohort rho=-0.49, *p*<0.0001; Supplementary Figure 3A,C) (He, 2013). This negative relationship was also observed in 6/7 of the HCP tasks when analyzed separately (FDR-corrected *p*<0.01; Supplementary Figure 8B). To link regional task activations with FC decreases, we tested for a correlation between regional task-evoked activity and the average FC change during task states for each region. Consistent with our hypothesis, we found that regions with greater task-evoked activity (averaged across tasks) reduced their average FC more during task states (exploratory cohort rho=-0.25, *p*<0.0001; replication cohort rho=-0.20, *p*=0.0002; Supplementary Figure 3B,D). When tasks were analyzed separately, this negative correlation was observed in 4/7 of the HCP tasks

(FDR-corrected $p<0.05$; Supplementary Figure 9B). Thus, brain areas with increased levels of task-evoked activity tend to reduce both their task-evoked variability and FC.

The information-theoretic relevance of task state reduction of neural correlations

Results from our empirical data converged across imaging modalities and species, illustrating that task states increased mean-evoked activity while reducing neural variability and correlations. However, the theoretical implication of a decreased correlated task state remains unclear. Here we sought to better characterize the information-theoretic implication of a global reduction in neural correlations. In particular, we hypothesized that reductions in neural correlations increase the collective information across units by suppressing background spontaneous activity/noise and reducing information redundancy, supporting robust information representations (Averbeck et al., 2006). Consistent with previous large-scale computational models (Abbott et al., 2011; Deco et al., 2014), this predicts that the dimensionality of global neural activity would increase during task-evoked states, increasing its collective information. We note that an increase in dimensionality is not trivially implied by decreased global correlations. Because we found that regional time series variance also decreases during task states, the neural data dimensionality would increase only if inter-region covariance decreases more than local regional variance (i.e., off-diagonal is reduced more than the diagonal of the variance-covariance matrix).

We measured the dimensionality (Abbott et al., 2011; Litwin-Kumar et al., 2017) of the neural activity (for human fMRI and NHP spiking data) during rest and task states (see Methods). Consistent with our hypothesis, we found that task states increased their overall dimensionality relative to rest states (fMRI task versus rest, exploratory cohort difference = 16.13, $t_{175} = 19.31$, $p < 0.0001$, replication cohort difference = 15.78, $t_{175} = 21.66$, $p < 0.0001$; NHP task versus rest, exploratory subject difference = 0.13, $t_{308} = 6.52$, $p < 0.0001$, replication subject difference = 0.16, $t_{830} = 7.90$, $p < 0.0001$) (Figure 4). We also found that when analyzing each of the 7 HCP tasks separately, dimensionality increased in all 7 tasks relative to resting state (FDR-corrected $p < 0.0001$; Supplementary Figure 10). The present results suggest that task states are associated with a decrease in neural variability and correlations, reflecting a suppression of shared and private spontaneous activity, which increases the effective dimensionality of neural activity.

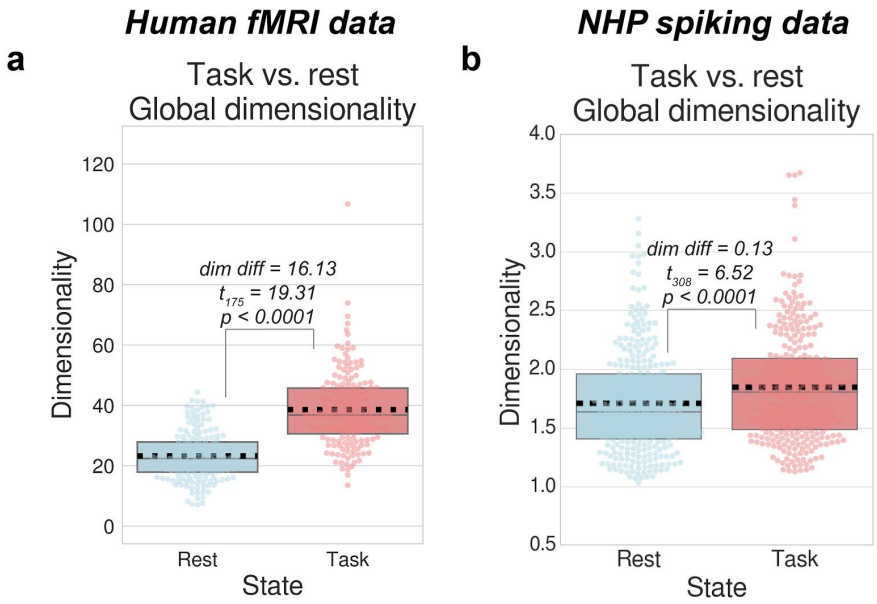


Figure 4. Neural dimensionality increases during task periods relative to resting-state activity. a) For each subject, we calculated the dimensionality during task and rest states and found that during task states, dimensionality significantly increased. **b)** We calculated the dimensionality of spiking activity across trials and found that during task states, dimensionality significantly increased. These findings provide a potential information-theoretic interpretation of neural correlation and variability reduction during task states. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, grey line indicates the median, and the distribution is visualized using a swarm plot.

From neurons to neural masses: Modeling neural dynamics of cortical areas

In the previous sections, we provided empirical evidence that task states reduce inter-area correlations and variability in multi-unit spiking and fMRI data. In this section, we construct a biologically plausible model that provides a parsimonious explanation of correlation and variability reductions in mean field spiking networks and cortical BOLD dynamics.

Neurophysiologically, functional brain areas are composed of local circuits with balanced excitatory and inhibitory neural activity (Figure 5A). In previous work, local circuits have been demonstrated to have clustered excitatory connections (Song et al., 2005), leading to slow dynamics and high variability in spiking networks simulated *in silico* (Litwin-Kumar and Doiron, 2012). Using this previously established model, we systematically perturbed this balanced network under a distribution of inputs to estimate the mean-field transfer function of a cortical population. We found that the population transfer function approximated a sigmoid activation function (Figure 5B). Approximating the mean-field transfer function of a cortical area allowed us to focus our modeling efforts on simplified excitatory networks across large cortical areas, since most inter-area cortical networks rely on excitatory connectivity (Joglekar et al., 2018).

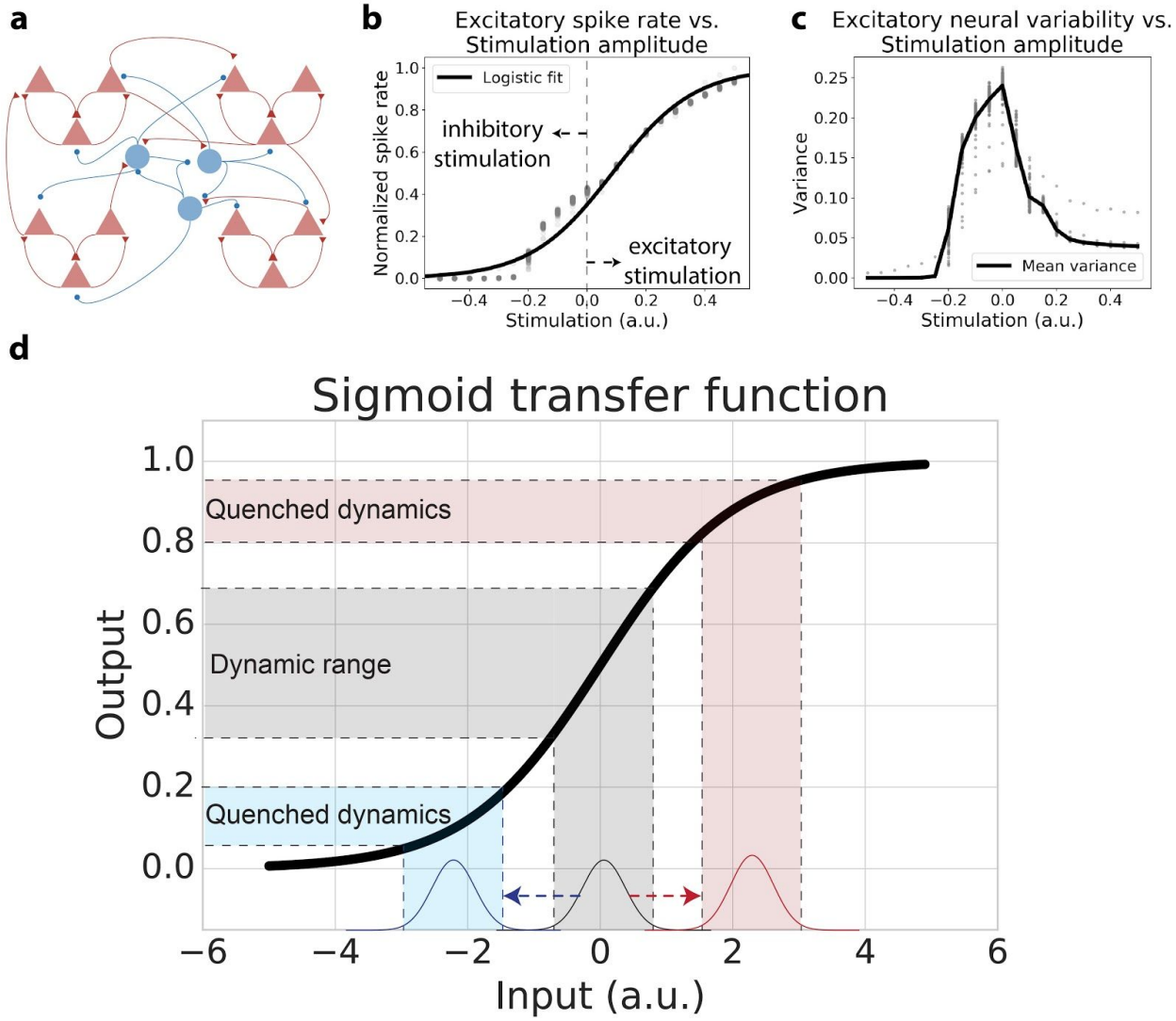


Figure 5. Inferring the mean-field transfer function of a neural population with clustered excitatory connections. a) Schematic illustration of the balanced spiking model with clustered excitatory connections. Network architecture and parameters are identical to those reported in (Litwin-Kumar and Doiron, 2012). Red triangles indicate excitatory cells, blue circles indicate inhibitory cells. **b)** The population spike rate (excitatory cells only) subject to inhibitory regulation. We systematically stimulated a subset of the neural population and measured the corresponding mean excitatory spike rate. Spike rates were normalized between 0 and 1. Excitatory stimulation was implemented by stimulating 400 excitatory neurons, and inhibitory stimulation was implemented by stimulating 400 inhibitory neurons. Spiking statistics were calculated across 30 trials, with each point in the scatter plot indicating a different 50ms time bin. **c)** Population neural variability (excitatory cells only), as a function of input stimulation. **d)** Based on panel b, we approximated the mean field neural transfer function as a sigmoid. A sigmoid transfer function produces optimal input-output dynamics for a narrow range of inputs (gray). The same input distribution mean shifted by some excitatory/inhibitory stimulation produces a quenched dynamic range.

In this balanced spiking network, any evoked stimulation, excitatory or inhibitory, would result in reduced variability (Figure 5C). Specifically, the magnitude of stimulation was

negatively correlated with spiking variability in the balanced spiking model (rho = -0.92; $p<0.0001$). While previous studies have suggested that the mean and variance of the spike rate may be independent of each other, those studies focused on mean-matching the spike rate of individual neurons within the same local population (Churchland et al., 2010; Litwin-Kumar and Doiron, 2012). Here we found a negative association between mean and variance under experimental perturbation, suggesting that at the mean-field level, mean and variance cannot be mechanistically dissociated. Based on these considerations, we hypothesized that during periods in which global neural activity levels are elevated, such as task states, both neural variability and correlations would be globally quenched.

Neural variability is quenched during task-evoked states in a neural mass model

Here we rigorously ground the intuition that task-evoked activity reduces output variability using neural mass modeling and dynamical systems theory. A recent study provided evidence that an evoked stimulus drives neural populations in sensory cortex around a stable fixed point attractor (Hennequin et al., 2018). We first extended these findings using a simplified neural mass model, which allows for a comprehensive dynamical systems analysis that is mathematically difficult in higher dimensions. Additionally, this enabled a simpler theoretical approach to investigating changes in neural dynamics that are generalizable across mean-field neural cortical areas (i.e., populations with sigmoidal transfer functions).

We first characterized the relationship between task-evoked and spontaneous activity in a large neural population using a single neural mass unit. We simulated the neural population's dynamics across a range of fixed input strengths (Figure 6A), finding a nonlinear relationship between stimulus strength and the observed variability of the neural population (Figure 6C). We found that variability was highest when there was no stimulation, while variability decreased for any type of evoked stimulation (e.g., negative or positive input amplitudes). Despite the model's simplicity, these findings are consistent with our (and others') empirical and model results demonstrating that task states quench time series variability in both human and animal data (Churchland et al., 2010; He, 2011; Hennequin et al., 2018).

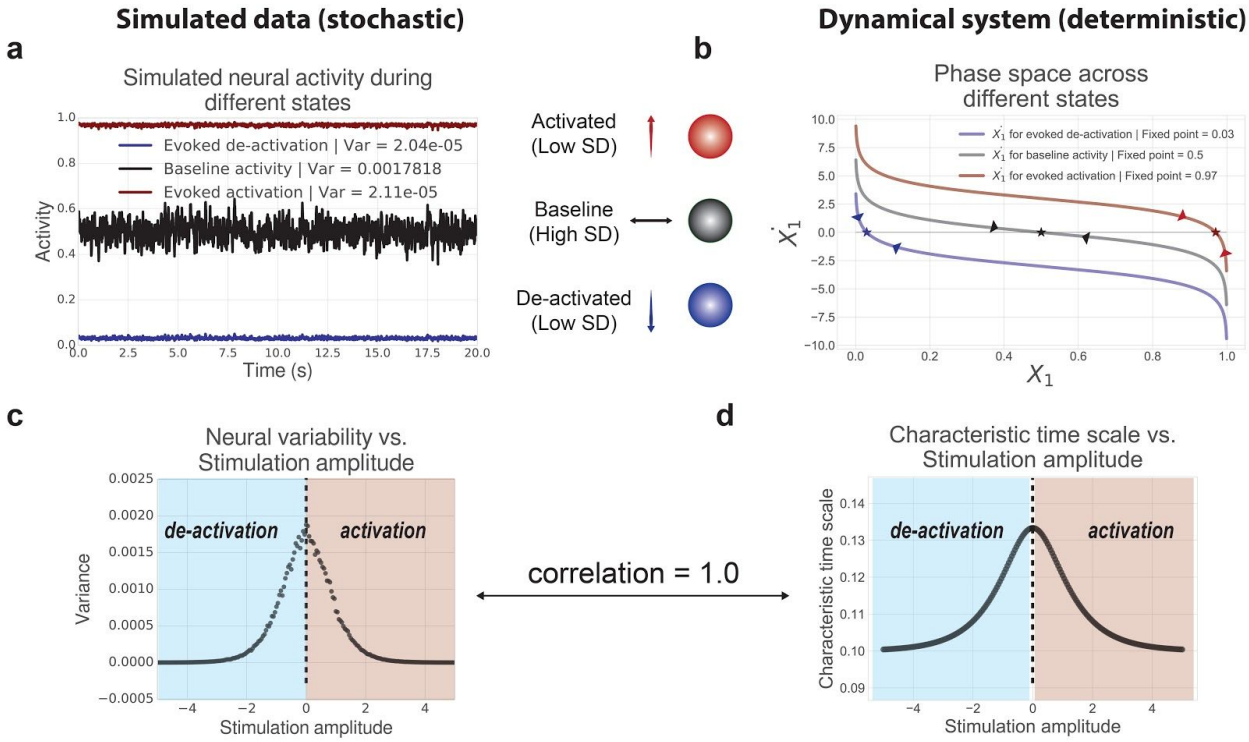


Figure 6. Task-evoked activity induces changes in neural variability and the underlying attractor dynamics. Our minimal modeling approach directly links descriptive statistics (e.g., time series variability) with rigorous dynamical systems analysis (e.g., attractor dynamics). **a)** During different evoked states (i.e., fixed inputs), there is a reduction in the observed time series variability (measured by variance across time). This is directly related to how input-output responses change due to the changing slope in the sigmoid transfer function. **b)** We visualized the phase space for each of the neural populations according to state by plotting the derivative of X_1 denoted by $\dot{X}_1$. For each state, we calculated the fixed point attractor (plotted as a star), denoting the level of mean activity the system is drawn to given some fixed input (or absence thereof). Arrows denote the direction/vector toward each fixed point, which specify the characteristic time scale (i.e., the speed) the system approaches the fixed point. **c)** We ran simulations across a range of stimulation amplitudes, calculating the variance across time at each amplitude. **d)** We characterized the shifting attractor dynamics for each stimulus by computing the characteristic time scale at the fixed point for each stimulation amplitude. The characteristic time scale across all fixed points is perfectly correlated with the neural variability of the simulated time series across all fixed inputs (rank correlation = 1.0).

We sought to leverage the model's simplicity to characterize dynamical systems properties governing the observed neural variability. This would provide rigorous evidence that shifting the underlying attractor dynamics alters the observed neural signals. We first performed a state space analysis in one dimension to identify the stable fixed point attractor (i.e., the equilibrium level of activity the system is drawn to during a particular state) for the intrinsic and evoked states (Figure 6B). The state space view enabled visualization of the system's full dynamics across different evoked states (Figure 6B). For example, dynamics around the fixed point attractor in the intrinsic baseline (rest) state appeared to approach equilibrium slowly. This can be identified by observing the angle where the curve intersects 0 on the y-axis (i.e., when

$\dot{x} = 0$; Figure 6B). The angle of this curve corresponds to the *characteristic time scale*, a dynamical property characterizing the speed with which the system approaches the attractor (a higher value reflects slower dynamics; see Methods) (Strogatz, 1994).

To quantify this more rigorously, we performed a linear stability analysis around the fixed point attractor of the system across the same range of stimulation amplitudes. For each input, we analytically calculated the characteristic time scale at each fixed point. Again, we found a nonlinear relationship between the amplitude of the stimulus and the characteristic time scale of the neural population (Figure 6D), and found that the characteristic time scale explained 100% of the variance of the simulated stimulus-evoked variability (Figure 6C). These results demonstrate that changes in observed neural variability can be directly attributed to changes in the underlying attractor dynamics.

To ensure that the model explanation would generalize to data obtained on a slower time scale (e.g., fMRI BOLD data), we transformed the simulated neural activity into fMRI BOLD activity using the Balloon-Windkessel model (Friston et al., 2003). The Balloon-Windkessel is a nonlinear transformation of neural activity to model the BOLD signal that takes into account the normalized blood volume, blood inflow, resting oxygen extraction fraction, and the normalized deoxyhemoglobin content. Consistent with previous accounts (He, 2013), we found that the characteristic time scale around the fixed point attractor was still strongly correlated with BOLD signal variability (rho=0.97; $p<0.0001$; Supplementary Figure 6).

Neural correlations are quenched during task states in a network model

We generalized the dynamical systems analysis in one dimension to two dimensions, allowing us to focus on correlations across cortical areas. We show illustrations of the state space for intrinsic and task-evoked states (Figure 7B,D), as well as the corresponding time series (Figure 7A,C) for our model. Induced negative activity produced qualitatively similar results to the activated state (Figure 7D) due to subthreshold levels of activity rather than saturating levels of activity.

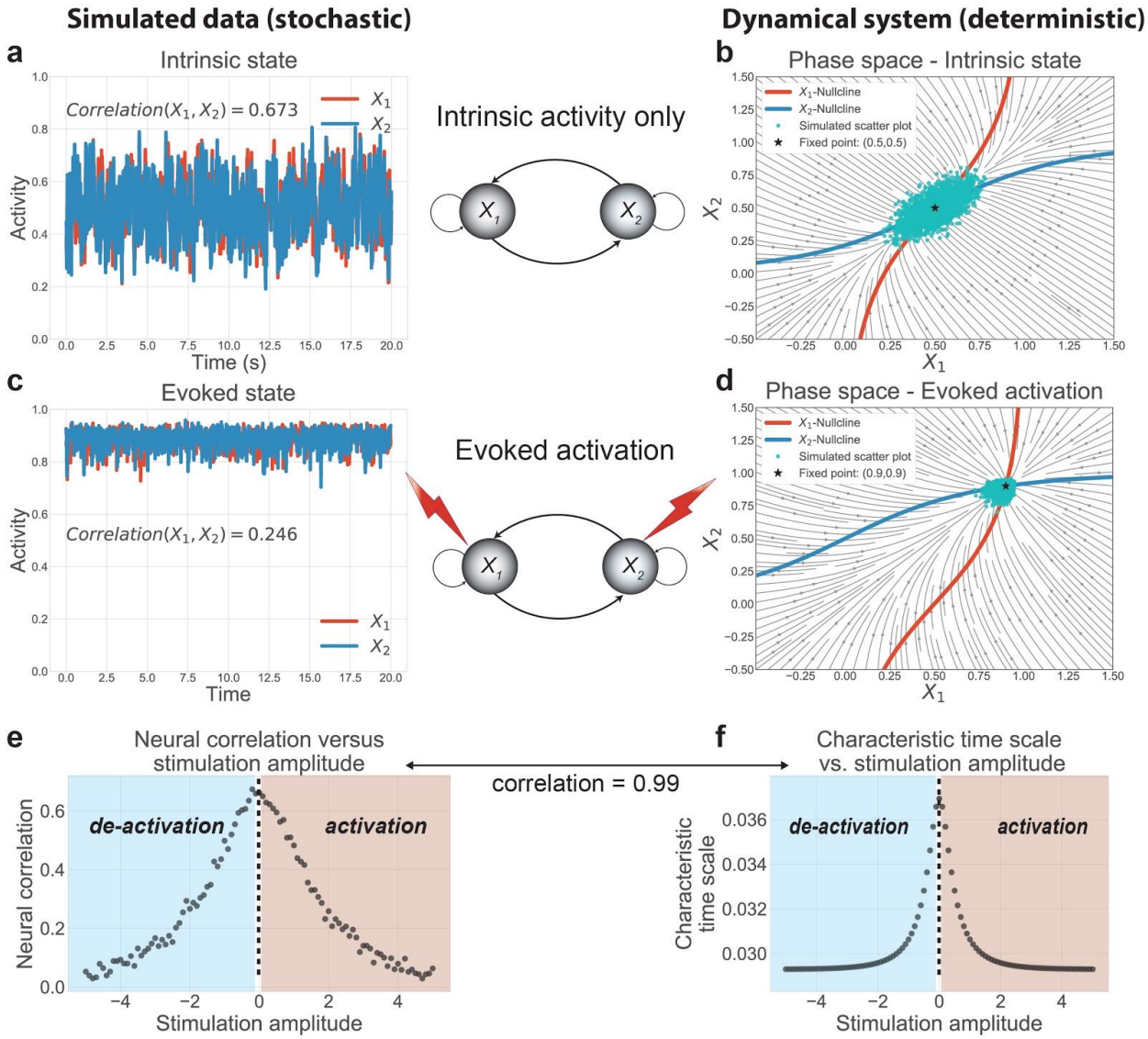


Figure 7. Task-evoked activity quenches neural correlations by altering the underlying attractor dynamics. We used a two unit network model, the minimal model necessary to study dynamic changes in neural correlations. **a)** At baseline, we observed slow, high amplitude fluctuations and high neural correlations. **b)** To characterize the underlying attractor dynamics, we visualized the two-dimensional state space, visualizing the flow field and the nullclines (blue and red curves, where the rate of change is 0) for each unit. The intersection of the two nullclines denote the fixed point attractor. We overlaid the simulated scatter plot (cyan dots) to illustrate the correspondence between the attractor dynamics and simulation. **c)** We injected a fixed input stimulation, shifting the network to an 'evoked' state, which caused a decrease in neural variability and correlation. **d)** The external input transiently moved the fixed point, altering the attractor dynamics and the corresponding scatter plot. **e)** We systematically injected a range of fixed inputs into the network. We found that neural correlations were optimal with no external stimulation, and decreased with any external stimulation. **f)** Across stimulation strengths, we found that the generalized characteristic time scale (see Methods) at the fixed point explained 98% of the neural correlation variance, providing a direct association between the network's attractor dynamics and observed neural correlations.

The state space analysis (Figure 7B,D) allowed us to track the simultaneous evolution of the two neural masses, providing a geometric interpretation of the system. We observed qualitatively that shifts in the attractor dynamics (i.e., changes to the flow field) due to stimulation were directly associated with changes to the correlation between the two neural masses. Specifically, we observed that intrinsic state dynamics supported slower, elongated trajectories along a diagonal axis, consistent with correlated neural activity between the two masses (Figure 7B). This was due to a slower characteristic timescale near the fixed point attractor, which corresponds mathematically to eigenvalues with smaller magnitudes. In contrast, during evoked states, the system approached the fixed point attractor at a faster speed, quenching trajectories in state space that support correlated variability (Figure 7D). Thus, the visualization of the state space demonstrated that changes in neural correlations were associated with changes to the flow field around the fixed point attractor.

To more carefully test the relationship between state-dependent neural correlations, we simulated our network model across a range of fixed input amplitudes. Despite no changes to the network's connectivity structure, we found that neural correlations systematically changed (decreased) as a function of evoked stimulation (Figure 7E). Further, using dynamical systems analysis, we found that a generalization of the characteristic time scale in higher dimensions accounted for changes in neural correlations (rho=0.99; p<0.0001; Figure 7F). In other words, we analytically determined that evoked stimulation shifted the attractor dynamics, changing the neural correlations in a network model with fixed synaptic connections. We found consistent results after transforming the neural activity to fMRI BOLD activity using the Balloon-Windkessel model (Friston et al., 2003), finding that changes to the characteristic time scale accounted for changes in BOLD correlations (rho=0.97; *p*<0.0001; Supplementary Figure 7). These results were reproduced using mutual information (rho=0.94; p<0.0001), a nonlinear measure of statistical dependence (MacKay, 2003), and non-parametric rank correlation (rho=0.99; p<0.0001). This suggests that the quenching of shared variance encompasses both parametric and non-parametric linear and nonlinear measures of statistical dependencies.

Discussion

The present results suggest that task-evoked neural activity globally quenches neural time series variability and correlations. We showed this in NHP spiking and human fMRI data, illustrating the generality of the phenomena. This supports the hypothesis that during task states, decreases in neural variability and correlations suppress ongoing spontaneous activity, better supporting information coding (Averbeck et al., 2006). We subsequently provided a dynamical systems model to demonstrate that evoked activity strengthened the system's fixed point attractor, quenching neural variability and correlations. This provided a mechanistic framework to interpret the empirical results. Importantly, the use of a sigmoid transfer function to model mean-field cortical dynamics revealed a simple interpretation underlying neural variability and correlation suppression widely applicable to many types of neural data. During task states, the slope of the neural transfer function decreases, reducing the dynamic range of input-output responses. This results in reduced overall output variability, as well as reduced shared variability (e.g., correlations) from connected neural populations. The collective empirical and

theoretical results provide strong evidence that observed neural variability and correlations are state-dependent, and these changes emerge from the activity dynamics governed by the transfer functions of large neural masses.

The relationship between neural correlations and neural communication (or FC) is complex. For example, it appears that a decrease in the neural correlation between a pair of brain regions does not imply a reduction in communication. In the spiking literature, this interpretation is attributed to a reduction of shared spontaneous activity (or, neural noise). This is because the cross-trial mean evoked response (i.e., “the signal” associated with the task or stimulus) is removed prior to calculating the correlation, leaving only “neural noise” or spontaneous, moment-to-moment activity (Cohen and Kohn, 2011). Notably, this “neural noise” can still be important, since some portion of it drives trial-by-trial variability in cognition and behavior. In the fMRI literature, this is equivalent to regressing out the cross-trial mean task-evoked activity associated with the task/stimulus prior to calculating FC (Cole et al., 2019). The primary reason for this is to target neural-to-neural correlations, rather than task-to-neural correlations.

Our theoretical and empirical results clarify the interpretation of correlation changes from rest to task states in large-scale neural systems. Though empirical studies in large-scale functional networks using fMRI have reported FC increases during task states (Fiebach et al., 2006), we recently found that task-evoked activity inappropriately inflates FC estimates if the mean-evoked activity response is not properly removed (Cole et al., 2019). Indeed, when properly accounting for the mean-evoked response, we found that FC changes from rest to task states were dominated by FC decreases (see Figure 3I). The correction of the mean-evoked response in our paradigm brought the empirical results in line with our modeling results, suggesting a counterintuitive interpretation of FC changes during tasks: task co-activation in the presence of neural correlation quenching is consistent with task-related signal communication with background noise suppression. This can be understood from an information-theoretic perspective: during task communication, ongoing spontaneous activity will be suppressed (i.e., neural variability and correlations), increasing the fidelity of the task signal (mean evoked response). Importantly, our results were consistent using both correlation and covariance measures, suggesting that these decreases were due to reductions in shared variance rather than changes in unshared variance (Cole et al., 2016a; Duff et al., 2018; Siegel et al., 2012).

Though we largely focused on FC decreases during task states in both data sets, we identified a small number of correlations that increased during task state. Most of these correlation increases were primarily between regions belonging to different functional networks, which is consistent with previous literature (Cole et al., 2014; Gonzalez-Castillo and Bandettini, 2017; Krienen et al., 2014). Correlation increases have also been reported in the NHP spiking literature, where spike count correlations between units with more similar task tuning (e.g., similar receptive fields) tend to decrease, while spike count correlations between units with dissimilar task tuning increase (Ruff and Cohen, 2014). This appears to be conceptually consistent with the present fMRI findings, where regions in the same network tend to decrease their correlations, while regions across functionally distinct networks tend to increase their correlations. For example, our results reliably show correlation increases between the cingulo-opercular (a task positive network) and the default mode network (a task negative

network) during tasks (Figure 3F) (Power and Petersen, 2013). However, such findings are observational, and future work will be necessary to directly evaluate this hypothesis. Moreover, the simplicity of our current computational model is unable to account for these complex correlation changes. Our current model is limited to only two neural masses and therefore unable to account for potential differences in task tuning and receptive fields. Thus, it will be important for future work to investigate the theoretical basis (and functional relevance) of neural correlation increases in more complex networks.

We propose that a sigmoid transfer function is an effective model of the activation dynamics of large cortical areas. This was based on causally manipulating a locally balanced E-I circuit with clustered excitatory connectivity (Litwin-Kumar and Doiron, 2012). Moreover, the implementation of the sigmoid transfer function is consistent with prior computational studies demonstrating that resting-state activity (i.e., the absence of task-evoked activity) corresponds to dynamic regimes with large amplitude, slow fluctuations (Deco et al., 2013). In contrast, during task-evoked states, the output dynamics of the sigmoid transfer function are reduced, which correspond to evoked states (e.g., cortical “Up” states) that exhibit quenched variability (Harris and Thiele, 2011; Renart et al., 2010). The quenching of output variability can be explained by different biological mechanisms, such as clustered excitatory connectivity in local circuits, tightening of E-I balance due to inhibitory activation, neural adaptation, and/or irregular synaptic vesicle release (Doiron et al., 2016; Hennequin et al., 2018; Litwin-Kumar and Doiron, 2012; Rosenbaum et al., 2012). The manifestation of these biological mechanisms can be summarized at the mean-field by the reduction of the response variability due to the decreased slope in the sigmoid transfer function during highly active or inactive states.

The present results may appear to contradict some reports that task engagement increases (rather than decreases) overall neural communication. Yet there are several key differences between those previous findings and the present results. First, many of the previous results focused on communication through coherence, which often involves frequency-specific coupling of neural signals (Fries, 2005). This involves the phase-alignment of neural activity on faster timescales, which relates only indirectly to the slower correlation measures of spike rate activity and metabolic demand focused on here (Kahn et al., 2013; Pesaran et al., 2018). A second key difference between the present and most previous results is our emphasis on the absolute amount of correlation change from rest to task, rather than changes in network organization. Previous studies have also acknowledged that global signal regression, a common fMRI preprocessing step, shifts the baseline of rest- and task-state correlations and artificially induces negative correlations (Fox et al., 2009; Murphy et al., 2009). This preprocessing step confounds the comparison of the magnitude of correlation changes during independent rest- and task-state fMRI. In the present study, we rigorously preprocessed our fMRI data while ensuring to not remove the global signal. Along with the recent finding that incorrect removal of the mean-evoked response can inflate FC estimates, we suggest that rest to task correlation increases in previous fMRI studies should be interpreted with care.

Despite converging results, there are several key differences in our two empirical data sets. First, the time scale of fMRI BOLD activity is much slower than the NHP spiking activity. However, these differences were mitigated by measuring spiking variability across trials, which is comparable to the time scale of fMRI’s sampling rate (in the hundreds of milliseconds). In

addition, our computational model results demonstrated that reductions in neural variability and correlations were preserved after nonlinearly transforming the spike rate signal to the fMRI BOLD signal with the Balloon-Windkessel model (Friston et al., 2003). Another key difference in the data sets is the lack of a true resting-state data set in our NHP data. However, to better compare these two data sets, we demonstrated in the human fMRI data set that the task block periods showed reduced variability relative to inter-block intervals, which is a more analogous comparison to the NHP data set. Despite these efforts, it will be important for future work to more thoroughly investigate the differences in variability and correlation quenching using experimental paradigms that are consistent across the data acquisition techniques.

In conclusion, we propose a mechanistic framework for interpreting changes in neural variability and correlations by investigating the effects of task-state activity on the underlying neural attractor dynamics. Using theoretical modeling and empirical data analysis across two highly distinct neural data sets, we demonstrated convergent evidence suggesting that task states quench neural variability and correlations due to strengthening neural attractor dynamics across large-scale neural systems. Our work extends previous research establishing similar attractor mechanisms in sensory cortex (Hennequin et al., 2018) to characterize the role of attractor dynamics across large-scale cortical areas. We expect these findings to spur new investigations to better understand how we can interpret neural variability and correlations during task states, providing a deeper understanding of dynamic processes in the brain.

Methods

Spiking data: Data collection

The behavioral paradigm for each monkey was a motion-color categorization task (Figure 1B). Experimental methods for electrophysiology data collected for NHP was previously reported in (Siegel et al., 2015) and (Brincat et al., 2018). Data was collected *in vivo* from two (one female) behaving adult rhesus macaques (*Macaca mulatta*) across 55 sessions. Data from six distinct cortical regions were recorded simultaneously from acutely inserted electrodes. Cortical regions included: MT, V4, PIT, LIP, FEF, and LPFC (Figure 1A). Spikes from each region were sorted offline into isolated neurons. However, given our interest in inter-region neural correlations across large scale neural systems, we pooled spikes from each functional area into a single spike rate time series. For each trial, spikes were sorted for a 5s period, beginning 2.5s prior to stimulus onset, and until 3.5s after stimulus onset. Further details regarding electrophysiological data collection can be found here: http://www.sciencemag.org/content/348/6241/1352/suppl/DC1 and here: http://www.pnas.org/content/pnas/suppl/2018/07/09/1717075115.DCSupplemental/pnas.1717075115.sapp.pdf

All statistical analyses in the main article (detailed below) were performed on a single monkey. Independent replication was performed on the second monkey, and is reported in Supplementary Figure 1.

Spiking data: Task versus rest variability analysis

Neural variability analysis was analyzed using an analogous approach to both the computational model and fMRI data. However, since we had no true 'resting-state' activity for the monkey data set, we used the inter-trial interval (ITI; 0.5s - 1s variable duration, see Figure 1B) as "resting-state activity". We used the 0.5s - 1s interval immediately preceding the trial's fixation period to avoid any reward/feedback signals from the previous trial. (Reward/feedback from the previous trial was provided more than 1.5s prior to the fixation period.) Spike counts were calculated by taking a 50ms sliding window with 10ms increments, consistent with previous studies (Churchland et al., 2010). The mean-evoked response across all trials for a given task rule (e.g., motion rule versus color rule) was calculated and removed from each trial, as is common in the spike count literature (Cohen and Kohn, 2011) and the fMRI literature (Cole et al., 2019). (Statistically this is equivalent to performing the task activation regression in the fMRI data, described below.) Trials with less than 250ms (or 50 time points) worth of spiking data for either the ITI and/or task cue presentation were excluded. This was done to reduce variability of the estimated spike count correlations, since correlations with few observations are highly variable.

We computed the variance across 25 consecutive trials using the average spike rate from each cortical recording during either the ITI or the task cue period. This was repeated for all trials for each subject. We used across-trial variance to calculate variability rather than Fano factor (Churchland et al., 2010). This choice was due to the insight from our model illustrating that the mean-evoked activity and the corresponding variance interact in a nonlinear manner, and that the Fano factor is computed as the variance over the mean. Cross-trial variance was computed as

$$Var = \sum_{trial=x}^{x+n} \frac{(r_{trial} - \bar{r})^2}{n-1} \quad (1)$$

Where $n = 25$ trials, r_{trial} reflected the average spike rate of each trial, and $\bar{r}$ is the mean firing rate for the task condition across all trials (i.e., either the mean firing rate for the color or motion task cue period).

The statistical difference in task versus rest neural variability was computed by using a two-way, paired t-test across all bins of 25 consecutive trials. The global neural variability change was computed by averaging the variance across all recording areas for each bin. Statistics for the regional neural variability change were corrected for multiple comparisons using an FDR-corrected threshold of $p<0.05$.

In addition, we computed the variance during the ITI and task cue period within trial (across time points) (Supplementary Figure 2). This analysis demonstrated that moment-to-moment variability (rather than trial-to-trial variability) was also quenched from rest to task periods, suggesting that variability quenching also occurs at faster timescales. The statistical difference in task versus rest neural variability was computed by using a two-way, paired t-test (paired by trial) across all trials for each monkey separately. The global neural variability change was computed by averaging the variance across all recording areas for each

trial (Supplementary Figure 2A). Statistics for the regional neural variability change were corrected for multiple comparisons using an FDR-corrected threshold of $p<0.05$.

Spiking data: Task versus rest state correlation analysis

Neural correlations for spiking data using the same preprocessing steps mentioned above for spike rate variability analysis. Specifically, the mean-evoked response across all trials for each task condition was removed from each trial. Spike count correlations were then computed across trials, using groups of 25 trials as described above (Figure 2E-F).

The difference in task versus rest neural correlations was calculated using a two-way, paired t-test (paired by each bin of 25 trials) for each subject separately using Fisher's *z*-transformed correlation values. The global neural correlation change was computed by averaging the Fisher's *z*-transformed correlation values between all pairs of cortical regions, and comparing the averaged task versus rest correlation for each bin (Figure 2D). Statistics for the pairwise neural correlation change (Figure 2E-G) were corrected for multiple comparisons using an FDR-corrected threshold of $p<0.05$.

In addition, we computed the spike count correlation during the ITI and task cue period separately within trial (across time points) (Supplementary Figure 2). This analysis demonstrated that moment-to-moment correlations (rather than trial-to-trial correlations) were also quenched from rest to task periods, suggesting that correlation quenching also occurs at faster timescales. The statistical difference in task versus rest neural correlations was computed by using a two-way, paired t-test (paired by trial) across all trials for each monkey separately. The global neural correlation change was computed by averaging the correlation across all pairs of recording areas for each trial (Supplementary Figure 2B).

fMRI: Data and paradigm

The present study was approved by the Rutgers University institutional review board. Data were collected as part of the Washington University-Minnesota Consortium of the Human Connectome Project (HCP) (Van Essen et al., 2013). A subset of data (n = 352) from the HCP 1200 release was used for empirical analyses. Specific details and procedures of subject recruitment can be found in (Van Essen et al., 2013). The subset of 352 participants was selected based on: quality control assessments (i.e., any participants with any quality control flags were excluded, including 1) focal anatomical anomaly found in T1w and/or T2w scans, 2) focal segmentation or surface errors, as output from the HCP structural pipeline, 3) data collected during periods of known problems with the head coil, 4) data in which some of the FIX-ICA components were manually reclassified; low-motion participants (i.e., exclusion of participants that had any fMRI run in which more than 50% of TRs had greater than 0.25mm framewise displacement); removal according to family relations (unrelated participants were selected only, and those with no genotype testing were excluded). A full list of the 352 participants used in this study will be included as part of the code release.

All participants were recruited from Washington University in St. Louis and the surrounding area. We split the 352 subjects into two cohorts of 176 subjects: an exploratory cohort (99 females) and a replication cohort (84 females). The exploratory cohort had a mean age of 29 years of age (range=22-36 years of age), and the replication cohort had a mean age of 28 years of age (range=22-36 years of age). All subjects gave signed, informed consent in accordance with the protocol approved by the Washington University institutional review board. Whole-brain multiband echo-planar imaging acquisitions were collected on a 32-channel head coil on a modified 3T Siemens Skyra with TR=720 ms, TE=33.1 ms, flip angle=52º, Bandwidth=2,290 Hz/Px, in-plane FOV=208x180 mm, 72 slices, 2.0 mm isotropic voxels, with a multiband acceleration factor of 8. Data for each subject were collected over the span of two days. On the first day, anatomical scans were collected (including T1-weighted and T2-weighted images acquired at 0.7 mm isotropic voxels) followed by two resting-state fMRI scans (each lasting 14.4 minutes), and ending with a task fMRI component. The second day consisted with first collecting a diffusion imaging scan, followed by a second set of two resting-state fMRI scans (each lasting 14.4 minutes), and again ending with a task fMRI session.

Each of the seven tasks was collected over two consecutive fMRI runs. The seven tasks consisted of an emotion cognition task, a gambling reward task, a language task, a motor task, a relational reasoning task, a social cognition task, and a working memory task. Briefly, the emotion cognition task required making valence judgements on negative (fearful and angry) and neutral faces. The gambling reward task consisted of a card guessing game, where subjects were asked to guess the number on the card to win or lose money. The language processing task consisted of interleaving a language condition, which involved answering questions related to a story presented aurally, and a math condition, which involved basic arithmetic questions presented aurally. The motor task involved asking subjects to either tap their left/right fingers, squeeze their left/right toes, or move their tongue. The reasoning task involved asking subjects to determine whether two sets of objects differed from each other in the same dimension (e.g., shape or texture). The social cognition task was a theory of mind task, where objects (squares, circles, triangles) interacted with each other in a video clip, and subjects were subsequently asked whether the objects interacted in a social manner. Lastly, the working memory task was a variant of the N-back task.

Further details on the resting-state fMRI portion can be found in (Smith et al., 2013), and additional details on the task fMRI components can be found in (Barch et al., 2013). All fMRI results reported in the main article reflect results found with the first cohort of subjects. Independent replication of these effects are reported in Supplementary Figure 4 with the replication cohort.

fMRI: Preprocessing

Minimally preprocessed data for both resting-state and task fMRI were obtained from the publicly available HCP data. Minimally preprocessed surface data was then parcellated into 360 brain regions using the (Glasser et al., 2016) atlas. We performed additional standard preprocessing steps on the parcellated data for resting-state fMRI and task state fMRI to

conduct neural variability and FC analyses. This included removing the first five frames of each run, de-meaning and de-trending the time series, and performing nuisance regression on the minimally preprocessed data (Ciric et al., 2017). Nuisance regression removed motion parameters and physiological noise. Specifically, six primary motion parameters were removed, along with their derivatives, and the quadratics of all regressors (24 motion regressors in total). Physiological noise was modeled using aCompCor on time series extracted from the white matter and ventricles (Behzadi et al., 2007). For aCompCor, the first 5 principal components from the white matter and ventricles were extracted separately and included in the nuisance regression. In addition, we included the derivatives of each of those components, and the quadratics of all physiological noise regressors (40 physiological noise regressors in total). The nuisance regression model contained a total of 64 nuisance parameters. This was a variant of previously benchmarked nuisance regression models reported in (Ciric et al., 2017).

We excluded global signal regression (GSR), given that GSR artificially induces negative correlations (Murphy et al., 2009; Power et al., 2014), which would bias analyses of the difference of the magnitude of correlations between rest and task. We included aCompCor as a preprocessing step here given that aCompCor does not include the circularity of GSR (regressing out some global gray matter signal of interest) while including some of the benefits of GSR (some extracted components are highly similar to the global signal) (Power et al., 2018). This logic is similar to a recently-developed temporal-ICA-based artifact removal procedure that seeks to remove global artifact without removing global neural signals, which contains behaviorally relevant information such as vigilance (Glasser et al., 2017; Wong et al., 2013). We extended aCompCor to include the derivatives and quadratics of each of the component time series to further reduce artifacts. Code to perform this regression is publicly available online using python code (version 2.7.15) (https://github.com/ito-takuya/fmriNuisanceRegression).

Task data for task FC analyses were additionally preprocessed using a standard general linear model (GLM) for fMRI analysis. For each task paradigm, we removed the mean evoked task-related activity for each task condition by fitting the task timing (block design) for each condition using a finite impulse response (FIR) model (Cole et al., 2019). (There were 24 task conditions across seven cognitive tasks.) We used an FIR model instead of a canonical hemodynamic response function given recent evidence suggesting that the FIR model reduces both false positives and false negatives in the identification of FC estimates (Cole et al., 2019). This is due to the FIR model's ability to flexibly fit the mean-evoked response across all blocks. Removing the mean-evoked response of a task condition is critical to isolate second order effects such as correlations, and is standard when performing in spike count correlations (Cohen and Kohn, 2011; Cole et al., 2019). Analogous statistical preprocessing steps were critical when comparing neural correlation measures across human fMRI data and NHP spiking data.

FIR modeled task blocks were modeled separately for task conditions within each of the seven tasks. In particular, two conditions were fit for the emotion cognition task, where coefficients were fit to either the face condition or shape condition. For the gambling reward task, one condition was fit to trials with the punishment condition, and the other condition was fit to trials with the reward condition. For the language task, one condition was fit for the story condition, and the other condition was fit to the math condition. For the motor task, six

conditions were fit: (1) cue; (2) right hand trials; (3) left hand trials; (4) right foot trials; (5) left foot trials; (6) tongue trials. For the relational reasoning task, one condition was fit to trials when the sets of objects were matched, and the other condition was fit to trials when the objects were not matched. For the social cognition task, one condition was fit if the objects were interacting socially (theory of mind), and the other condition was fit to trials where objects were moving randomly. Lastly, for the working memory task, 8 conditions were fit: (1) 2-back body trials; (2) 2-back face trials; (3) 2-back tool trials; (4) 2-back place trials; (5) 0-back body trials; (6) 0-back face trials; (7) 0-back tool trials; (8) 0-back place trials. Since all tasks were block designs, each time point for each block was modeled separately for each task condition (i.e., FIR model), with a lag extending up to 25 TRs after task block offset.

fMRI: Task state activation analysis

We performed a task GLM analysis on fMRI task data to evaluate the evoked global brain activity during task states. The task timing for each of the 24 task conditions was convolved with the SPM canonical hemodynamic response function to obtain task-evoked activity estimates (Friston et al., 1994). FIR modeling was not used when modeling task-evoked activity. Coefficients were obtained for each parcel in the Glasser et al. (2016) cortical atlas for each of the 24 task conditions.

fMRI: Task state versus resting-state variability analysis

To compare task state versus resting-state variability, we regressed out the exact same FIR design matrix used on task-state regression on resting-state data. This was possible given that the number of timepoints of the combined resting-state scans in the HCP data set exceeded the number of timepoints of the combined task-state scans (4800 resting-state TRs > 3880 task-state TRs). This step was to ensure that any spurious change to the residual task time series through the removal of the mean-evoked response through FIR regression would also induce spurious changes in the resting-state data. However, results were qualitatively identical without the regression of the task design matrix on resting-state data.

After task regression, we obtained the residual time series for both resting-state and task state fMRI data. We then z-normalized each task run with zero-mean and unit variance such that we could appropriately compare the neural variability of task blocks across different runs. We emphasize that task activation regression (removal of the mean-evoked task response) was removed *prior* to z-scoring the time series. Additionally, Supplementary Figure 5 shows results without *z*-normalization, and the results are qualitatively identical.) This enabled us to evaluate whether the variability during task blocks significantly decreased relative to inter-block intervals by evaluating the variance of task blocks relative to 1. We then extracted the time series variance during task blocks, and then averaged the variance across all task conditions to obtain our statistic of task-evoked neural variability. To identify the resting-state neural variability, we applied the same exact procedure to resting-state time series using the task-state design matrix. A sanity check for our analysis was that the ‘intrinsic-state’ neural variability is close to 1 (given that the time series was normalized to have unit variance), while the task-state neural variability

is significantly less than 1 (Figure 3A). This ensured that variability measures were not biased by the normalization step.

We compared the neural variability of the entire brain during task state periods versus resting-state periods. For each subject, we computed the variance during task and resting state separately, and then averaged across all brain regions. This resulted in two values per subject, representing task state and resting-state variability. We then performed a two-way group paired t-test across subjects to assess statistical significance (Figure 3A). We also computed the task state versus resting-state difference in neural variability for each brain region separately (Figure 3B). We corrected for multiple comparisons using an FDR-corrected threshold of $p<0.05$ (Figure 3B,C). Cortical surface visualizations were constructed using Connectome Workbench (version 1.2.3) (Van Essen et al., 2013).

fMRI: Task state versus resting-state correlation analysis

We compared task state versus resting-state FC (i.e., neural correlations), after performing the exact same preprocessing steps as mentioned above. Results without *z*-normalization (and using covariance rather than correlations) on the task and rest residual time series are reported in Supplementary Figure 5.

We computed the correlation between all pairs of brain regions for each task condition during task block periods. We then averaged the Fisher's *z*-transformed correlation values across all task conditions to obtain a general task state FC matrix (Figure 3E). We repeated the same procedure (i.e., using the same task-timed blocks) on resting-state FC to obtain an equivalent resting-state FC matrix for each subject (Figure 3D). We directly compared task-state FC to resting-state FC by performing two-way group paired t-tests for every pair of brain regions using the Fisher's *z*-transformed correlation values. Statistical significance was assessed using an FDR-corrected threshold of $p<0.05$ (Figure 3F). To compare the average global correlation during task state and resting-state, we computed the average correlation between all pairs of brain regions during task and resting-state, performing a group paired t-test (Figure 3G). To compare the average global connectivity profile of every brain region (Cole et al., 2010), we computed the average Fisher *z*-transformed correlation of a single region to all other brain regions during task and rest and performed a two-way group paired t-test between task and rest (Figure 3H,I). Statistical significance was assessed using an FDR-corrected threshold of $p<0.05$.

Information-theoretic analysis

We evaluated the information-theoretic relevance of rest and task states by characterizing the dimensionality of neural activity. To estimate the statistical dimensionality of neural data, we used the formulation of dimensionality previously described in (Litwin-Kumar et al., 2017). We first obtain the covariance matrix W of activity for rest and task states separately. We then calculated

$$dim_W = \frac{(\sum_{i=1}^{m} \lambda_i)^2}{\sum_{i=1}^{m} \lambda_i^2} \qquad (2)$$

Where dim_W corresponds to the statistical dimensionality of W, and λ_i corresponds to the eigenvalues of the covariance matrix W. Intuitively, this is related to finding the number of components needed to explain variance greater than some fixed threshold, with more needed components reflecting a higher dimensionality of the data.

For human fMRI data, we estimated the task-state and resting-state dimensionality by calculating the whole-brain covariance matrix for each state. For task state this was done by estimating the covariance matrix using task block periods. For resting state this was done by calculating the covariance matrix across the equivalently lengthed resting-state periods (using the same data in the FC analysis above). We applied equation 2 to the task-state and resting-state covariance matrix for each subject. Finally, we applied a two-way, group paired t-test comparing the dimensionality of task-state activity to resting-state activity (Figure 4A). We replicated this finding in the replication cohort. In addition, we performed this analysis for each fMRI task separately (Supplementary Figure 10).

For NHP spiking data, we estimated the task (task cue period) and rest (ITI) dimensionality by calculating the covariance matrix between all pairs of population recordings. We then applied equation 2 to task and rest periods for each covariance matrix. (Each covariance matrix was calculated using bins of 25 consecutive trials.) Finally, we applied a two-way group paired t-test (across bins) comparing the dimensionality of task activity to rest activity (Figure 5B). We replicated this effect in the held-out second monkey.

Spiking model: Estimating the transfer function of a neural population with a balanced spiking model

Our goal was to evaluate the effects of evoked activity across large neural populations, rather than within populations. Thus, we first estimated the transfer function of a neural population using a previously established balanced neural spiking model, with 4000 excitatory and 1000 inhibitory units (Litwin-Kumar and Doiron, 2012). All parameters are taken directly from (Litwin-Kumar and Doiron, 2012) with the description paraphrased below. Units within the network were modeled as leaky integrate-and-fire neurons whose membrane voltages obeyed the equation

$$\frac{dV}{dt} = \frac{1}{\tau}(\mu - V) + I_{syn} \qquad (3)$$

where τ indicates the membrane time constant, μ is the bias term, and I_{syn} is the synaptic input. When neurons reached $V_{th} = 1$, a spike was emitted, and voltages were reset to $V_{re} = 0$ for an absolute refractory period of 5ms. τ was 15ms and 10ms for excitatory and inhibitory neurons, respectively. For excitatory neurons, μ was randomly sampled from a uniform distribution between 1.1 and 1.2. For inhibitory neurons, μ was randomly sampled from a uniform distribution between 1 and 1.05.

Synapses to a neuron were modeled as the sum of excitatory and inhibitory synaptic trains x_E and x_I, respectively, and was calculated as the normalized difference of exponentials describing the synaptic rise and decay times caused by each presynaptic event. This effectively

captured the weighted effect of all presynaptic neurons to a target neuron, and specifically obeyed the equations

$$I_{y,syn} = x_E(t) + x_I(t) \quad (4)$$

$$x_Z(t) = \frac{x_{Z,\,decay} - x_{Z,\,rise}}{\tau_{Z,\,decay} - \tau_{Z,\,rise}}, \qquad Z \in \{E, I\} \quad (5)$$

where the synaptic rise and decay of x_E and x_I was modeled as the first order differential equations

$$\frac{dx_{Z,\,decay}}{dt} = \sum_j J_{ij} s_j - \frac{x_{Z,\,decay}}{\tau_{Z,\,decay}} \quad (6)$$

$$\frac{dx_{Z,\,rise}}{dt} = \sum_j J_{ij} s_j - \frac{x_{Z,\,rise}}{\tau_{Z,\,rise}} \quad (7)$$

J_{ij} refers to the synaptic weight from neuron j to i, s_j indicates whether neuron j emitted a spike. Synaptic rise times were the same for excitatory and inhibitory neurons, with $\tau_{E,\,rise} = \tau_{I,\,rise} = 1ms$, while $\tau_{E,\,decay} = 3ms$ and $\tau_{I,\,decay} = 2ms$. Connection probabilities p^{xy} from neurons in population y to x were $p^{EI} = p^{IE} = p^{II} = 0.5$, and on average, $p^{EE} = 0.2$. However, if two neurons were both excitatory and belonged to the same cluster, the connection strength was multiplied by 1.9. (We employed only the homogenous clustered networks, as described by (Litwin-Kumar and Doiron, 2012) $J^{EE} = 0.024$, $J^{EI} = -0.045$, $J^{IE} = 0.014$, and $J^{II} = -0.057$. Excitatory stimulation was performed by increasing μ to the first 400 excitatory neurons from 0.05 to 0.5 in 0.05 increments. Inhibitory stimulation was performed by decreasing μ by 0.05 to 0.5 in 0.05 increments to 400 inhibitory neurons.

To estimate the population transfer function, we simulated 30 trials lasting 2s each at each stimulation amplitude. Spike train statistics were estimated across trials in 50ms sliding windows with 10ms shifts. Only excitatory neurons were included when calculating the population spike train statistics (i.e., mean and variance at each stimulation amplitude).

Model code was originally adapted from (Litwin-Kumar and Doiron, 2012), and was simulated with Julia (version 1.1.1).

Model: One-dimensional minimal network model

We use the simplest model to mathematically characterize the relationship between evoked activity and neural variability: a one-dimensional mean-field model. We used Wilson-Cowan-type firing rate dynamics to simulate neural population activity (Wilson and Cowan, 1972). Specifically, our population's activity obeyed the equation

$$\tau_i \frac{dx_i}{dt} = -x_i + f(w_{ii} x_i + b_i + s_i + I) \quad (8)$$

where x_i denotes the firing rate (or a measure of activity), $\tau_i = 0.1$ denotes the time constant, $w_{ii} = 1$ refers to local coupling (auto-correlation), $b_i = -0.5$ refers to the input threshold for optimal activity (or a bias term), s_i refers to the evoked stimulation ($s_i = 0$ for intrinsic activity), I refers to background spontaneous activity sampled from a Gaussian distribution with mean 0

and standard deviation 0.25, and f is a sigmoid input-output activation function, which is defined as

$$f(x) = \frac{1}{1+e^{-k*x}} \qquad (9)$$

where $k = 1$. Numerical simulations were computed using a Runge-Kutta second order method with a time step of dt=10ms (Burden and Faires, 2001). We simulated neural population activity injecting a fixed input (boxcar input) with amplitudes ranging from -5 to 5 in 0.01 increments (Figure 6C). Neural variability for each input strength was calculated using the standard deviation of the time series following the input onset and preceding input offset. Each trial was run for 20 seconds. Figure 6C was generated using input amplitudes of -3, 0, and 3.

To visualize the full dynamics of our single neural population, we visualized the one-dimensional phase space (i.e., flow field on a line) (Strogatz, 1994). In particular, we calculated the flow field by plotting $\dot{x}$ (i.e., $\frac{dx}{dt}$) as a function of x_i (equation 8). Notably, fixed point attractors (equilibrium states) are defined where $\dot{x} = 0$ (Fig. 6B).

Model: Two-dimensional minimal network model

To characterize the effects of evoked activity on neural correlations, we use a two-dimensional neural population model. We extended the one-dimensional network model to include two neural populations. The network dynamics obeyed the equations

$$\tau_1 \frac{dx_1}{dt} = -x_1 + f(w_{11}x_1 + w_{21}x_2 + b_1 + s_1 + I_1) \quad (10)$$

$$\tau_2 \frac{dx_2}{dt} = -x_2 + f(w_{22}x_2 + w_{12}x_1 + b_2 + s_2 + I_2) \quad (11)$$

where x_1 and x_2 describe the activity of each population, and all other variables are as described above. Inter-regional coupling was set to be greater than local coupling, given evidence from previous studies that global coupling is greater than local coupling (Cole et al., 2016b; Deco et al., 2013; Ito et al., 2017). Specific network parameters for this network model were: $w_{11} = w_{22} = 2$, $w_{12} = w_{21} = 4$, $b_1 = b_2 = -3$, $\tau_1 = \tau_2 = 0.1$. I_1 and I_2 were sampled from a Gaussian distribution with mean 0 and standard deviation 1. For this network model, we decreased the slope of the sigmoid $k = 0.5$ to allow for a larger dynamic, linear response range. The full model code will be made publicly available.

To quantify the relationship between evoked activity and neural correlations, we systematically simulated the network under different stimulation states (input strengths). Using the same methods as above, we simulated network activity for 50 seconds. We injected fixed input into both neural populations with amplitudes from -5 to 5 in 0.01 increments (Fig. 7E). Notably, given that the injected stimulation is uncorrelated (due to 0-variance in a fixed input), it is non-trivial that the FC between two nodes would change in response to different inputs. Neural correlations were calculated using a Pearson correlation of the two time series following input onset and preceding input offset.

The use of a minimal model constrained our network to two dimensions. This allowed us to leverage dynamical systems tools to visualize the flow field in the two-dimensional phase plane. To identify the fixed point attractors, we first calculated the nullclines for x_1 and x_2. Nullclines are defined as the values of x_1 and x_2 such that $\dot{x_2} = 0$ and $\dot{x_1} = 0$, respectively.

The fixed points lie at the intersection of the nullclines. For our particular system of equations, the nullclines of x_1 and x_2 were defined, respectively, as

$$x_2 = \frac{f^{-1}(-x_1) - w_{11}x_1 - b_1 - s_1}{w_{21}} \qquad (12)$$

$$x_1 = \frac{f^{-1}(-x_2) - w_{22}x_2 - b_2 - s_2}{w_{12}} \qquad (13)$$

where parameters are identical to those used in equations 10 and 11. However, the background noise, parameter *I*, was removed when calculating the nullclines. Fixed point attractors (equilibrium states) are defined where $\dot{x} = 0$ (Figure 7B). The full flow field was obtained by applying the system of equations (equations 10 and 11) to every point in the phase space (e.g., all values of x_1 and x_2).

Model: Evaluating fixed point attractor dynamics and the characteristic time scale

Our models accurately demonstrated that evoked activity decreased the neural variability and correlations from a stochastic dynamical network model. Since our network model was governed by firing rate equations which provided us full access to the system's dynamics, we sought to link dynamical mechanisms (in the absence of spontaneous activity) with changes in the descriptive statistics. Such an analysis would provide us with a mechanistic understanding between descriptive neural statistics used in empirical data analysis and the governing neural dynamics.

To understand how attractor dynamics influenced simulated activity in a network model, we characterized the dynamics around the network's fixed point attractor. Specifically, we performed a linear stability analysis around the fixed point (i.e., the equilibrium level of activity the system is drawn to during a particular state or input, e.g., Figure 6B) in both the one-dimensional and two-dimensional network models. In the one-dimensional case, this analysis is equivalent to evaluating the first derivative of equation 8 at the fixed point (e.g., the slope of the line at the starred locations in Figure 6B). We then calculate the *characteristic time scale* T at the fixed point x^* (in one-dimension) with the equation

$$T = \frac{1}{|f'(x^*)|} \qquad (14)$$

where f represents equation 8 (Strogatz, 1994). The characteristic time scale captures the speed with which the system approaches the fixed point attractor. We calculated the characteristic time scale across the same range of evoked stimulation strengths as in the neural variability analysis. Fixed points were computed numerically by running the network model until it reached a steady state in the absence of noise/spontaneous activity.

The characteristic time scale is an established measure for one-dimensional systems. However, we sought to extend the characteristic time scale beyond a single dimension to evaluate shifting attractor dynamics in higher dimensions. We first performed a linear stability analysis in two dimensions by evaluating the Jacobian matrix for our two-dimensional system at the fixed point (x^*_1, x^*_2)

$$\mathbf{J}(x_1^*, x_2^*) = \begin{pmatrix} \frac{df_1}{dx_1} & \frac{df_1}{dx_2} \\ \frac{df_2}{dx_1} & \frac{df_2}{dx_2} \end{pmatrix} \qquad (15)$$

Where f_1 and f_2 refer to the equations governing neural populations 1 and 2 (equations 10 and 11, respectively). For our particular system of equations, the Jacobian was calculated as

$$\mathbf{J}(x_1^*, x_2^*) = \begin{pmatrix} \left(-1 + f'(w_{11}x_1 + w_{21}x_2 + b_1 + s_1)\right)\frac{1}{\tau_1} & \left(f'(w_{11}x_1 + w_{21}x_2 + b_1 + s_1)\right)\frac{1}{\tau_1} \\ \left(f'(w_{22}x_2 + w_{12}x_1 + b_2 + s_2)\right)\frac{1}{\tau_2} & \left(-1 + f'(w_{22}x_2 + w_{12}x_1 + b_2 + s_2)\right)\frac{1}{\tau_2} \end{pmatrix} \quad (16)$$

For each input strength (i.e., differing evoked states), we evaluated the Jacobian at the fixed point attractor. We then calculated the two eigenvalues (denoted λ_1 and λ_2) and eigenvectors (denoted v_1 and v_2) of the Jacobian using an eigendecomposition. To calculate the generalized characteristic time scale in two dimensions, we first calculated the linear combination of the eigenvectors weighted by the real eigenvalues, and computed the magnitude of the vector, such that

$$v_{sum}(x, y) = re(\lambda_1)v_1 + re(\lambda_2)v_2 \quad (17)$$

We then define the two dimensional characteristic time scale T as the reciprocal of the magnitude of $v_{sum}(x, y)$, such that

$$T = \frac{1}{|\sqrt{x^2+y^2}|} \quad (18)$$

We calculated T for a range of values $s_1, s_2 \in [-5, 5]$ in 0.01 increments, and correlated T across all values of s_1 and s_2 with the corresponding neural correlations (Strogatz, 1994).

Model: Simulating fMRI BOLD activity

We used the above model to simulate fMRI BOLD activity to demonstrate that changes in neural variability and correlations would extend to fMRI BOLD dynamics (Supplementary Figures 6-7). Neural activity generated from our model simulations was transformed to fMRI BOLD activity using the Balloon-Windkessel model, a nonlinear transformation from neural activity to the fMRI BOLD signal (Buxton et al., 1998; Friston et al., 2003). Notably, the transformation assumes a nonlinear transformation of the normalized deoxyhemoglobin content, normalized blood inflow, resting oxygen extraction fraction, and the normalized blood volume. All state equations and biophysical parameters were taken directly from (Friston et al., 2003) (equations 4-5). The Balloon-Windkessel model was implemented in Python (version 2.7.13), and the implementation code has been made publicly available on GitHub (https://github.com/ito-takuya/HemodynamicResponseModeling).

Code and data availability

All code related to analyses and models in this study will be publicly released on GitHub. Code will provide supplementary details on data analysis, computational modeling, and figure generation. All fMRI human data is publicly available through the Human Connectome Project (http://www.humanconnectomeproject.org) (Van Essen et al., 2013). Inquiries related to monkey electrophysiology data can be made to authors S.L.B. and E.K.M.

References

Abbott, L.F., Rajan, K., and Sompolinsky, H. (2011). Interactions between Intrinsic and Stimulus-Evoked Activity in Recurrent Neural Networks. The Dynamic Brain: An Exploration of Neuronal Variability and Its Functional Significance 1–16.

Averbeck, B.B., Latham, P.E., and Pouget, A. (2006). Neural correlations, population coding and computation. Nat. Rev. Neurosci. *7*, 358.

Barch, D.M., Burgess, G.C., Harms, M.P., Petersen, S.E., Schlaggar, B.L., Corbetta, M., Glasser, M.F., Curtiss, S., Dixit, S., Feldt, C., et al. (2013). Function in the human connectome: task-fMRI and individual differences in behavior. Neuroimage *80*, 169–189.

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

Biswal, B., Yetkin, F.Z., Haughton, V.M., and Hyde, J.S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn. Reson. Med. *34*, 537–541.

Brincat, S.L., Siegel, M., von Nicolai, C., and Miller, E.K. (2018). Gradual progression from sensory to task-related processing in cerebral cortex. Proceedings of the National Academy of Sciences *115*, E7202 LP – E7211.

Burden, R.L., and Faires, J.D. (2001). Numerical analysis. 2001. Brooks/Cole, USA.

Buxton, R.B., Wong, E.C., and Frank, L.R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. Magn. Reson. Med. *39*, 855–864.

Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B.B., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat. Neurosci. *13*, 369.

Ciric, R., Wolf, D.H., Power, J.D., Roalf, D.R., Baum, G.L., Ruparel, K., Shinohara, R.T., Elliott, M.A., Eickhoff, S.B., Davatzikos, C., et al. (2017). Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. Neuroimage *154*, 174–187.

Cohen, M.R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. Nat. Neurosci. *14*, 811.

Cohen, M.R., and Maunsell, J.H.R. (2009). Attention improves performance primarily by reducing interneuronal correlations. Nat. Neurosci. *12*, 1594–1600.

Cole, M., Bassett, D., Power, J., Braver, T., and Petersen, S. (2014). Intrinsic and task-evoked network architectures of the human brain. *83*.

Cole, M.W., Yang, G.J., Murray, J.D., Repovš, G., and Anticevic, A. (2016a). Functional connectivity change as shared signal dynamics. J. Neurosci. Methods *259*, 22–39.

Cole, M.W., Ito, T., Bassett, D.S., and Schultz, D.H. (2016b). Activity flow over resting-state networks shapes cognitive task activations. Nat. Neurosci. *19*, 1718–1726.

Cole, M.W., Ito, T., Schultz, D., Mill, R., Chen, R., and Cocuzza, C. (2019). Task activations produce spurious but systematic inflation of task functional connectivity estimates. Neuroimage *189*, 1–18.

Deco, G., and Hugues, E. (2012). Neural Network Mechanisms Underlying Stimulus Driven Variability Reduction. PLoS Comput. Biol. *8*, e1002395.

Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G.L., Hagmann, P., and Corbetta, M. (2013). Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. J. Neurosci. *33*, 11239–11252.

Deco, G., Ponce-Alvarez, A., Hagmann, P., Romani, G.L., Mantini, D., and Corbetta, M. (2014). How local excitation-inhibition ratio impacts the whole brain dynamics. J. Neurosci. *34*, 7886–7898.

Doiron, B., Litwin-Kumar, A., Rosenbaum, R., Ocker, G.K., and Josić, K. (2016). The mechanics of state-dependent neural correlations. Nat. Neurosci. *19*, 383.

Duff, E.P., Makin, T., Cottaar, M., Smith, S.M., and Woolrich, M.W. (2018). Disambiguating brain functional connectivity. Neuroimage *173*, 540–550.

Ecker, A.S., Berens, P., Keliris, G.A., Bethge, M., Logothetis, N.K., and Tolias, A.S. (2010). Decorrelated neuronal firing in cortical microcircuits. Science *327*, 584–587.

Fiebach, C.J., Rissman, J., and D'Esposito, M. (2006). Modulation of inferotemporal cortex activation during verbal working memory maintenance. Neuron *51*, 251–261.

Fox, M.D., Zhang, D., Snyder, A.Z., and Raichle, M.E. (2009). The Global Signal and Observed Anticorrelated Resting State Brain Networks. J. Neurophysiol. *101*, 3270–3283.

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. Trends Cogn. Sci. *9*, 474–480.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.-P., Frith, C.D., and Frackowiak, R.S.J. (1994). Statistical parametric maps in functional imaging: A general linear approach. Hum. Brain Mapp. *2*, 189–210.

Friston, K.J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. Neuroimage *19*, 1273–1302.

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., et al. (2016). A multi-modal parcellation of human cerebral cortex. Nature *536*, 171–178.

Glasser, M.F., Coalson, T.S., Bijsterbosch, J.D., Harrison, S.J., Harms, M.P., Anticevic, A., Van Essen, D.C., and Smith, S.M. (2017). Using Temporal ICA to Selectively Remove Global Noise While Preserving Global Signal in Functional MRI Data.

Gonzalez-Castillo, J., and Bandettini, P.A. (2017). Task-based dynamic functional connectivity: Recent findings and open questions. Neuroimage 1–8.

Harris, K.D., and Thiele, A. (2011). Cortical state and attention. Nat. Rev. Neurosci. *12*, 509.

He, B.J. (2011). Scale-Free Properties of the Functional Magnetic Resonance Imaging Signal during Rest and Task. J. Neurosci. *31*, 13786 LP – 13795.

He, B.J. (2013). Spontaneous and task-evoked brain activity negatively interact. J. Neurosci. *33*, 4672–4682.

Hennequin, G., Ahmadian, Y., Rubin, D.B., Lengyel, M., and Miller, K.D. (2018). The Dynamical Regime of Sensory Cortex: Stable Dynamics around a Single Stimulus-Tuned Attractor Account for Patterns of Noise Variability. Neuron *98*, 846–860.e5.

Ito, T., Kulkarni, K.R., Schultz, D.H., Mill, R.D., Chen, R.H., Solomyak, L.I., and Cole, M.W. (2017). Cognitive task information is transferred between brain regions via resting-state network topology. Nat. Commun. *8*, 1027.

Jacobs, E.A.K., Steinmetz, N.A., Carandini, M., and Harris, K.D. (2018). Cortical state fluctuations during sensory decision making.

Ji, J.L., Spronk, M., Kulkarni, K., Repovš, G., Anticevic, A., and Cole, M.W. (2018). Mapping the human brain's cortical-subcortical functional network organization. Neuroimage.

Joglekar, M.R., Mejias, J.F., Yang, G.R., and Wang, X.-J. (2018). Inter-areal Balanced Amplification Enhances Signal Propagation in a Large-Scale Circuit Model of the Primate Cortex. Neuron *98*, 222–234.e8.

Kahn, I., Knoblich, U., Desai, M., Bernstein, J., Graybiel, A.M., Boyden, E.S., Buckner, R.L., and Moore, C.I. (2013). Optogenetic drive of neocortical pyramidal neurons generates fMRI signals that are correlated with spiking activity. Brain Res. *1511*, 33–45.

Krienen, F.M., Yeo, B.T.T., and Buckner, R.L. (2014). Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. Philos. Trans. R. Soc. Lond. B Biol. Sci. *369*, 20130526–20130526.

Litwin-Kumar, A., and Doiron, B. (2012). Slow dynamics and high variability in balanced cortical networks with clustered connections. Nat. Neurosci. *15*, 1498–1505.

Litwin-Kumar, A., Harris, K.D., Axel, R., Sompolinsky, H., and Abbott, L.F. (2017). Optimal Degrees of Synaptic Connectivity. Neuron *0*, 1153–1164.e7.

MacKay, D.J.C. (2003). Information theory, inference and learning algorithms.

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

Pesaran, B., Vinck, M., Einevoll, G.T., Sirota, A., Fries, P., Siegel, M., Truccolo, W., Schroeder, C.E., and Srinivasan, R. (2018). Investigating large-scale brain dynamics using field potential

recordings: analysis and interpretation. Nat. Neurosci. *21*, 903–919.

Pinto, L., Rajan, K., DePasquale, B., Thiberge, S.Y., Tank, D.W., and Brody, C.D. (2019). Task-Dependent Changes in the Large-Scale Dynamics and Necessity of Cortical Regions. Neuron.

Ponce-alvarez, A., He, B.J., Hagmann, P., and Deco, G. (2015). Task-Driven Activity Reduces the Cortical Activity Space of the Brain : Experiment and Whole-Brain Modeling. 1–26.

Power, J.D., and Petersen, S.E. (2013). Control-related systems in the human brain. Curr. Opin. Neurobiol. *23*, 223–228.

Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J. a., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., et al. (2011). Functional Network Organization of the Human Brain. Neuron *72*, 665–678.

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

Power, J.D., Plitt, M., Gotts, S.J., Kundu, P., Voon, V., Bandettini, P.A., and Martin, A. (2018). Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data. Proceedings of the National Academy of Sciences 201720985.

Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., and Harris, K.D. (2010). The Asynchronous State in Cortical Circuits. Science *327*, 587 LP – 590.

Rosenbaum, R., Rubin, J.E., and Doiron, B. (2012). Short-term synaptic depression and stochastic vesicle dynamics reduce and shape neuronal correlations. J. Neurophysiol. *109*, 475–484.

Ruff, D.A., and Cohen, M.R. (2014). Attention can either increase or decrease spike count correlations in visual cortex. Nat. Neurosci. *17*, 1591–1597.

Ruff, D.A., and Cohen, M.R. (2016). Stimulus Dependence of Correlated Variability across Cortical Areas. J. Neurosci. *36*, 7546–7556.

Siegel, M., Donner, T.H., and Engel, A.K. (2012). Spectral fingerprints of large-scale neuronal interactions. Nat. Rev. Neurosci.

Siegel, M., Buschman, T.J., and Miller, E.K. (2015). Cortical information flow during flexible sensorimotor decisions. Science *348*, 1352–1355.

Smith, S.M., Beckmann, C.F., Andersson, J., Auerbach, E.J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D.A., Griffanti, L., Harms, M.P., et al. (2013). Resting-state fMRI in the Human Connectome Project. Neuroimage *80*, 144–168.

Song, S., Sjöström, P.J., Reigl, M., Nelson, S., and Chklovskii, D.B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS Biol. *3*, e68.

Strogatz, S.H. (1994). Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering (Cambridge, MA.

Tomasi, D., Wang, R., Wang, G.-J., and Volkow, N.D. (2014). Functional Connectivity and Brain Activation: A Synergistic Approach. Cereb. Cortex *24*, 2619–2629.

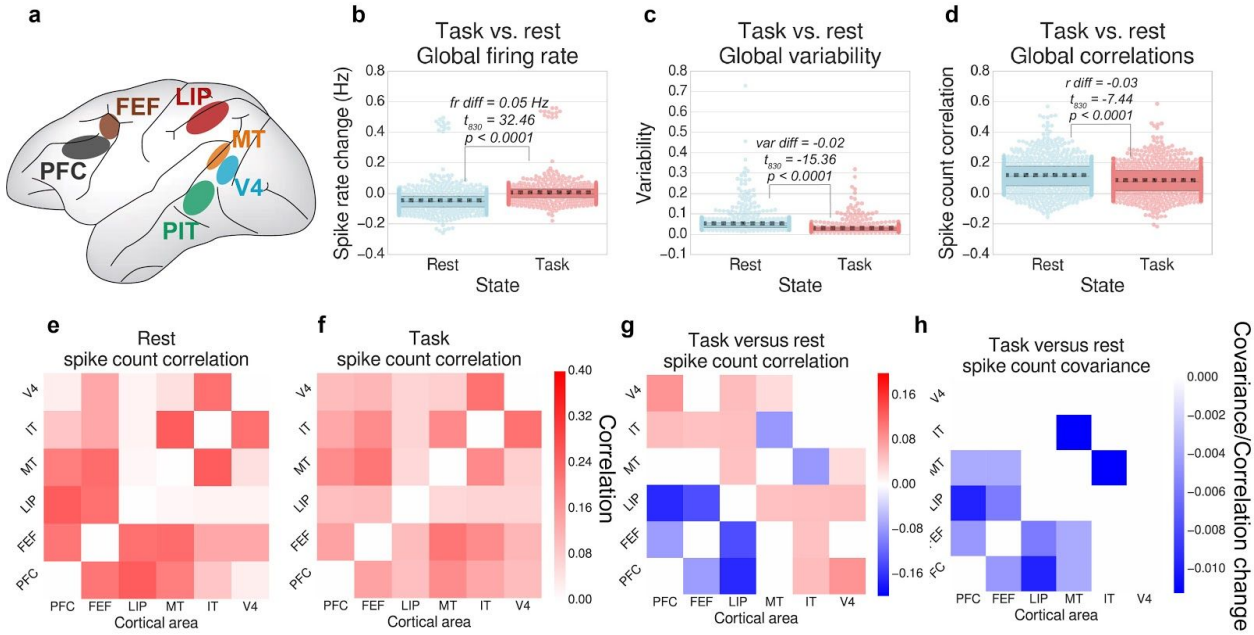
Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., and Ugurbil, K. (2013). The WU-Minn Human Connectome Project: an overview. Neuroimage *80*, 62–79.

Wilson, H.R., and Cowan, J.D. (1972). Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. Biophys. J. *12*, 1–24.

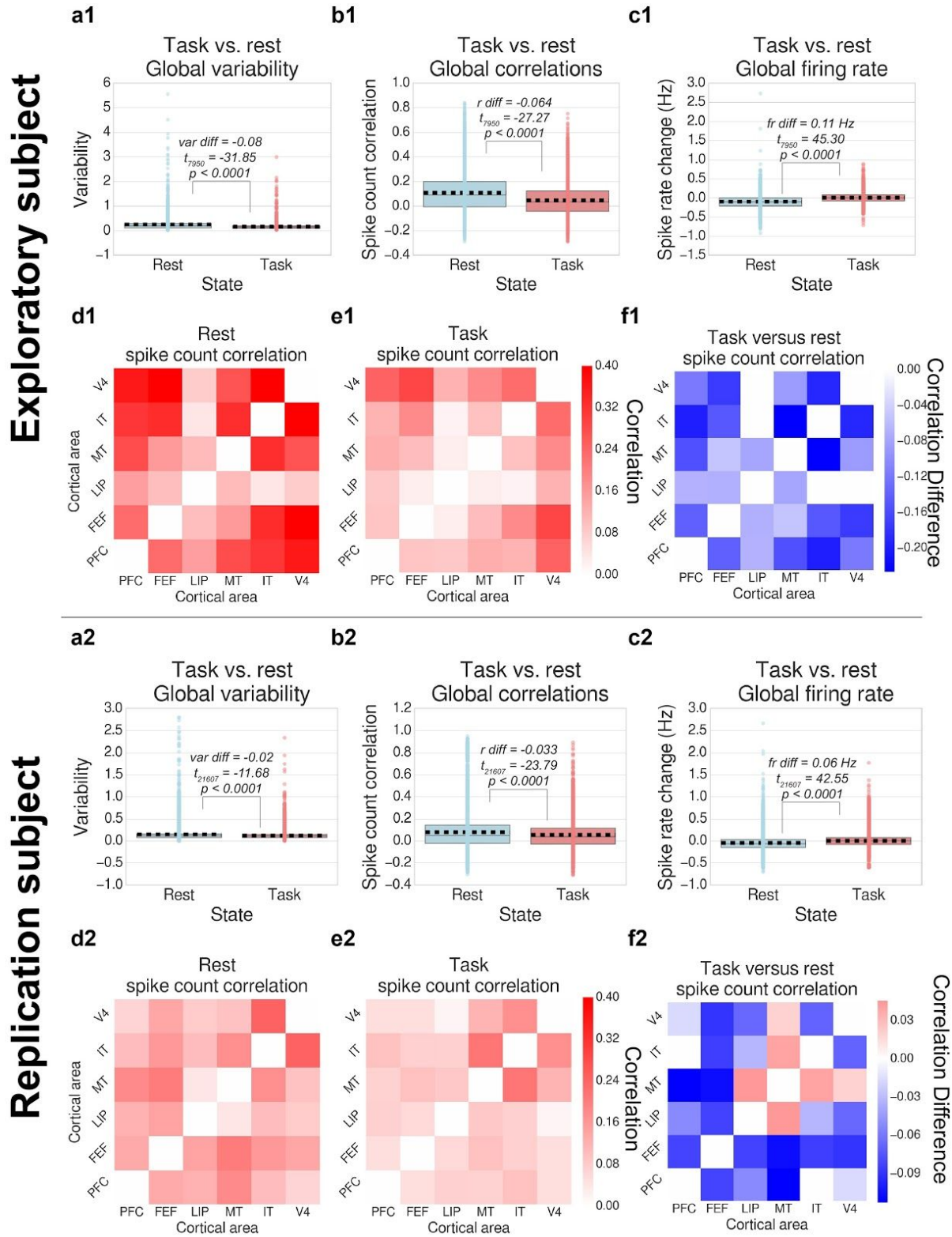
Wong, C.W., Olafsson, V., Tal, O., and Liu, T.T. (2013). The amplitude of the resting-state fMRI global signal is related to EEG vigilance measures. Neuroimage *83*, 983–990.

Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zollei, L., Polimeni, J.R., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. *106*, 1125–1165.

Supplementary Figures

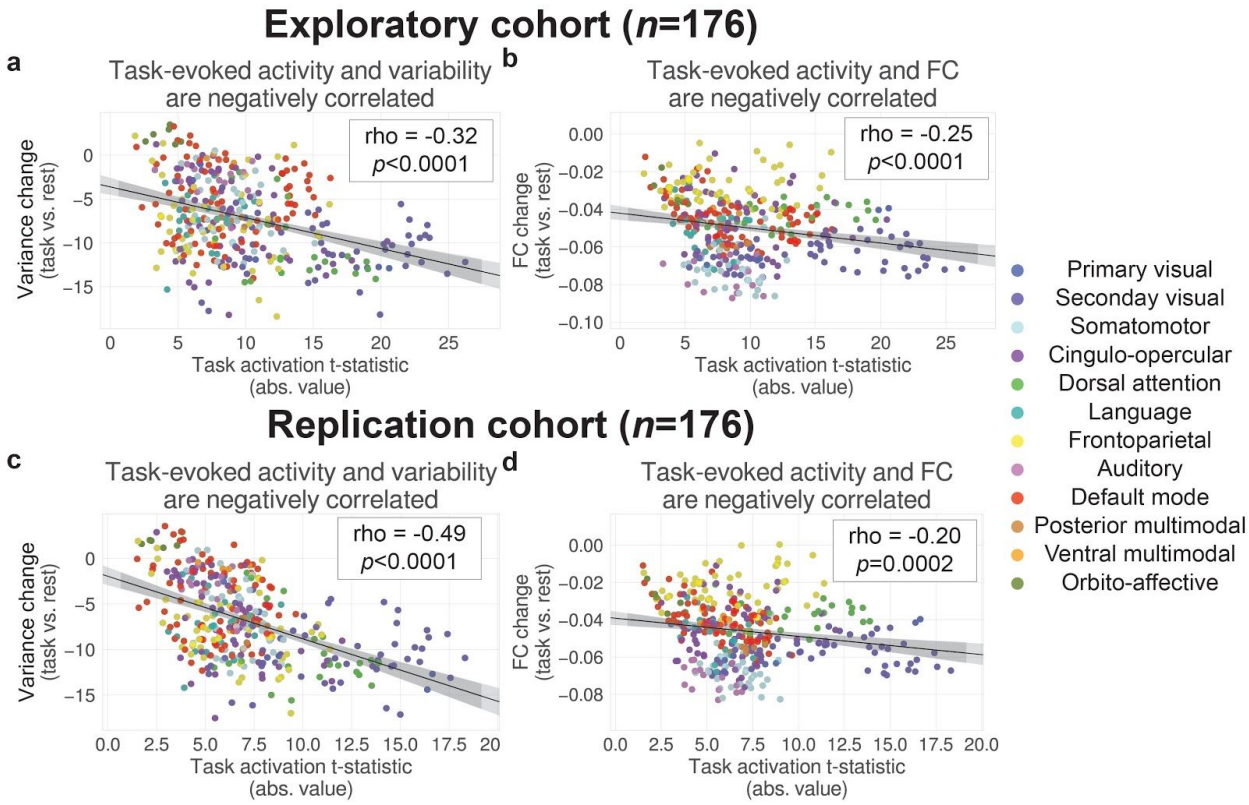


Supplementary Figure 1. Replication analysis for the excluded NHP subject. This figure is organized identically to Figure 4, but using data from a replication subject. We find nearly identical patterns between the exploratory and replication subjects, with the exception that we did not replicate any correlation increases. **a)** Multi-unit spike recordings from six different cortical regions fed into our analyses. **b)** As in our empirical fMRI data set, we calculated the global variability across task and rest states (estimated using the standard deviation across trials). **c)** We then calculated the global neural correlation (i.e., the spike count correlation across trials) for task and rest states between all pairs of recorded brain regions. (Spike rates were averaged within each cortical area.) **d-f)** For each pair of brain regions, we visualized the correlation matrices between each recording site for the averaged rest, task, and the differences between task versus rest state spike count correlations. For panels **d-f**, plots were thresholded and tested for multiple comparisons using an FDR-corrected p<0.05 threshold. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, grey line indicates the median, and the distribution is visualized using a swarm plot.

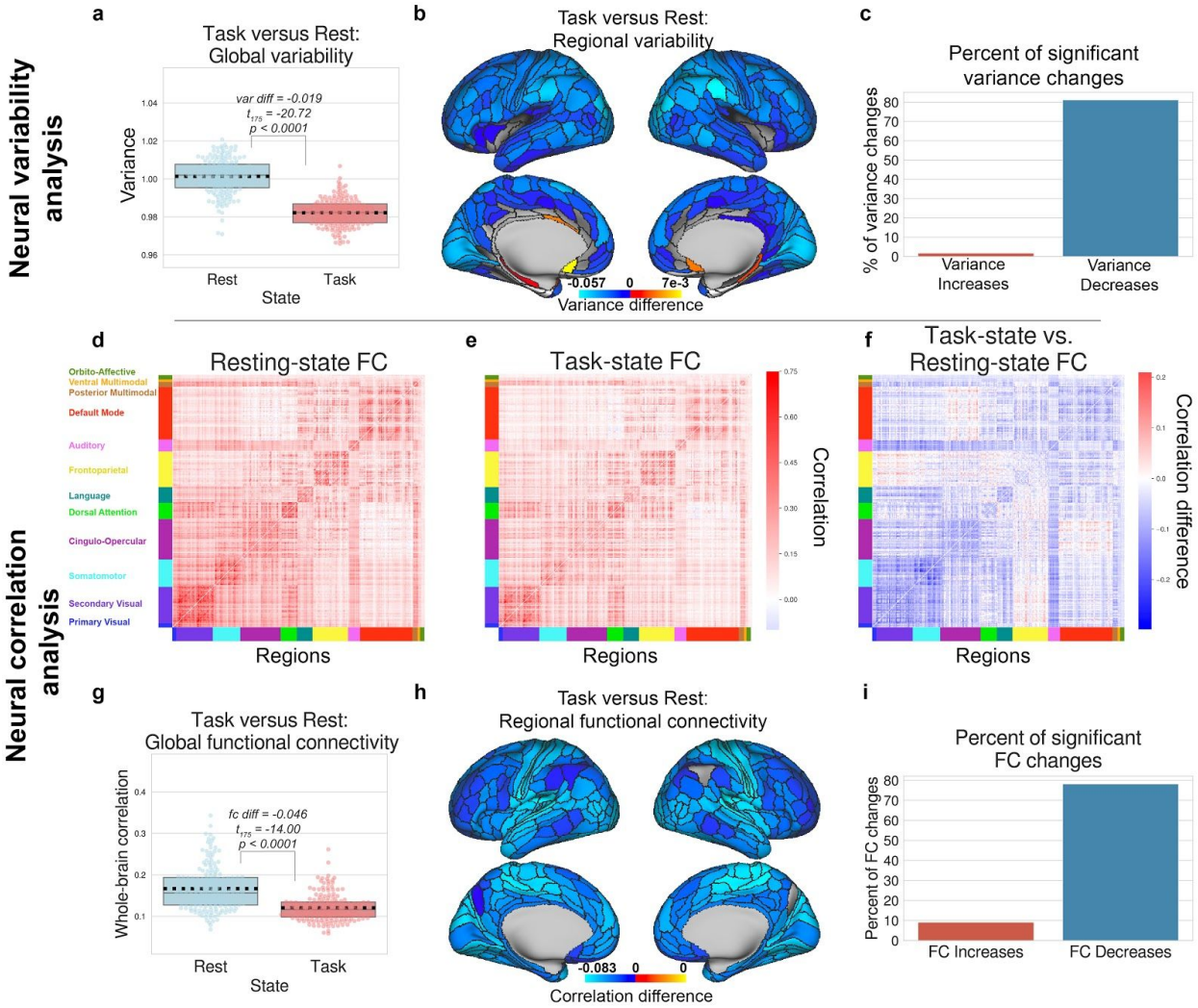


Supplementary Figure 2. Neural correlations and variability are quenched within trials from rest to task intervals. We analyzed the variability across time points (within trial) during ITIs and task cue periods to evaluate whether correlation and variability quenching also occurred on a moment-to-moment

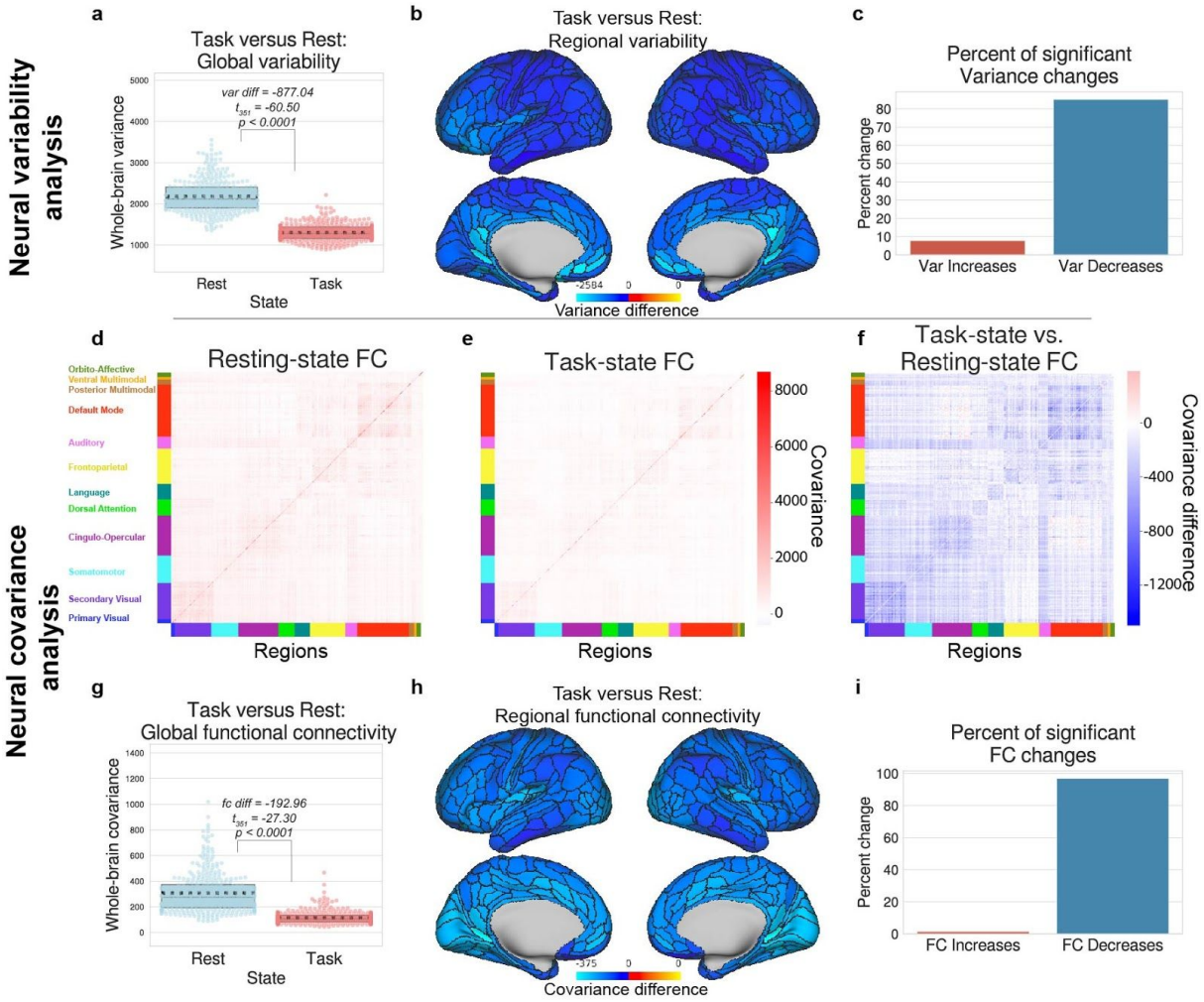
basis (i.e., faster timescale). Task cue intervals and ITIs were matched to have equivalent time points on a trial-by-trial basis. **a1,a2)** Global variability across the two states (estimated using the variance across time points) between task and rest state windows. **b1,b2)** We then calculated the global spike count correlation between the exact same task cue intervals with equivalent rest intervals between all pairs of recorded brain regions. (Spike rates were averaged within each cortical area.) **c1,c2)** We also calculated the global firing rate (averaged across all recording areas) during the task interval and rest interval. **d1-f1,d2-f2)** For each pair of brain regions, we visualize the spike count correlation matrices between each recording site for the averaged rest, task, and the differences between task versus rest state spike count correlation. For panels **d-f**, plots were thresholded and tested for multiple comparisons using an FDR-corrected $p<0.05$ threshold. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, grey line indicates the median, and the distribution is visualized using a strip plot.



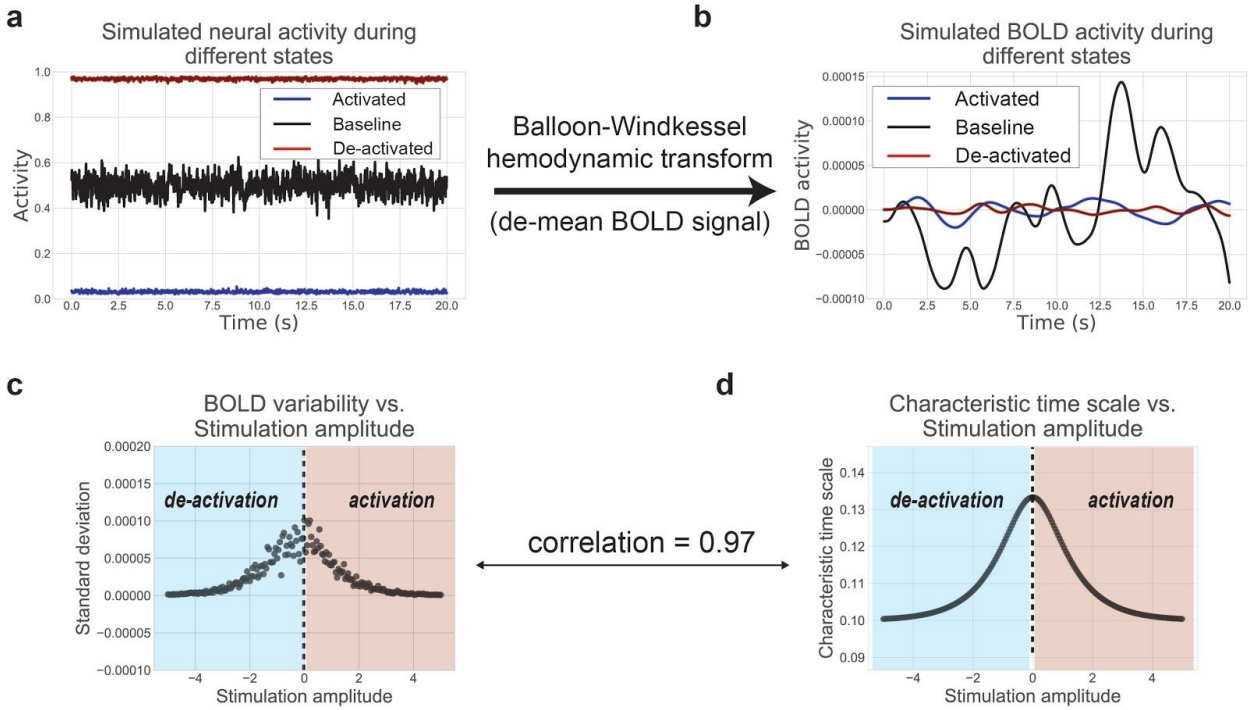
Supplementary Figure 3. Task-evoked activity is negatively correlated with neural variability and correlations across regions in fMRI data. a) We replicated a previous result (He, 2013), demonstrating that regions that activated more during tasks tend to decrease their neural variability more during task states. **b)** We extended those results to evaluate the relationship between task-evoked activity and FC across regions. We found that regions that activated more during tasks tend to decrease their global functional FC accordingly during task states. Scatter plots reflect each parcel in the Glasser atlas (Glasser et al., 2016), and are colored according to network affiliation (Ji et al., 2018). Best fit lines were estimated using linear regression, but correlations were calculated using a non-parametric rank correlation. **c,d)** Replication of panels a,b, respectively using the replication cohort of subjects. Statistics were calculated using the same steps as in (He, 2013). To calculate the averaged regional task activation, we first performed a group t-test for each task against 0, took the absolute value of the t-statistic, and then averaged across tasks. To calculate the averaged regional FC and SD, we performed a group t-test against 0 for each region. We then correlated these values across regions to measure the relationship between activity and FC, and activity and SD.



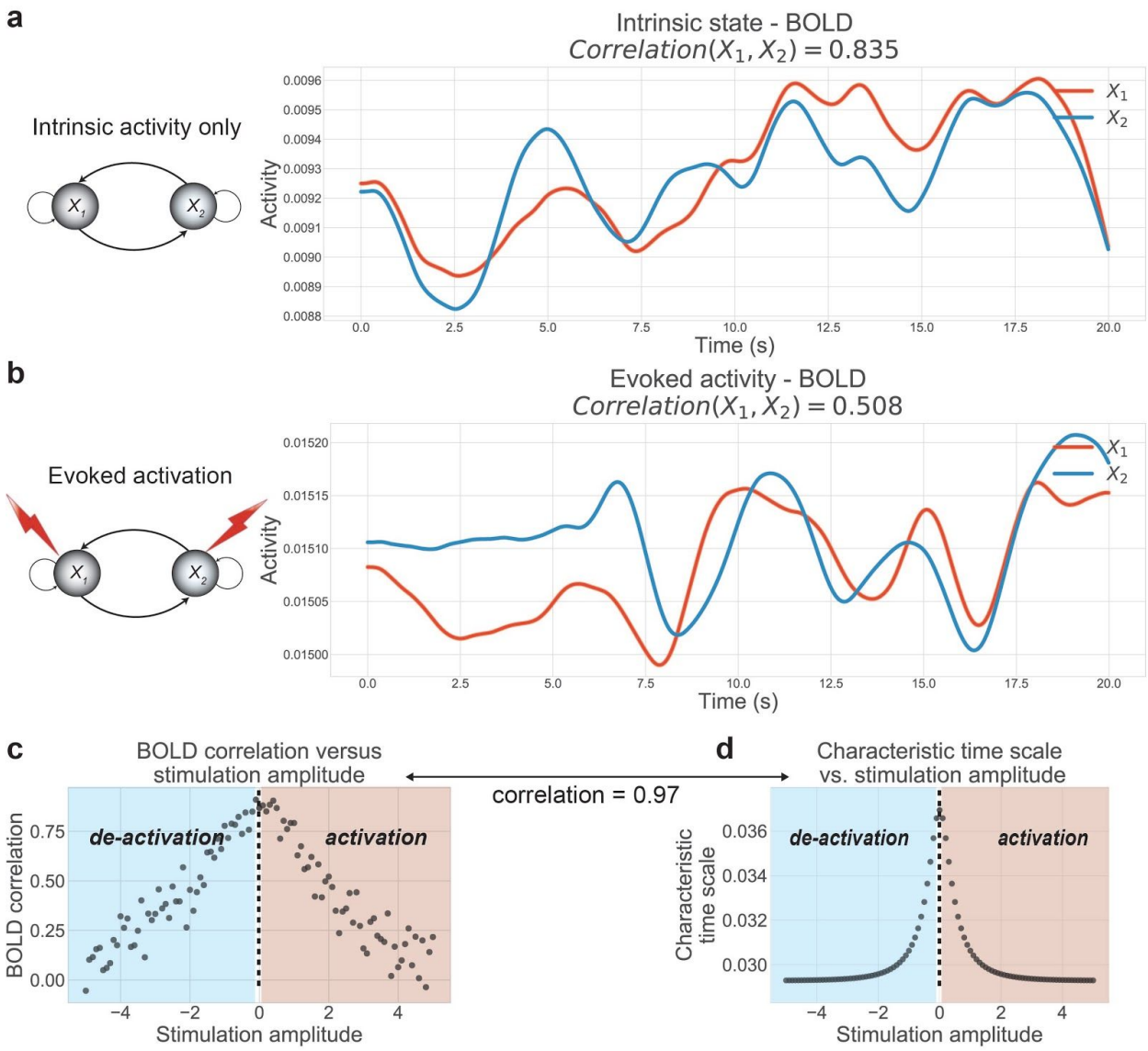
Supplementary Figure 4. Replication data set: Neural variability and correlations decrease during task states in human fMRI data. We successfully replicated results from Figure 3 using our held-out cohort of 176 subjects. **a)** We first compared the global variability during task and rest states, which is averaged across all brain regions, and then **b)** computed the task- versus rest-state variability for each brain region. **c)** The percentage of brain regions with significant increases/decreases of neural variability during task. **d)** We next compared the correlation matrices for resting state blocks with **(e)** task state blocks, and **(f)** computed the task- versus rest-state correlation matrix difference. **g)** We found that the average FC between all pairs of brain regions is significantly reduced during task state. **h)** We found that the average correlation for each brain region, decreased for each brain region during task state. **i)** We calculated the percentage of correlation increases/decreases for all pairs of brain regions. For panels **b-f, h,** and **i**, plots were tested for multiple comparisons using an FDR-corrected $p<0.05$ threshold. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, grey line indicates the median, and the distribution is visualized using a swarm plot.



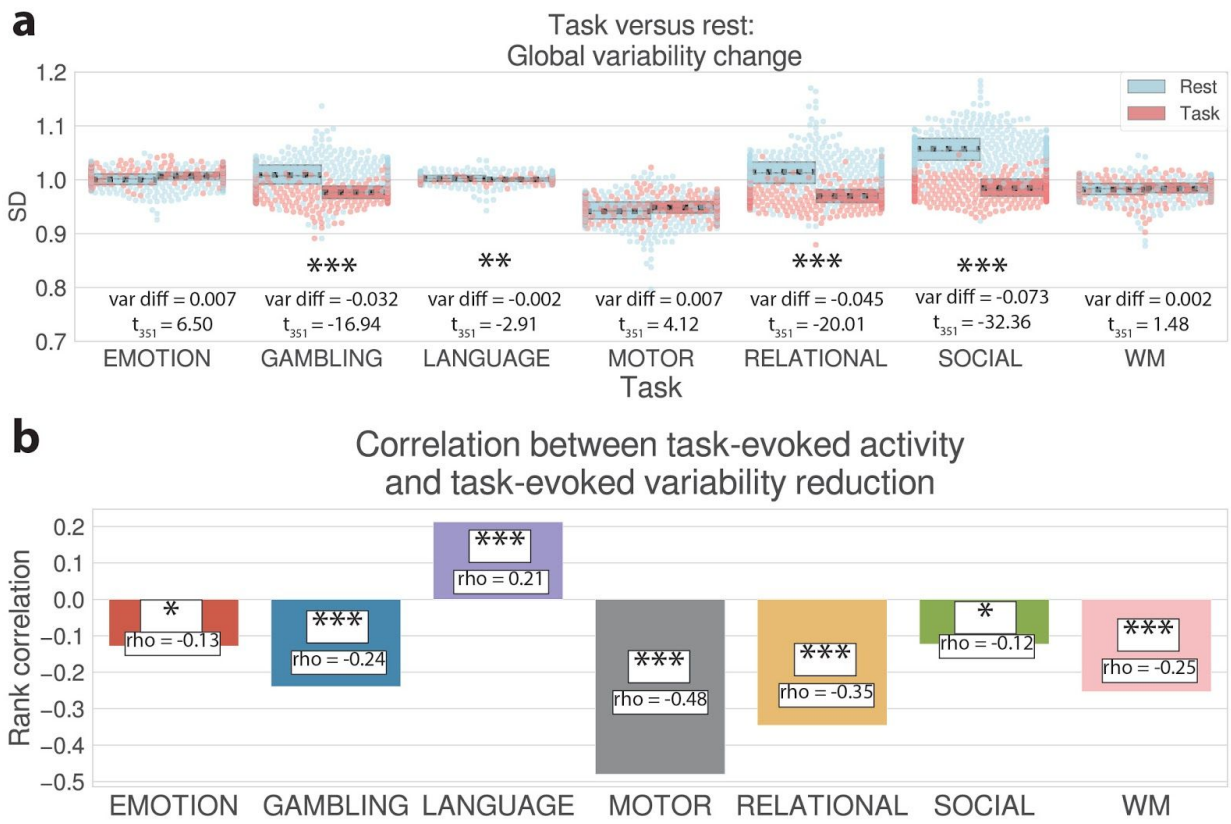
Supplementary Figure 5. Non-normalized data using variance and covariance, using the full set of 352 subjects. Neural variance and covariance decreased during task states in human fMRI data. We successfully replicated results from Figure 3 using, but without z-normalizing the time series (and using covariance instead of correlation). The combination of reduced correlations (Figure 3) and covariance measures suggested that shared signal dynamics is reduced from task to rest (Cole et al., 2016a; Duff et al., 2018; Siegel et al., 2012). **a)** We first compared the global variability during task and rest states, which is averaged across all brain regions, and then **b)** computed the task- versus rest-state variability for each brain region. **c)** The percentage of brain regions with significant increases/decreases of neural variability during task. **d)** We next compared the covariance matrices for resting state blocks with **(e)** task state blocks, and **(f)** computed the task- versus rest-state covariance matrix difference. **g)** We found that the average covariance between all pairs of brain regions is significantly reduced during task state. **h)** We found that the average covariance for each brain region, decreased for each brain region during task state. **i)** We calculated the percentage of covariance increases/decreases for all pairs of brain regions. For panels **b-f, h,** and **i**, plots were tested for multiple comparisons using an FDR-corrected p<0.05 threshold. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, grey line indicates the median, and the distribution is visualized using a swarm plot.



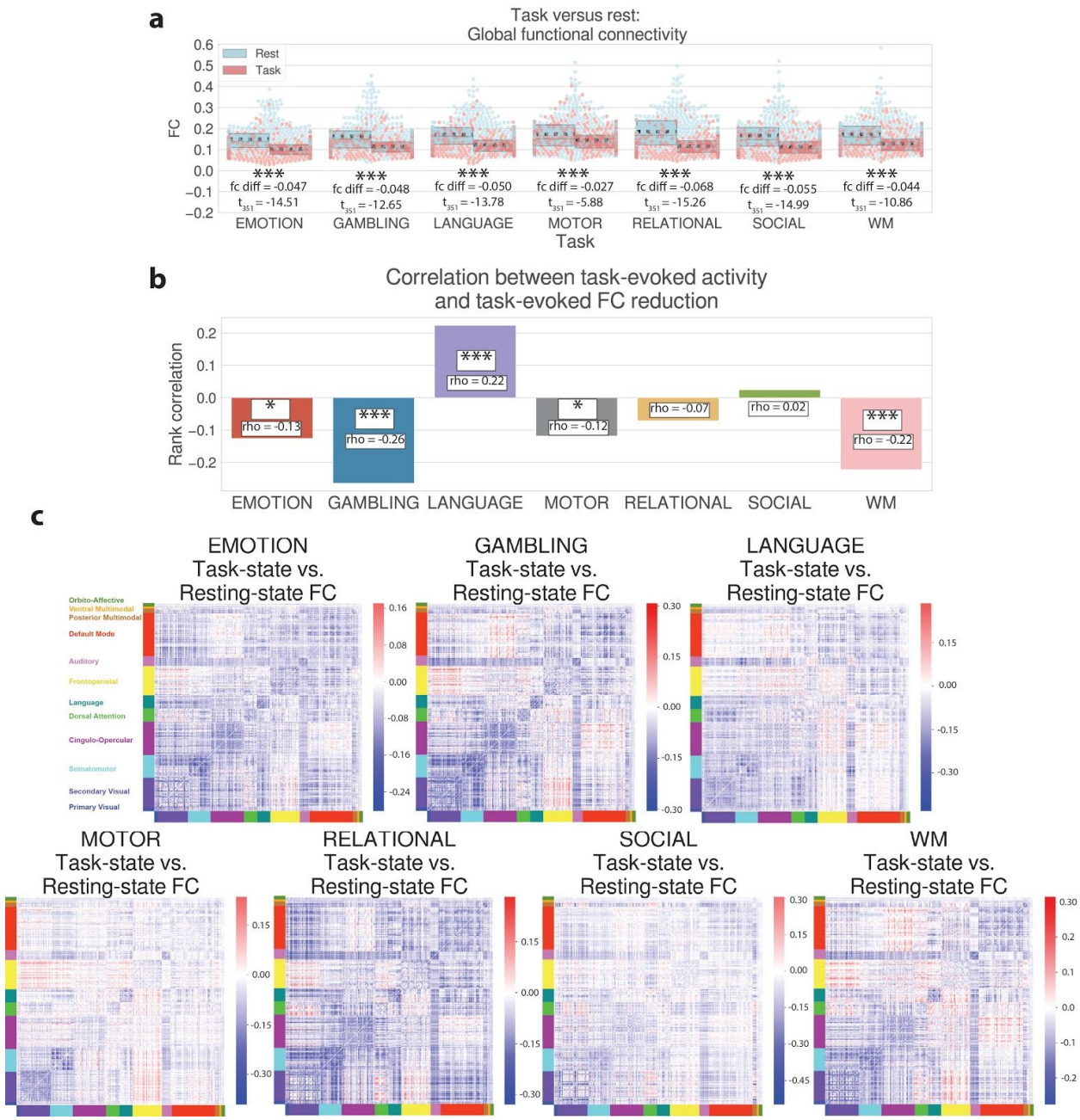
Supplementary Figure 6. Task-state neural variability reduction is preserved in the BOLD signal in the neural mass model. a) We simulated the neural mass model under the same three stimulus conditions (de-activated, baseline, and activated states) as in Figure 6A. **b)** We subsequently applied the Balloon-Windkessel transformation to the simulated neural activity, a nonlinear transformation from neural activity to the fMRI BOLD signal (Friston et al., 2003). Notably, the transformation assumes a nonlinear transformation of the normalized deoxyhemoglobin content, normalized blood inflow, resting oxygen extraction fraction, and the normalized blood volume. All BOLD signals were de-meaned such that it is possible to visually compare the time series variance of each stimulus condition. **c)** We simulated BOLD activity under a range of stimulus conditions and calculated the standard deviation of each time series. **d)** We calculated the rank correlation of the standard deviation of the BOLD signal across stimulus conditions with the characteristic time scale at each condition.



Supplementary Figure 7. Task-state neural correlation reduction is preserved in the BOLD signal in the two-unit neural mass model. a-b) Using the simulated the neural mass data in Figure 7, we applied the Balloon-Windkessel transform to convert our neural data into BOLD data (Friston et al., 2003). **c)** We simulated BOLD activity under a range of stimulus conditions and calculated the neural correlation between the two units of each. **d)** We calculated the rank correlation of the neural correlation of the BOLD signal across stimulus conditions with the characteristic time scale at each condition.

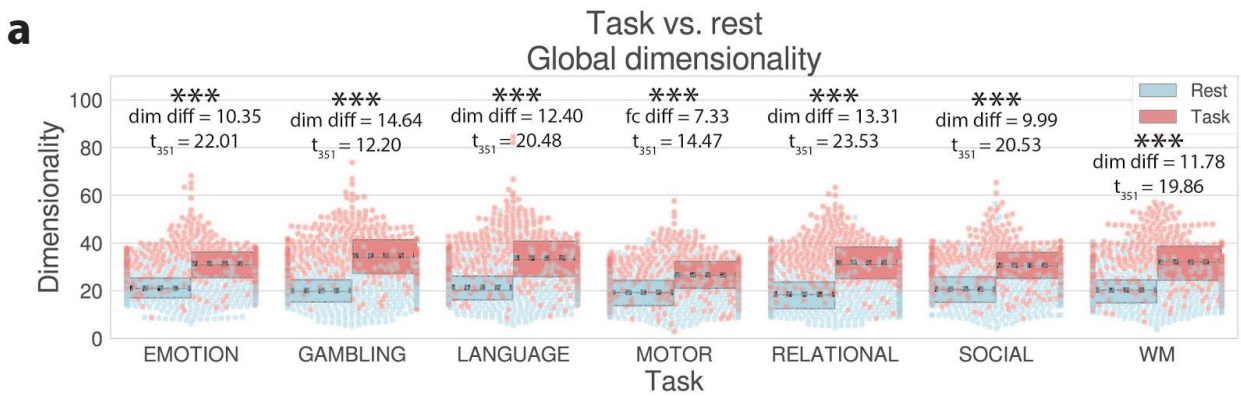


Supplementary Figure 8. fMRI variability reduction analysis for each of the 7 HCP tasks separately. a) This panel is identical to the analysis performed in Figure 3A, except that it was performed on each HCP task separately. Global variability, averaged across all regions, was reduced for 4/7 of the HCP tasks. Global variability was not reduced for the Emotion and Motor tasks, though task-evoked activity was correlated with task-evoked variability reduction across space (see next panel). **b)** This panel is identical to the analysis performed in Supplementary Figure 3A, except that the spatial correlation was performed on each HCP task separately (and is visualized as a bar plot rather than a scatter plot). Regional task-evoked variability was significantly negatively correlated with the magnitude of task-evoked activation (absolute value) for 6/7 of the HCP tasks. All analyses (in panels a and b) were corrected for multiple comparisons using FDR correction. (***= FDR-corrected $p<0.0001$; **=FDR-corrected $p<0.01$; *=FDR-corrected $p<0.05$). Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, and the distribution is visualized using a swarm plot.



Supplementary Figure 9. Task versus rest fcMRI analysis for each of the 7 HCP tasks separately. **a)** This panel is identical to the analysis performed in Figure 3G, except that it was performed on each HCP task separately. Whole-brain FC, averaged across all pairs of regions, was reduced for 7/7 of the HCP tasks. **b)** This panel is identical to the analysis performed in Supplementary Figure 3B, except that the spatial correlation was performed on each HCP task separately (and is visualized as a bar plot). Regional task-evoked FC was significantly negatively correlated with the magnitude of task-evoked activation (absolute value) for 4/7 of the HCP tasks. All analyses (in panels A and B) were corrected for multiple comparisons using FDR correction. (***= FDR-corrected $p<0.0001$; **=FDR-corrected $p<0.01$; *=FDR-corrected $p<0.05$). **c)** Task- versus rest-state FC analysis for each of the 7 HCP tasks separately. (This figure is identical to Fig. 3f, except that the statistics were performed on each task separately.) Though whole-brain FC differences from task to rest are different for each task, there are mostly FC

decreases during task state relative to rest state. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, and the distribution is visualized using a swarm plot.



Supplementary Figure 10. Task versus rest dimensionality comparison for each of the 7 HCP tasks separately. a) This panel is identical to the analysis performed in Figure 4A, except that it was performed on each HCP task separately. Whole-brain dimensionality increased from rest to task states for each of the 7 HCP tasks. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, and the distribution is visualized using a swarm plot. (***= FDR-corrected $p<0.0001$; **=FDR-corrected $p<0.01$; *=FDR-corrected $p<0.05$)