Behavioral/Cognitive

Edge-Based General Linear Models Capture Moment-to-Moment Fluctuations in Attention

Henry M. Jones,[1,2] Kwangsun Yoo,[3,4,5] Marvin M. Chun,[3,6,7] and Monica D. Rosenberg[1,2,8]

[1]Department of Psychology, The University of Chicago, Chicago, Illinois 60637, [2]Institute for Mind and Biology, The University of Chicago, Chicago, Illinois 60637, [3]Department of Psychology, Yale University, New Haven, Connecticut 06520, [4]Department of Digital Health, Samsung Advanced Institute for Health Sciences and Technology, Sungkyunkwan University, Seoul 06355, Korea, [5]Data Science Research Institute, Research Institute for Future Medicine, Samsung Medical Center, Seoul 06351, Korea, [6]Wu Tsai Institute, Yale University, New Haven, Connecticut 06520, [7]Department of Neuroscience, Yale University, New Haven, Connecticut 06520, and [8]Neuroscience Institute, The University of Chicago, Chicago, Illinois 60637

Although we must prioritize the processing of task-relevant information to navigate life, our ability to do so fluctuates across time. Previous work has identified fMRI functional connectivity (FC) networks that predict an individual's ability to sustain attention and vary with attentional state from 1 min to the next. However, traditional dynamic FC approaches typically lack the temporal precision to capture moment-to-moment network fluctuations. Recently, researchers have "unfurled" traditional FC matrices in "edge cofluctuation time series" which measure timepoint-by-timepoint cofluctuations between regions. Here we apply event-based and parametric fMRI analyses to edge time series to capture moment-to-moment fluctuations in networks related to attention. In two independent fMRI datasets examining young adults of both sexes in which participants performed a sustained attention task, we identified a reliable set of edges that rapidly deflects in response to rare task events. Another set of edges varies with continuous fluctuations in attention and overlaps with a previously defined set of edges associated with individual differences in sustained attention. Demonstrating that edge-based analyses are not simply redundant with traditional regions-of-interest–based approaches, up to one-third of reliably deflected edges were not predicted from univariate activity patterns alone. These results reveal the large potential in combining traditional fMRI analyses with edge time series to identify rapid reconfigurations in networks across the brain.

Key words: attention fluctuations; edge cofluctuations; fMRI; replication; vigilance

---

Significance Statement

For over 30 years, task-based fMRI has been a key pillar in the development of cognitive-neuroscientific theories. However, the standard analysis—fitting general linear models (GLMs) with regressors of cognition and behavior to activation time series—treats each significant cluster as an independent unit. In contrast, the modern field of network neuroscience characterizes relationships between these units (called edges) and how they change across time or tasks. Here, we combine these frameworks in a general-use method, fitting GLMs to moment-to-moment estimates of edge cofluctuations. Results reveal edges that reliably predict rapid changes in attention across independent datasets. This method offers a new framework for understanding the neural underpinnings of fluctuating cognitive process.

Introduction

Our ability to selectively attend to task-relevant information varies across time. For example, while reading a book, you might periodically wonder if you have a message from a friend and check your phone. You may have to jump back a few paragraphs to find the place where you lost focus. Such attention fluctuations have been quantified with tasks in which participants attempt to detect rare target stimuli, either across prolonged periods (vigilance tasks; Mackworth, 1948) or within a rapid stimulus stream (continuous performance tasks, CPTs; Robertson et al., 1997).

The gradual-onset continuous performance task (gradCPT; Esterman et al., 2013; Rosenberg et al., 2013; Fortenbaugh et al., 2015), a modified CPT in which image stimuli gradually fade from one to the next, was developed to assess sustained attention fluctuations while avoiding attention-capturing abrupt stimulus onsets (but see Jun and Lee, 2021). The gradCPT requires participants to make button presses in response to images from a frequently presented category while withholding responses to images from a rare target category. Like other "not-X" CPTs, the gradCPT is a powerful tool for investigating sustained attention because it offers two methods for investigating attention fluctuations. First, participants' responses to rare targets can be analyzed to assess successful withholding (reflecting engaged attention) or failures to inhibit the prepotent response (indicating an attention lapse). Second, the variability of response times (RTs) to frequent-category images can be analyzed to produce a higher-frequency measure of attentional state known as the variance time course (VTC). Evidence suggests that more consistent RTs correspond to engaged attention. In contrast, more erratic responding (including both overly fast and slow responses) predict attention lapses and disengaged focus (Esterman et al., 2013; Rosenberg et al., 2013; Fortenbaugh et al., 2015; Kucyi et al., 2017; Jayakumar et al., 2023). Thus, VTC values, calculated as the absolute deviation of each RT from the participant mean, are thought to reflect fluctuating sustained attention.

Analyses of the neural activity underpinning sustained attention fluctuations have revealed seemingly counterintuitive results. When gradCPT RTs are consistent and performance is accurate, functional MRI activity in the default mode network (DMN)—typically associated with off-task and internally directed thought and mind wandering (Buckner et al., 2008, 2019)—tends to rise above baseline (Esterman et al., 2013, 2014; Kucyi et al., 2016, 2017; Fortenbaugh et al., 2018; Song et al., 2022). When RTs are erratic and performance suffers, fMRI activity in brain regions associated with top-down control, such as regions of the dorsal attention network (DAN), tends to increase (Esterman et al., 2013, 2014; Kucyi et al., 2016, 2017; Fortenbaugh et al., 2018; Song et al., 2022). This pattern of results may reflect practiced (Mason et al., 2007), less effortful performance during high attentional states and more effortful performance that relies on top-down control mechanisms during low attentional states (Esterman et al., 2013).

The activity approach has been informative but is limited in two ways. First, univariate BOLD activity has been shown to be a limited description of the signals present in the brain (Serences et al., 2009). Second, although results are often examined at the network level, activation approaches cannot always speak to the functional configuration of networks. For example, activity in regions within a network may become more synchronized with increasing attention, without changing in overall activation.

An alternative approach has involved studying functional networks more directly by relating individual functional connectivity (FC) matrices, computed using time series data collected while participants performed attention tasks or rest, to attention function (Kessler et al., 2016; Poole et al., 2016; Wu et al., 2020; Kucyi et al., 2021; Yoo et al., 2022). For example, one approach identified sets of functional connections, or edges, whose strength predicts individual performance on the gradCPT (Rosenberg et al., 2016). In addition, FC matrices from small time windows can predict ongoing sustained attention performance across a scale of 10 of seconds to minutes (Rosenberg et al., 2020; Kardan et al., 2022). Windowed FC approaches, however, are limited by the window size used, preventing them from capturing moment-to-moment fluctuations in attentional state. This contrasts with the relative temporal granularity of GLM fitting to univariate activity. While the hemodynamic response function (HRF) is generally too smooth and slow to capture subsecond attentional shifts (Fiebelkorn and Kastner, 2019), it can still capture relatively fast changes that occur over seconds.

In summary, although the two fMRI approaches described above have related univariate activity and FC to sustained attention, neither method directly tests for the type of relatively rapid network reconfigurations that may underlie fluctuations in attention. Traditional activation-based approaches are limited in the sorts of changes they can identify, while FC approaches are limited in their temporal granularity. These limitations are not restricted to research on sustained attention but apply to any cognitive phenomenon that can be understood at the network level and which varies from one moment to the next.

Recent advances in network neuroscience provide the possibility of combining the FC perspective with the ability to capture fluctuations on the order of seconds via GLM fitting. Specifically, researchers have begun to "unfurl" the FC matrix into its time-point-by-timepoint contributions, called edge cofluctuations (Faskowitz et al., 2020; Zamani Esfahlani et al., 2020). These edge cofluctuations reflect estimates of covariance between the pair of brain regions comprising the edge at each moment in the time series. The mean of these estimates across time is Pearson's r between the two regions' time series, the most common FC metric. Edge cofluctuations capture higher-frequency changes in covariance between brain regions than traditional dynamic FC approaches, which incorporate information across multiple observations in a time window.

Here, we combine the analytical methods for univariate activity with a FC perspective to ask how moment-to-moment network reconfigurations relate to fluctuations in sustained attention. To address this question, we examined two independent gradCPT fMRI datasets. We first examine the replicability of standard univariate activity approaches (Fortenbaugh et al., 2018). Following this, we fit standard GLMs to edge time series to capture attentional lapses and correlate the VTC with edges. In almost all cases, we identified a set of edges significantly deflected in both datasets. A large proportion of these edges are not predicted from just univariate activity. Furthermore, edges that reliably fluctuate with attentional performance significantly overlap with an existing FC model of sustained attention (Rosenberg et al., 2016). Finally, we compare significance thresholding approaches with varying assumptions and consider how the edge response functions may differ from the canonical HRF. We propose that fitting GLMs to edge time series offers a powerful lens into rapid network reconfigurations underlying cognitive operations.

Materials and Methods

Dataset 1

To assess the reproducibility of edge-behavior associations, we examined fMRI data from two independent datasets collected while participants performed the gradCPT (Esterman et al., 2013), which measures sustained attention and inhibitory control. The first dataset, dataset 1, was described in Rosenberg et al. (2016).

Participants
Dataset 1 included 25 participants with normal or corrected-to-normal vision (13 female, 12 male, age 18–32 years, mean age 22.7 years). One participant was excluded due to a task lag in one run, leaving 24 participants for analysis. Participants gave written informed consent in accordance with the Yale University Human Subjects Committee and were paid for their participation.

Experimental design
Participants completed three fMRI runs of the gradCPT. Runs consisted of four 3 min blocks (225 800-ms trials/block). Blocks were separated with 32-second rest blocks. Participants were instructed to respond to city scenes (90% of trials) and to withhold responses to mountain scenes (10%) with an emphasis on accuracy. An iterative algorithm assigned key presses to trials as described previously (Esterman et al., 2013). The algorithm first assigned unambiguous button presses, which occurred after image n was 80% cohered but before image $n + 1$ was 40% cohered. Ambiguous presses were then assigned to an adjacent trial if one of the two had no response, to the closest trial if both had no response, or to an adjacent city trial if the other adjacent trial was a mountain. If multiple presses could be assigned to a trial, the fastest response was selected. On rare occasions, the original algorithm incorrectly assigned responses from the first trial of a block to the last trial of the preceding block, resulting in long RTs. We recalculated RTs from the original dataset.

fMRI acquisition parameters
MRI data were collected at the Yale Magnetic Resonance Research Center on a 3 T Siemens Trio TIM system using a 32-channel head coil. Before task runs, an anatomical scan was collected using a magnetization-prepared rapid gradient echo (MPRAGE) sequence. Task runs consisted of 824 whole-brain volumes acquired using a multiband echo-planar imaging (EPI) sequence: repetition time (TR), 1,000 ms; echo time (TE), 30 ms; flip angle, 62°; acquisition matrix, 84 × 84; in-plane resolution, 2.5 mm^2; 51 axial-oblique slices parallel to the ac–pc line; slice thickness, 2.5 mm; and multiband, 3; acceleration factor, 2. A 2D T1-weighted image with the same slice prescription as the EPI images was collected for registration.

fMRI data preprocessing and parcellation
Data were preprocessed using BioImage Suite (Joshi et al., 2011) and custom Matlab scripts (MathWorks) as described previously (Rosenberg et al., 2016). The first eight TRs of each run were excluded from analysis. Motion was corrected using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/). Linear and quadratic drift, mean signal from cerebrospinal fluid, white matter, and gray matter were regressed from the data along with a 24-parameter motion model (6 motion parameters, 6 temporal derivatives, and their squares). Data were then temporally smoothed with a zero mean unit variance Gaussian filter. Functional MRI data were then parcellated using the 268-node Shen atlas (Shen et al., 2013).

Dataset 2
Dataset 2 was described in Yoo et al. (2022). Although rest and other task runs were also collected, we only analyze gradCPT runs here. Dataset 2 is available at https://nda.nih.gov/edit_collection.html?id=2402.

Participants
A total of 127 participants performed the gradCPT during two MRI sessions separated by ∼2 weeks on average. The original authors excluded 35 participants for excessive head motion (>3 mm maximum head displacement or >0.15 mm mean framewise displacement), fewer than 120 TRs after censoring, low data quality via visual quality assurance, or task performance 2.5 SD from the group mean in both sessions, leaving 92 participants (60 females, 32 males, ages 18–35 years, mean age 22.79 years). Exclusion criteria were applied to each run, and participants were excluded if they had no run for any of the five run types (three tasks, one rest, and one move-watching). Of these, 65 participants had behavioral performance measures in both gradCPT sessions. We analyze data from the 58 participants in this group with fewer than 10% of censored whole-brain volumes. Participants gave written informed consent in accordance with the Yale University Human Subjects Committee and were paid for their participation.

Experimental design
Participants completed one 10 min gradCPT run containing 740 trials in a single block in each MRI session. The task parameters are the same as those used in dataset 1.

fMRI acquisition parameters
fMRI data were collected at the Yale Magnetic Resonance Research Center and the Brain Imaging Center at Yale on 3 T Siemens Prisma systems using a 64-channel head coil. Before task runs, an anatomical MPRAGE scan was collected. Task runs consisted of 600 whole-brain volumes acquired using a multiband EPI sequence: TR, 1,000 ms; TE, 30 ms; flip angle, 62°; acquisition matrix, 84 × 84; in-plane resolution, 2.5 mm^2; 52 axial-oblique slices parallel to the ac–pc line; slice thickness, 2.5 mm; multiband, 4; and acceleration factor, 1.

fMRI data preprocessing and parcellation
Data were preprocessed using Analysis of Functional NeuroImages (AFNI version 17.2.07) as described previously (Yoo et al., 2022). The first three TRs of each run were excluded from analysis. Volumes were censored if they contained outliers in >10% of voxels or if the Euclidean distance of the head motion parameter derivatives were >0.2 mm. Data were then despiked, slice time corrected, and motion corrected. Mean signal from the cerebrospinal fluid, white matter, and gray matter were regressed from the data along with six motion parameters, six temporal derivatives, and their squares. Data were then aligned to the MPRAGE anatomical image and normalized to MNI space. fMRI data were then parcellated using the 268-node Shen atlas (Shen et al., 2013). Thirteen nodes were excluded from analysis in at least one participant due to imperfect acquisition of the voxels within the region, predominantly within the frontoparietal and motor networks (5 nodes each), and within the medial frontal (2 nodes) and subcortical-cerebellar (1 node) networks.

Analyses
Post hoc analyses and deviations from the preregistration are noted as such below. Throughout this paper, we use the term first-level model to refer to within-individual analyses, second-level model to refer to assessments of significance at the group level within each dataset, and third-level to refer to assessments of overlap between datasets. Visualization of connectivity matrices were built using seaborn's heat-map plotting functions (Waskom, 2021).

Code accessibility
Analyses were conducted in Python 3.10.2 using a combination of nilearn (https://nilearn.github.io/stable/index.html), scikit-learn (Pedregosa et al., 2011), and custom scripts. Our main analyses were preregistered (https://osf.io/ftzr4/) and the code is available (github.com/henrymj/edge_GLM).

First-level analyses: ROI-based activation
We first asked how univariate fMRI activity in the 268 Shen atlas ROIs covaried with attention fluctuations following analyses described in Fortenbaugh et al. (2018). Fortenbaugh et al. (2018) examined how changes in activity reflected fluctuations in attention and attention lapses during the gradCPT, replicating original results from Esterman et al. (2013) in a larger sample ($n = 140$; one 8 min gradCPT run per participant). They examined attentional lapses by contrasting pretrial and trial-evoked activity to erroneous responses to rare mountain scenes (commission errors, CEs) against correct withholding of responses (correct omissions, or COs). To characterize attention fluctuations, they correlated fluctuations in RT with activation time series. These analyses contextualize edge-based results by indicating the reproducibility of each analysis and demonstrating what edge-based analyses add beyond traditional ROI-based approaches (Our data and preprocessing steps differ from that of Fortenbaugh et al. (2018) in a few ways. For example, Fortenbaugh et al. (2018) regressed out the mean signal from a white matter and CSF mask, while Rosenberg et al. (2016) and Yoo et al. (2022) regressed out the global signal, which includes the gray matter. Therefore, this analysis also assesses whether the results of Fortenbaugh et al. (2018) are robust to processing differences).

Event-based contrasts. We built a first-level model with the following regressors: (1) incorrect presses to rare mountain scenes (commission errors, CEs), which occur when participants fail to inhibit their prepotent response and indicate a lapse in attention; (2) correct withholding of responses to rare mountain scenes (correct omissions, COs), indicating

that participants successfully inhibited their prepotent response and thus are likely in a focused, high-attention state; and (3) incorrect failures to respond to common city scenes (omission errors, OEs), which may result from a variety of sources.

As in Fortenbaugh et al. (2018), we did not model correct responses to common city scenes (correct commissions) due to their high frequency.

Stick regressors for each modeled trial type were convolved with the SPM HRF model and its temporal derivative to create the design matrix. Because nuisance regressors (motion, drift, white matter, global signal) were regressed out during preprocessing, they were not included in these models. To better equate the baseline period between dataset 1, which included rest breaks between task blocks, and dataset 2, which did not contain breaks, we included a single block break regressor per run which lasted the duration of each break in dataset 1.

We fit each participant's time series for each run to the design matrix using nilearn's autoregressive (AR) model, which whitens the data according to the autoregressive covariance structure in the time series. Censored timepoints in dataset 2 were interpolated from adjacent timepoints, and an additional impulse regressor was included in the model for each censored timepoint. We also confirmed that excluding censored timepoints and fitting an ordinary least squares model did not qualitatively change results. Statistics were averaged across runs within each subject.

Using this model, we computed the four main contrasts of Fortenbaugh et al. (2018): CE against baseline, CO against baseline, OE against baseline, and CO − CE. One run from one participant in dataset 1 is excluded, as they did not have any COs for that run; the participant's remaining two runs were kept. One participant in dataset 2 is excluded from these contrasts as they did not make any OEs, leaving a sample size of 57 in dataset 2 for these analyses.

Correlating activity with RT variability. The contrasts above ask how activity differs between trial types. To characterize how activity fluctuates with a continuous measure of attentional state, we separately correlated activation time series with each run's VTC. The VTC is computed using the RTs on correct frequent-category trials (i.e., correct presses to city trials or correct commissions) as described previously (Esterman et al., 2013; Rosenberg et al., 2013). RTs were z scored and converted to absolute values. Trials missing responses (OEs and COs) were linearly interpolated from the values of the two surrounding trials. This time course was then smoothed using a Gaussian kernel with a FWHM of nine trials, producing the VTC. The VTC was computed separately within each of the four blocks per run in dataset 1. In dataset 2, each run consisted of a single block. To account for the lag of the HRF, we shifted the VTC by 6 s following Fortenbaugh et al. (2018).

We then correlated the VTC with residuals of each ROI's BOLD-signal time series after fitting the event-based GLM described above. Timepoints from the between-block breaks were excised from the residuals in dataset 1. Correlation values were converted to z values using Fisher's transformation to be assessed using our second- and third-level paradigms below.

Identifying precursors of attention lapses. The above analyses asked how ROI activity corresponds to attention lapses, task events, and continuous fluctuations in sustained attention. We next asked whether activity preceding rare mountain trials *predict* upcoming COs versus CEs. To do so, we replicated the trial precursor analysis of Fortenbaugh et al. (2018), applying the details of their whole-brain analysis. ROI time courses were linearly interpolated in time to the onset of each trial. For each CO or CE, the preceding timepoints from −4.8 s (6 trials) to the onset of the trial were averaged. These preceding averages were averaged across trials of the same type, and a difference map (CO minus CE) was generated for each participant. As in the contrast analyses above, one run from one participant in dataset 1 was excluded, as they did not produce any COs for that run; the participant's remaining 2 runs were kept.

First-level analyses: edge cofluctuations

We next asked how edges rather than activity in a single region alone covary with attention. To this end, we computed edge cofluctuation

time series for a given pair of nodes (ROIs), i and j, following Faskowitz et al. (2020). Specifically, each ROI's time series vector (i or j) was z scored, and the elementwise product between the two standardized vectors was computed. Mathematically, the edge time series e_{ij} is computed as follows:

$$e_{ij} = \frac{(i - \bar{i})}{S_i} \cdot \frac{(j - \bar{j})}{S_j} = Z(i) \cdot Z(j),$$

where $\bar{x}$ is the mean and S_x is the standard deviation across the time series for node x:

$$S_x = \sqrt{\frac{\sum_{t=0}^{N}(x_t - \bar{x})^2}{N-1}}$$ and N is the number of timepoints in the time series.

We replicated all activation analyses with edge (i.e., cofluctuation) time series as described in detail below.

Relating cofluctuations to attention lapses and salient events. We first asked how edges covary with attention lapses and fluctuations. To do so, we built first-level models with regressors for the following trial types: CEs, COs, and OEs. As in our ROI activity analyses, we computed the four main contrasts of Fortenbaugh et al. (2018): CE against baseline, CO against baseline, OE against baseline (We did not mention OE against baseline in our preregistration, but we included the contrast to help ground our replication of Fortenbaugh et al. (2018)), and CO − CE. In addition to these four contrasts, we included a CO + CE contrast, which, given the structure of the task, is equivalent to a no-go over baseline contrast or an oddball (i.e., rare mountain trial) versus baseline contrast. This analysis allows us to identify whether certain task-based events (attention lapses, correct inhibition, no-go, and oddball processing) are accompanied by momentary changes in the functional interactions between brain regions.

Relating cofluctuations to RT variability. We next explored whether edges varied systematically with the VTC, the continuous measure of attention computed from RT deviance (Esterman et al., 2013; Rosenberg et al., 2013). As in our ROI activity analyses, we correlated the lag-shifted VTC with edge time series after excising timepoints from the between-block breaks in dataset 1 and converted the correlations to z values using Fisher's transformation for group-level assessment.

In addition to this time-shifted smoothed VTC analysis, we asked whether a more traditional parametric regressor analysis (Esterman et al., 2013) might be better powered than the original correlational approach. Rather than smoothing the VTC with a Gaussian kernel, we convolved the nonsmoothed VTC with the SPM HRF function and its temporal derivative and included these in a model with trial event regressors, similar to that described above. We included an event regressor for each of the four trial types (correct commissions, COs, CEs, and OEs). A block break regressor was included for dataset 1. We fit this model to the edge time series and computed the contrast VTC over baseline (In the preregistration, we also planned to fit a similar model, but with the VTC demeaned before convolution. However, this transformation only changes the beta values of the unmodulated event regressors, not that of the VTC regressor. Because we constrain our focus to the VTC results to compare with previous literature (Fortenbaugh et al., 2018, Esterman et al., 2013), we have omitted the demeaned approach).

Identifying precursors of attention lapses in edge cofluctuations. As in our ROI activity analyses, we asked whether edge cofluctuations preceding rare mountain trials predict upcoming COs versus CEs. First, edge time courses were linearly interpolated in time to the onset of each trial and then averaged across the preceding timepoints from −4.8 s to the onset of each rare mountain trial. We produced a difference map (CO − CE) of these average preceding values for each participant.

Second-level significance testing via permutation

We tested the significance of each first-level analyses at the group level using two different family-wise error rate corrections: (1) max-T for

both ROI and edge analyses and (2) the network-based statistic (NBS; Zalesky et al., 2010) for edge analyses. These methods rely on different assumptions to assess statistical significance. We provide both results so that other researchers interested in applying edge-based GLMs can use the one best suited for their research question.

Max-T controls for type 1 error by producing a null distribution of the maximum test statistics (or minimum p values) across random permutations of the data, to which the observed edge test statistics can be compared. It makes no assumptions about the relationships between tests (in other words, about the relationship between edges, or ROIs), so it can be applied to both ROI activity and edge cofluctuation first-level maps.

NBS is a graph-specific procedure that also controls for type 1 error by constructing a null distribution across random permutations of the data, but it makes the additional assumption that edges that are significantly different from the null will be interconnected in subcomponents of the full graph. This is analogous to the assumption that significant voxels will be spatially autocorrelated which underlies cluster correction in traditional fMRI contrasts. Thus, the null distribution for NBS is formed of maximum subcomponent sizes across permutations. In both cases, we used the sign-flipping paradigm from FSL's randomise algorithm (Winkler et al., 2014) to randomly permute edges or ROIs.

First, we fit a second-level model to the first-level outputs and obtained an observed test statistic for each edge. Because this work is exploratory and we have no directional hypotheses, we chose to use the F statistic. We then ran 10,000 permutations of the second-level model by randomly flipping the signs of the intercept column (i.e., multiplying each participant's vector of edge statistics by 1 or −1) and computing new F statistics for each edge.

To build the max-T distribution, we recorded the largest F value. To build the NBS distribution, we thresholded the permutation's second-level graph to edge-behavior associations with $p < 0.01$ and recorded the size of the largest fully connected component, defined by the number of edges in the component (Although we preregistered 0.05 as our initial threshold, 0.01 is more in keeping with other works using NBS (Zalesky et al., 2010, Serin et al., 2021) and should reduce the number of false positives in the final components, making them more interpretable). For edges, the max-T and NBS null distributions were computed in parallel. The observed second-level maps were then thresholded using these two null distributions, producing two sets of results. For max-T, observed ROI-behavior or edge-behavior associations were thresholded to those with F statistics greater than the 95th percentile of the null distribution. For NBS, observed edge-behavior associations were thresholded to those with $p < 0.01$, and any components larger than the 95th percentile of the NBS null distribution were retained.

Third-level analyses for edges: between dataset comparisons via permutation

The reproducibility of the ROI results can be assessed via comparison with previous literature, such as Esterman et al. (2013) and Fortenbaugh et al. (2018). In contrast, there is no existing work to compare the edge results against. Therefore, we rely on comparisons between our two independent datasets to assess the reliability and reproducibility of these results. To assess whether the edges identified as significant were consistent across datasets, we relied on a permutation test with slight differences in implementation to account for the different assumptions of our two second-level paradigms.

To assess overlap in the max-T results, we recorded the number of overlapping significant edges between our two datasets. Across 10,000 permutations, we shuffled the locations of significant edges in both datasets and recorded the number of overlapping edges.

To assess overlap in the NBS results, we repeated the above process, but we shuffled node locations rather than edges, as NBS makes assumptions about the connections within components. Shuffling nodes therefore randomizes edge locations while preserving the structure of the observed components.

For both max-T and NBS, a result was considered to have significant consistency if the number of observed overlapping edges was larger than

the 95th percentile of the null distribution of overlapping edges (Here we note two differences from our preregistration. First, the permutation testing approach differs from that described in our preregistration, which involved assessing the overlap of significant edges between datasets by assuming a hypergeometric distribution. The nonparametric permutation testing approach reduces the number of assumptions required when assessing significance with this new method. Second, in our preregistration we described assessing which networks contained more significant edges than expected by chance and comparing the overlap of networks across datasets. However, this approach is redundant with testing the overlap of edges across datasets and describing the distribution of overlapping edges across networks).

Assessing dataset-level overlap with edges underlying individual differences in attentional control

This work aims to identify edges associated with sustained attentional performance by looking for differences from the null (i.e., no association) at the group level. In contrast, previous work has identified edges associated with attentional performance by examining individual differences in performance. Rosenberg et al. (2016) applied connectome-based predictive modeling (CPM; Finn et al., 2015; Shen et al., 2017) to predict individuals' overall gradCPT performance (d'). CPM uses a cross-validation approach to identify edges whose strengths are most positively or negatively correlated with attentional performance. The set of most positively correlated edges is termed the high-attention network (HAN), and the set of most negatively correlated edges is termed the low-attention network (LAN). CPM identifies the HAN and LAN in a training set of data and then fits a simple linear model to that training set to predict d' for a given individual as the weighted combination of their average high-attention edge strength and their average low-attention edge strength. This model is then tested on held out data.

In this exploratory analysis, we asked whether the edges identified by our group-level approach overlapped with those identified via CPM. To do this, we assessed the overlap between preregistered HAN and LAN edges and the second-level results from three analyses isolating attentional control: (1) the CO − CE contrast, (2) the CO − CE precursor analysis, and (3) correlations with the VTC. We focused on overlap with NBS-based second levels, as the max-T results produced very sparse graphs. Because the CPM edges were identified in dataset 1, we examined the results for each dataset separately and focus on significant overlap between the CPM edges and dataset 2 results.

We assessed significant overlap between second-levels and CPM edges in nearly the same manner that we assessed significant overlap between datasets (see above, Second-level significance testing via permutation). We first recorded the observed overlap with the following assumptions. We assumed that positively deflected edges for CO − CE (for either the contrast or precursor analysis) should map onto HAN edges and negatively deflected edges should map onto LAN edges. In contrast, we expect HAN and LAN edges to map onto the edges correlated with the VTC differently. Increasingly positive VTC values reflect increasingly erratic responding, associated with reduced attention; edges which are positively correlated with the VTC reflect increased synchrony with reducing attention. Thus, we expect that edges which are positively correlated with the VTC should map onto LAN edges and edges which are negatively correlated with the VTC should map onto HAN edges. Then, across 10,000 permutations, the second-level maps were permuted at the node level to preserve the subcomponent structure and overlap with the CPM edges were recorded. Observed overlap was assessed against this null distribution.

Comparing observed edge results to ROI-based predictions

Do edge-based GLMs reveal new information about brain–behavior relationships, or are they redundant with the results of traditional ROI-based GLMs? To ask this question, we investigated whether the results based on edge cofluctuation time series contained information distinct from those contained in ROI time series and traditional analyses. In each dataset, we compared our second-level edge-based results against theoretical edge second levels predicted from ROI activity. To predict the edge-based results from the ROI results, we

converted each first-level ROI map to a "theoretical edge map" using the following formula:

$$\hat{e}_{ij} = \text{sign}(t_i \cdot t_j) \cdot \sqrt{|t_i \cdot t_j|},$$

where t is the t score for a given node (i or j in this example).

This formula captures a few straightforward predictions. First, if both nodes are deflected (i.e., have nonzero t scores), their edges are likely to also be deflected. Second, if both nodes are deflected in the same direction, this will likely lead to a positive deflection in the edge, and if the nodes are deflected in opposite directions, this will likely lead to a negative deflection in the edge.

We analyzed these maps using the second-level tests described above and compared contrasts with significant overlapping edges between datasets in both the ROI-based and edge-based maps. We predicted that there may be pairs of regions whose activation time series are not significantly deflected from baseline in response to task events, yet whose edge cofluctuation time series were.

Estimating HRF of edge time series

Though the timing of edge fluctuations should still be constrained by the speed of the HRF, it is possible that the standard SPM HRF provides a poor description of how edge time series are deflected by events. In this post hoc analysis, to estimate the HRF of edges in response to events, we employed finite impulse response (FIR) modeling and k-means clustering. We performed this estimation procedure using dataset 1 so that the estimated response functions may be applied to dataset 2 should these response functions differ in form from the HRF.

We fit our trial-type model to both the ROI and edge time series for each participant's first and second runs of dataset 1 separately. Rather than convolving the trial-type stick regressors with the HRF, we used 31 separate impulse regressors per trial-type (one for each TR of the first 30 s after an event). To increase power, we focused on parameter estimates for the CO + CE contrast, which includes the most data.

We then ran a group-level test for each ROI or edge at each timepoint and averaged across participants. Including only those ROIs or edges with a value significantly different from 0 for at least one timepoint produced similar HRFs to using all ROIs or edges. To avoid separating response functions with the same structure but different directions, we multiplied each time series by the sign of its largest magnitude value (i.e., if an ROI or edge's largest value was negative, the whole time series was flipped).

We then clustered the ROI or edge time series using k-means clustering in participants' first and second runs separately. We found an optimal number of clusters by testing all values of K between 2 and 30 (inclusive) and identifying the elbow in the distortion function across K values (yellowbrick; https://www.scikit-yb.org/en/latest/api/cluster/elbow.html). We took the mean, or centroid, of each cluster to be a representative HRF. We assessed the consistency of clusters by examining the adjusted mutual information (AMI) between the partitions and comparing observed AMI to a null distribution built by shuffling the cluster labels. Finally, we compared ROI and edge clusters to canonical network labels defined in Finn et al. (2015; e.g., DMN, frontoparietal network) to test whether ROIs and edges from the same network tended to show hemodynamic response profiles that clustered together.

Results

Activation patterns replicate previous work on sustained attention

In both datasets, the three baseline contrasts produced a similar pattern of results to that reported in Fortenbaugh et al. (2018; Fig. 1A; compare with their Fig. 4). COs were associated with greater activity in task-positive frontal and parietal regions previously associated with vigilant attention (Langner and Eickhoff, 2013) and lesser activity in regions of the DMN often associated with task-irrelevant or internally directed thought (Buckner et al., 2008, 2019). As in Fortenbaugh et al. (2018), lateral visual cortex activity increased with COs whereas medial visual cortex activity decreased. CEs were associated with similar activity patterns, potentially reflecting attention capture by and processing of rare target images. In addition, both trial types likely involve attempts at inhibitory control processes. OEs, which occurred rarely, were again accompanied by similar activity patterns including increased activity in the middle frontal gyrus, inferior parietal lobule, and insula and decreased activity in ventromedial prefrontal cortex. To facilitate visual comparison with the results of Fortenbaugh et al. (2018), we projected the ROI activation results onto the fsAverage surface mesh from FreeSurfer, using nilearn's functions *vol_to_surf* and *plot_surf_stat_map*. Vertices were assigned the value of the nearest voxel. To get a sense of the ROI boundaries on this surface projection, we mapped the ROI labels to a set of six circularly repeating, highly contrasting colors and surface projected this mapping in the same manner as well (Fig. 1C).

Contrasting COs with CEs (CO − CE) in dataset 2 replicated Fortenbaugh et al. (2018). Successful inhibition was associated with more activity in subcortical and cerebellar regions and visual and inferior parietal areas presumably driven by greater attention-to-task. Unsuccessful inhibition, on the other hand, was associated with greater medial prefrontal and dorsal anterior cingulate cortex and insula activity, potentially reflecting error processing (Dali et al., 2023). A sparser CO − CE map in dataset 1 suggests that this smaller sample may be underpowered to detect effects consistent in Fortenbaugh et al. (2018) and dataset 2, although this interpretation of the null effect is speculative.

Associations between continuous attention fluctuations and ROI activity replicated the overall pattern of results observed in Fortenbaugh et al. (2018; Fig. 1B; compare with their Fig. 7). In both datasets, the smoothed, lag-shifted VTC was negatively correlated with activity in regions of the DMNs, replicating previous associations between higher DMN activity and in-the-zone, more practiced sustained attention task performance (Mason et al., 2007; Esterman et al., 2013, 2014; Kucyi et al., 2017; Song et al., 2022). This relationship is hypothesized to reflect the fact that, during "in-the-zone" attentional states, good performance may be less effortful and need not rely on the suppression of DMN activity (Esterman et al., 2013). In contrast, higher activity in task-positive frontal and parietal regions was positively correlated with RT variability, potentially reflecting more effortful top-down attention during erratic periods of performance (Esterman et al., 2013; Kucyi et al., 2017; Fortenbaugh et al., 2018).

Pretrial activity did not significantly vary before COs versus CEs in dataset 1. In dataset 2, greater frontal and parietal activity preceded COs whereas greater ventromedial prefrontal cortex activity preceded CEs. This aligns with previous work suggesting that, while "task-positive" frontoparietal activity may decrease and "task-negative" DMN activity may increase during in-the-zone attentional performance, attention lapses can occur when frontoparietal activity is too low or DMN activity is too high (Esterman et al., 2013) (Whereas Esterman et al. (2013) observed more DMN and less DAN activity before CEs, Fortenbaugh et al. (2018) replicated the pretrial DMN but not DAN effect. We did not formally test for replication of these network-level effects as our analyses were performed at the ROI level using the Shen 268-node atlas. Shen atlas nodes have been grouped into a DMN mask but not a DAN mask). Notably, however, both datasets tested here have fewer participants than Fortenbaugh et al. (2018) and may be underpowered for detecting attention lapse precursors.
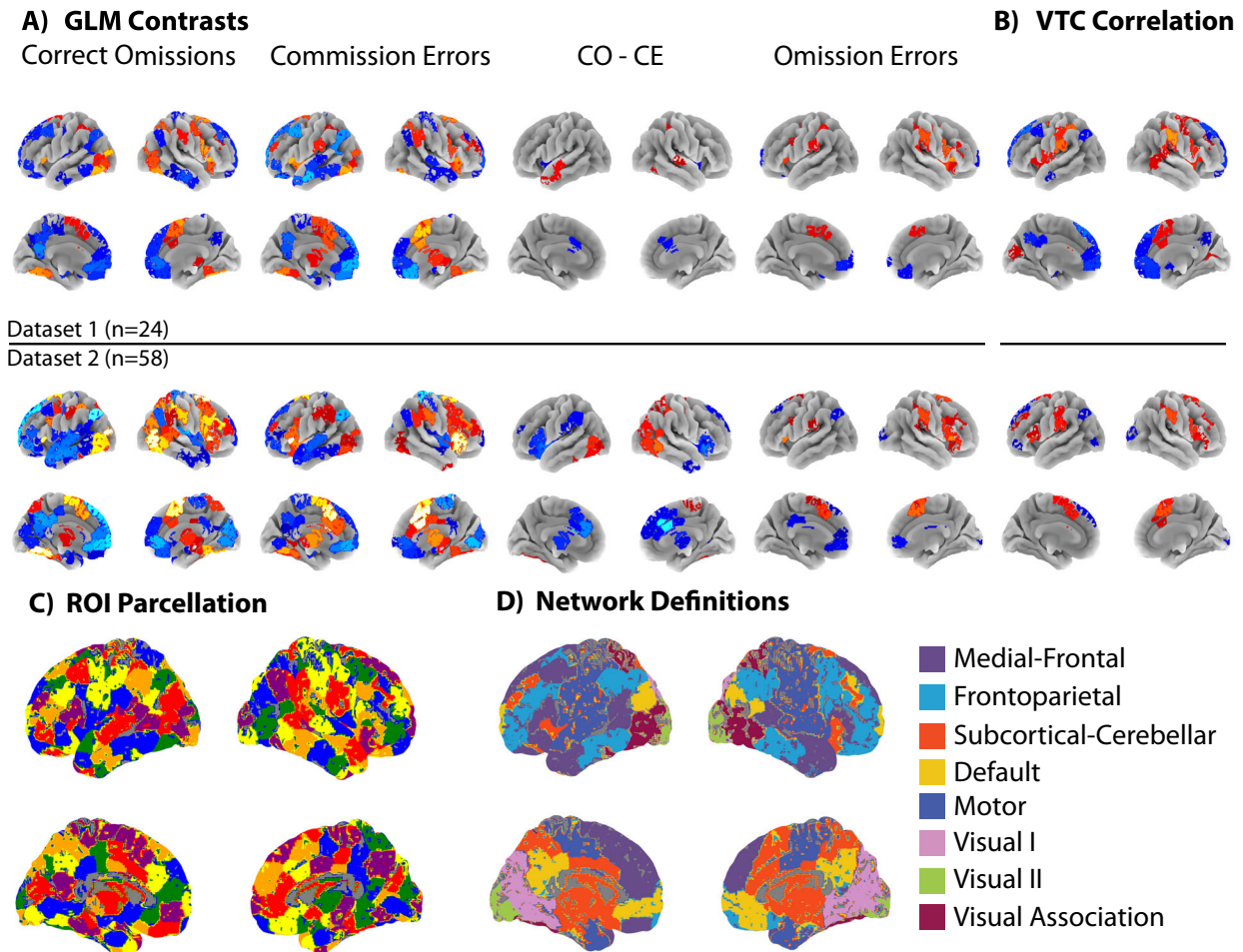
Figure 1. ROI-based replication of activation results. **A**, ROI-based activation contrasts for trial-type events. Each column reflects one of the four main event-based contrasts of Fortenbaugh et al. (2018) or (**B**) the correlation between the VTC and ROI time series. Each row reflects the dataset-specific second levels, which were thresholded using a max-T permutation testing approach (see Materials and Methods, Second-level significance testing via permutation). Note that because these analyses were conducted at the ROI level and then surface projected, there are shared contours to the activation sites. **C**, Shen atlas ROIs surface projected with high-contrast colors. Each ROI was mapped onto one of six highly contrasting colors around the color wheel (blue, green, yellow, orange, red, purple) in a circularly repeating manner (e.g., regions 1 and 7 are blue, regions 2 and 8 are green, and so on). Each vertex was matched with the label of the nearest voxel. **D**, Eight networks surface projected following a similar procedure to **C**. Colors were chosen to match that of Scheinost et al. (2016).

Edge-GLM reveals significant fluctuations in connectivity across networks

Associations between ROI activity and sustained attention performance generally replicated previous associations between in-the-zone performance and increased DMN and decreased frontoparietal activity while providing tentative support for the idea that *too much* DMN activity or *too little* frontoparietal activity can cause lapses. Errors caused by these lapses are accompanied by greater activity in regions associated with error processing, and rare targets in general are associated with more task-positive and less DMN activity.

Together, these findings and the body of work they replicate motivate the prediction that changes in interactions within and between *networks* of regions—not just individual regions themselves—vary with attentional performance. This hypothesis is not directly testable with traditional dynamic FC analyses that rely on measuring region interactions in windows of BOLD-signal time series. Thus, we applied GLMs to edge time series to assess relationships between attention task events and performance at the level of region pairs.

To do so, we applied the same analyses described previously to edge rather than BOLD-signal time series. We asked whether edge time series reliably covaried with task or cognitive events and assessed five contrasts reflecting attention lapses/inhibitory control failures (CE vs baseline), successful inhibition (CO vs baseline and CO − CE), failures to respond (OE vs baseline), and rare target processing (CO + CE vs baseline) using traditional trial-type regressors (Figs. 2, 3).

We first assessed edges that consistently deflected with COs. Activation results revealed that COs were associated with more frontoparietal and lateral visual cortical activity. The edge-based GLM revealed edges associated with these networks and others. Max-T permutation testing identified two reliable edges (where "reliable" refers to significant deflection in both datasets; probability of overlap given the number significant edges in each dataset is $p = 0.004$), both of which were negative deflections between the motor and visual association networks. NBS permutation testing identified 734 edges common to both datasets ($p = 1/10,001$). These consisted of negative deflections between the motor and visual and subcortical-cerebellar networks but positive deflections between motor and medial frontal, frontoparietal, and DMNs. Regions in these higher-order association networks also showed positive deflections with visual networks in response

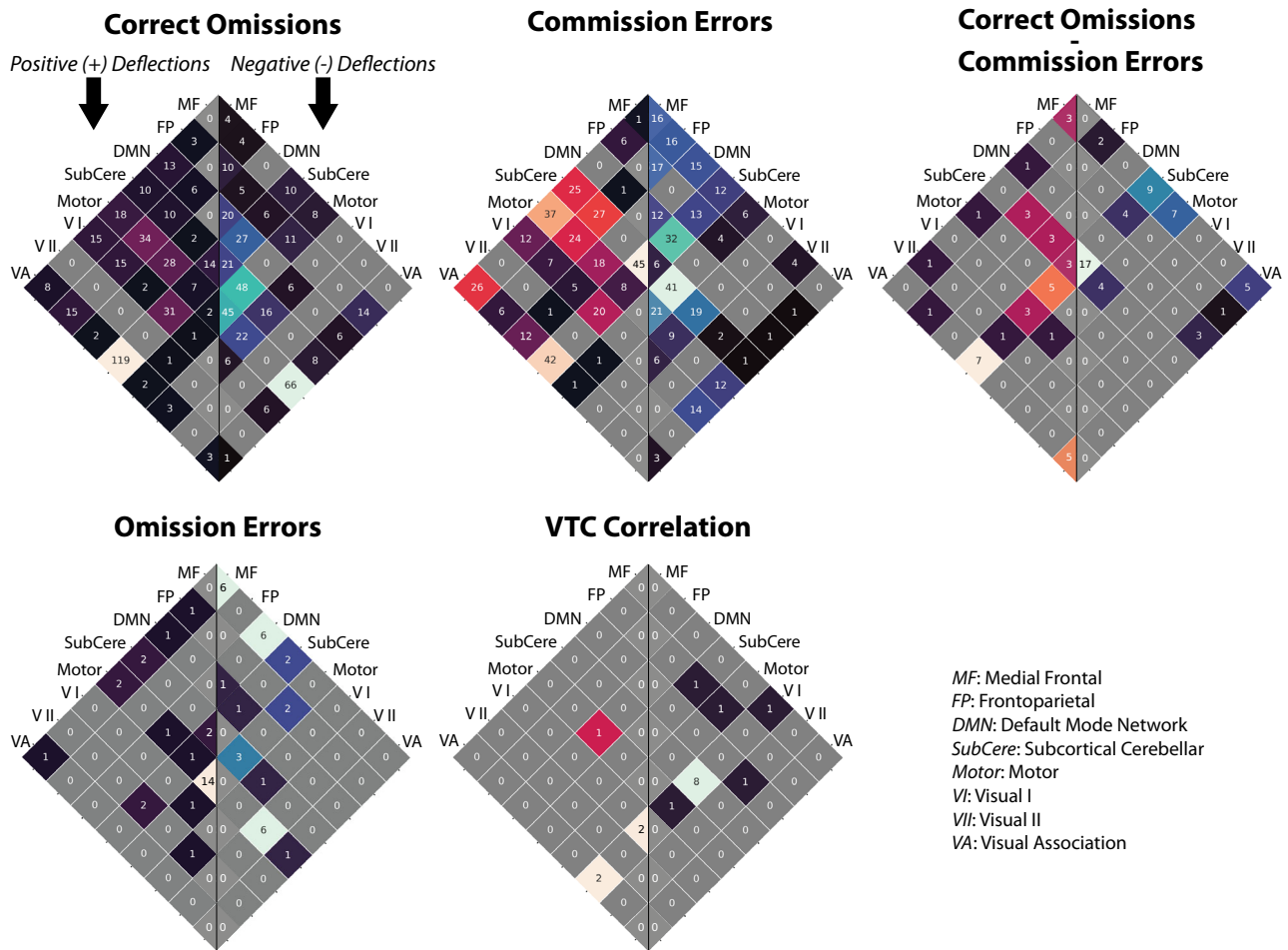## Between Dataset Overlap using NBS Permutation Testing

Figure 2. Cross-network distribution of reliable (i.e., significant in both datasets) edges from trial-type contrasts, based on second-levels computing using NBS permutation testing. Each of the first four matrices reflects the overlap of significant edges for a given contrast: COs, CEs, rare probe trials (CO + CE), and successful inhibitory control (CO − CE). The final matrix reflects edges that were positively or negatively correlated with the VTC. Each heat map shows the number of edges in a given set of network-network connections that were significant for a given analysis. The left side of each heatmap shows the number of edges with positive deflections, while the right side of each heatmap shows the number of edges with negative deflections.

to COs. Negative deflections within the medial frontal, frontoparietal, default mode, and motor networks—and a similar number of positive and negative deflections within the subcortical-cerebellar network—also accompanied COs. Thus, COs tended to be accompanied by increased cofluctuation between higher-order association networks and motor and visual regions and decreased cofluctuation between motor, visual, and subcortical-cerebellar networks. This may reflect attention capture by rare target stimuli and/or the engagement of top-down attentional control leading to successful inhibition.

Unsuccessful inhibition (CE against baseline) identified 0 overlapping edges using the max-T approach and 608 overlapping edges ($p = 1/10,001$) using the NBS approach. As in the activation-based results, the overall pattern was similar to that of COs, with some differences. For example, more edges within the medial frontal and frontal parietal networks were negatively deflected, whereas more edges within the subcortical-cerebellar network were positively deflected. The medial frontal network also showed more negative deflections with the frontoparietal and DMNs and positive deflections with the motor and visual networks.

We isolated successful inhibitory control and sustained attention by comparing COs against CEs (CO − CE). Activation analyses found increased activity in subcortical, cerebellar, visual, and inferior parietal regions in response to successful inhibition but increased activation in control and error processing related regions in response to unsuccessful inhibition. Max-T permutation testing on edges did not identify overlap between the datasets. NBS permutation testing identified 90 overlapping edges ($p = 1/10,001$), which primarily involved regions of the subcortical-cerebellar and medial frontal networks. Successful inhibition was associated with more positive deflections within medial frontal and visual association networks and more negative deflections within the subcortical-cerebellar network. Successful inhibition was also associated with more negative deflections between the medial frontal network and the frontoparietal, subcortical-cerebellar, motor, and visual association networks and more positive deflections between the subcortical-cerebellar network and every other network except for the medial frontal network. Speculatively, this pattern of results may suggest that greater medial frontal network segregation (more positive within- and negative between-network cofluctuation) and
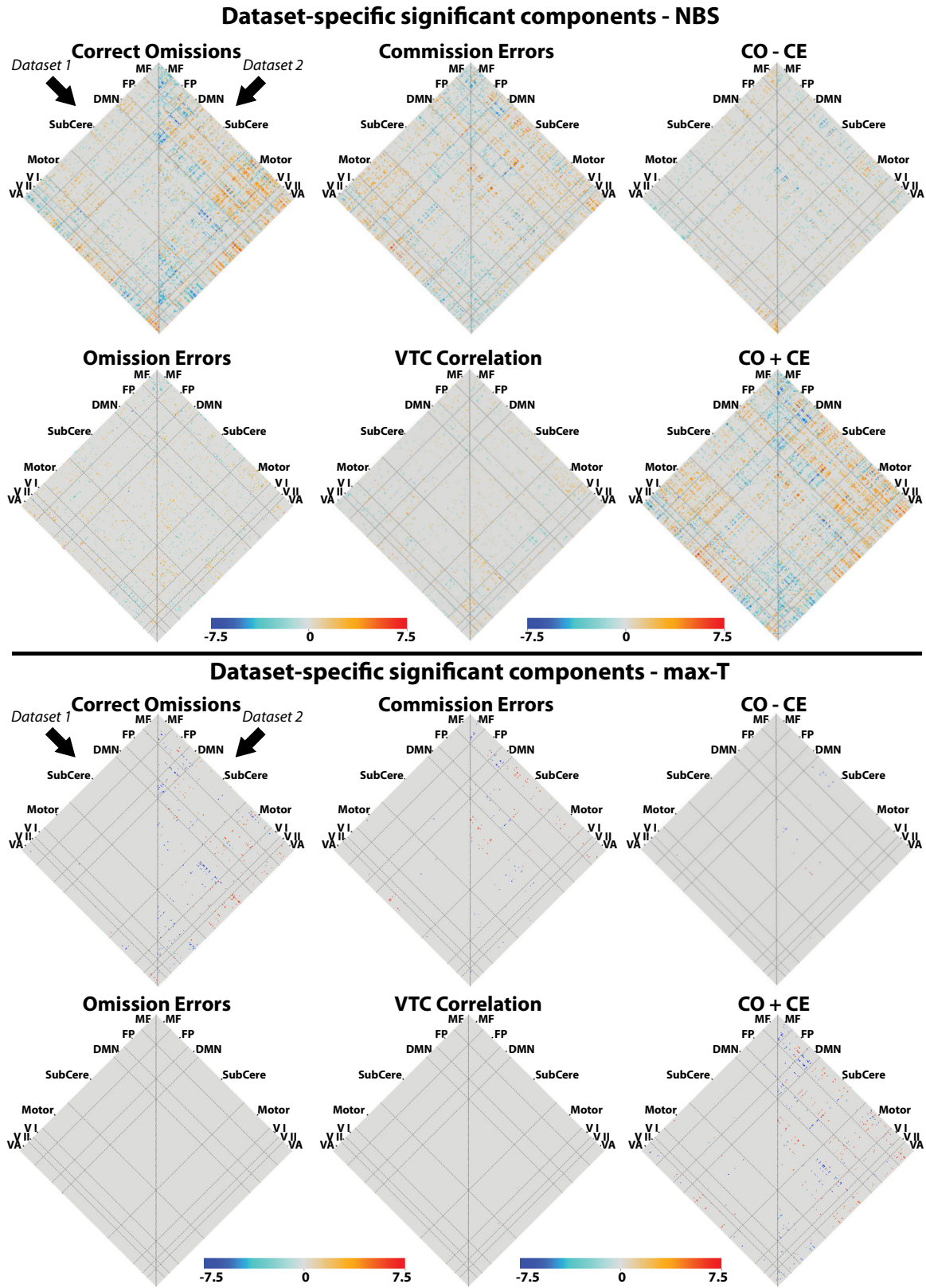
Figure 3. Dataset-specific second-level results for each of the five contrasts and the correlation with the VTC. Within each results matrix, the left half reflects dataset 1 and the right half reflects dataset 2. Colors reflect thresholded *t* scores, with blue indicating negative values and red indicating positive values. The top set reflects component-level thresholding using NBS permutation testing while the bottom reflects edge-level thresholding using max-T permutation testing.

greater subcortical-cerebellar network integration (more negative within- and positive between-network cofluctuation) accompany successful relative to unsuccessful inhibition.

We contrasted OEs (incorrectly withholding responses to common city trials) against baseline. Activation patterns found increased activity in the middle frontal gyrus, inferior parietal

lobule, and insula and decreased activity in the ventromedial pre-frontal cortex. While max-T–based permutation testing did not identify any overlapping edges, NBS-based permutation testing identified 58 overlapping edges ($p = 1/10,001$). The edges consisted mostly of positive deflections within the motor network and negative deflections distributed across the brain networks. Most of the negatively deflected edges were associated with the medial frontal network, both within-network and with the DMN and subcortical-cerebellar network or were associated with connections between the motor network and the subcortical-cerebellar and the visual I network.

Lastly, we examined rare trials against baseline (CO + CE). Max-T identified seven overlapping edges ($p = 1/10,001$). These consisted of negative deflections between the motor network and both the visual association and subcortical-cerebellar networks, along with positive deflections between the subcortical-cerebellar network and both high-level vision networks and the frontoparietal network. NBS permutation testing identified 1,252 overlapping edges ($p = 1/10,001$), with patterns of deflections roughly combining those seen separately in the CO against baseline and CE against baseline contrasts.

Edges covary with continuous fluctuations in attentional states across time

We investigated whether edge fluctuations track continuous fluctuations in attentional states in two ways. For each, we began by computing the VTC, a measure of attentional focus (Esterman et al., 2013; Fortenbaugh et al., 2018). In the first analysis, we replicated Fortenbaugh et al. (2018)'s analysis by correlating each edge time series with the VTC, shifted by 6 s to account for hemodynamic lag. We transformed the correlations into z scores via Fisher's transformation and applied our second-level and between-group analyses. While max-T permutation testing did not identify an overlapping set of edges, NBS-based permutation testing identified a significant overlapping set in both the correlation approach ($n = 18$; $p = 0.002$). Most correlations with the VTC were with edges within and between the visual I network. Edges within this network and with the visual association network were positively correlated with VTC values, while edges with the subcortical-cerebellar network were negatively correlated. There was some evidence of the involvement of higher-order association regions: a DMN-motor edge was positively correlated with the VTC while two frontoparietal-motor and frontoparietal-subcortical-cerebellar edges were negatively correlated. This suggests that, as attention waned, cofluctuations between frontoparietal, vision, and motor production networks became increasingly negative, while cofluctuations within vision networks and with the DMN became increasingly positive. In contrast, increases in attention are associated with increased FC between frontoparietal, vision, and motor production networks. The second planned parametric regressor analysis did not identify any significant edges using max-T or any significantly large components using NBS in either dataset.

Edges may not vary preceding successful versus unsuccessful responses

We examined edge time series preceding COs and CEs in response to rare mountain trials (Fortenbaugh et al., 2018). We averaged the preceding 4.6 s before each CO and CE trial and took the difference between trial types for each participant. Max-T–based permutation testing did not identify significant differences in either dataset. NBS-based permutation testing also did not identify any significantly different subcomponent

in dataset 1 but did identify a subcomponent containing 621 edges in dataset 2. These included both positive and negative deflections within the subcortical-cerebellar and motor networks as well as between these networks and higher-order association and visual networks.

Edges correlated with attentional fluctuations support individual performance predictions

We assessed whether edges that were significantly deflected in three key analyses isolating attentional control overlapped with edges previously identified as underlying individual differences in sustained attention and attentional control using CPM (Rosenberg et al., 2016). We selected the NBS-thresholded second levels of (1) the CO − CE contrast, (2) the CO − CE precursor analysis, and (3) correlations with the VTC. Because the CPM edges were originally identified in dataset 1, we focus on overlap with second-level results in dataset 2.

Neither of the CO − CE analyses (contrast and precursor analyses) identified a significantly overlapping set of edges (contrast, $n = 32$, $p = 0.28$; precursor, $n = 16$, $p = 0.17$). In contrast, there was significant overlap between the edges that correlated with the VTC and the attention networks identified via CPM ($n = 55$, $p = 1/10,001$; Fig. 4). This overlap analysis was built on the assumption the HAN edges should map onto edges negatively correlated with the VTC and LAN edges should map onto edges positively correlated with the VTC (see Materials and Methods, Assessing dataset-level overlap with edges underlying individual differences in attentional control). Only edges matching these mappings were considered during the permutation test. Of the 58 edges identified as overlapping between the VTC second-level and the CPM edges, 55 matched the mapping. A chi-squared test confirmed that the two discretizations (HAN/LAN and positive/negative correlations) were dependent ($\chi^2_{Yates} = 42.68$; $p = 6 \times 10^{-11}$), supporting our assumption. HAN/negatively correlated edges were largely distributed across edges between frontal networks and motor/visual networks (Fig. 4). In contrast, LAN/positively correlated edges were distributed across subcortical-cerebellar, motor, and visual association networks.

Edge fluctuations can be systematically mispredicted by ROI activation

To test whether the information captured by our edge time series analysis was redundant with information from traditional ROI activation analyses, we fit the ROI time series with the contrast models introduced above. We then converted the first-level results to predicted edge results by multiplying the t scores for each pair of ROIs and taking the signed square root (see Materials and Methods, Comparing observed edge results to ROI-based predictions). We applied our second- and third-level comparisons to these predicted edge values and identified edges *predicted* to be nonsignificant in either dataset but in reality significant in both datasets.

Across the five event-based contrasts within our NBS-based second-levels, 19–33% of all edges that were significant in both datasets were not predicted to be significant in either dataset based on their activation alone. This is not because the ROI-based predictions produced more sparse maps than the observed edge maps. In each dataset and in each of the five contrasts, as many or more edges were identified as significant when predicted from ROI activation maps compared with being actually observed. With the max-T permutation paradigm, this increase ranged from 0% (an equal number of edges were
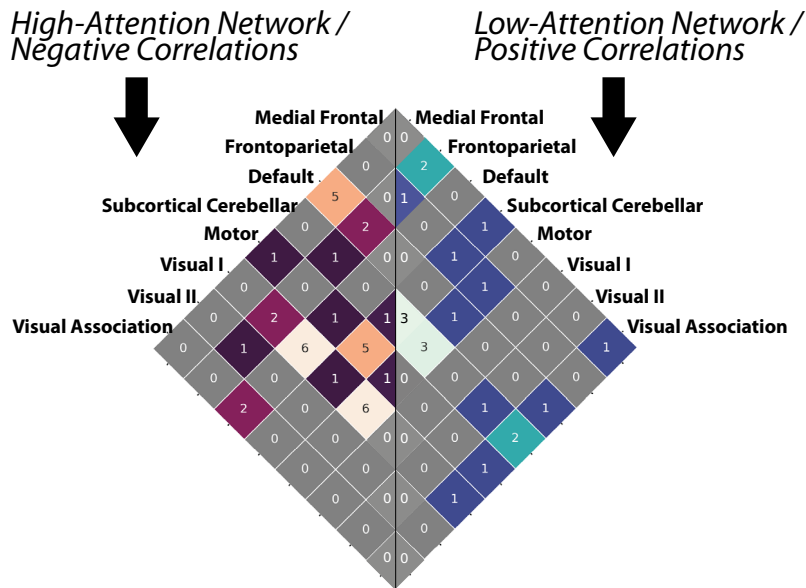
Dataset 2 VTC correlations overlap with CPM edges

Figure 4. Overlap of edges which significantly correlated with the VTC in dataset 2 and which contributed to predictions of individual differences in performance in dataset 1. Each cell of the matrix represents the number of significant edges between two networks (or within a network, if lying along the diagonal). The left half reflects the count of HAN edges that overlapped with edges that were significantly anticorrelated with the VTC. The right half reflects the count of LAN edges that were significantly positively correlated with the VTC.

identified as significant) to 4,317% (17 significant edges observed vs 751 predicted) in dataset 1 and from 772% (22 vs 192) to 1,361% (359 vs 5,248) in dataset 2. The increase was smaller when using the NBS permutation paradigm but still substantial (∼100–200%). Thus, edge-based GLMs are related to but not redundant with activation analyses.

Edge responses are similar to traditional HRFs
The above results indicate that we are able to capture changes in edge time series using models assuming the canonical HRF. However, it is possible that not all edge responses are described by the HRF, meaning that model designs may not be optimized. To test this, we estimated ROI and edge responses to rare trials versus baseline using FIR modeling of data from run 1 of dataset 1 (see Materials and Methods, Estimating HRF of edge time series).

Nine clusters were identified for succinctly capturing the ROI HRFs produced by our contrast (Fig. 5). Six of the centroids, accounting for 165 ROIs (61.6% of all ROIs), were shaped as the canonical HRF and varied mostly in amplitude and timing (peaking between 4 and 6 s). The remaining three centroids, accounting for 103 ROIs (38.4%), were shaped similarly to the canonical HRF but decayed slowly after peaking, rather than dropping roughly in proportion to the activation rate. Two of these three centroids, accounting for 84 ROIs, had a somewhat late peak (8 and 10 s). ROI response functions similar to these centroids would likely be partially matched by the canonical HRF, especially when modeled with its temporal derivative as an additional regressor, so we may assume that the canonical HRF is roughly appropriate for all ROIs.

Nine clusters were also identified for succinctly capturing the edge HRFs produced by our contrast (Fig. 5). Four of these, accounting for 15,923 edges (44.5% of all edges), appeared to match the canonical HRF with slight shifts in amplitude and timing (peaking between 2 and 9 s). Three of the remaining five

centroids, accounting for 13,289 edges (37.1%), were also shaped similar to the canonical HRF but with larger lags in their timing (peaking between 14 and 27 s). The final two centroids were not of the canonical HRF shape. One centroid, accounting for 4,110 edges (11.5%), contained two distinct peaks, one at 5 s, and one at 27 s. Edge response functions similar to this centroid would likely be partially matched by the canonical HRF, which peaks ∼5 s and returns to baseline ∼20 s after stimulus onset. The final centroid, accounting for 2,456 edges (6.9%), slowly rose, peaking ∼18 s, and then slowly fell. Thus, the canonical HRF, along with its temporal derivative, likely provided a strong-to-partial match for 44.5–56% of edges.

To assess the reliability of these results, we repeated the above analyses using run 2 data from dataset 1. An elbow of 10 clusters was identified for the ROI HRFs, while nine clusters were again identified for the edge HRFs. We assessed whether the cluster labels were significantly similar between runs by computing the AMI (Vinh et al., 2010). We assessed the significance of the mutual information by permuting the cluster labels 10,000 times and recomputing the AMI. Both clusterings were significantly similar across runs (ROI-AMI = 0.223, $p = 1/10,001$; edge-AMI = 0.004, $p = 1/10,001$). We also asked whether these cluster labels were related to network definitions by computing the AMI between the ROI clusters and eight canonical network labels (Finn et al., 2015) and between the edge clusters and the network-network labels. For both runs, the clusterings were significantly similar to the network definitions (ROIs: both AMIs > 0.075, ps = $1/10,001$; edges: both AMIs > 0.009, ps = $1/10,001$).

It is important to note that, though the canonical HRF is only appropriate for a smaller subset of edges compared with ROIs, this should only produce false negatives among edges with non-canonical response functions, not false positives. This suggests that the number of edges that were not predicted from univariate activity above may be an underestimate.
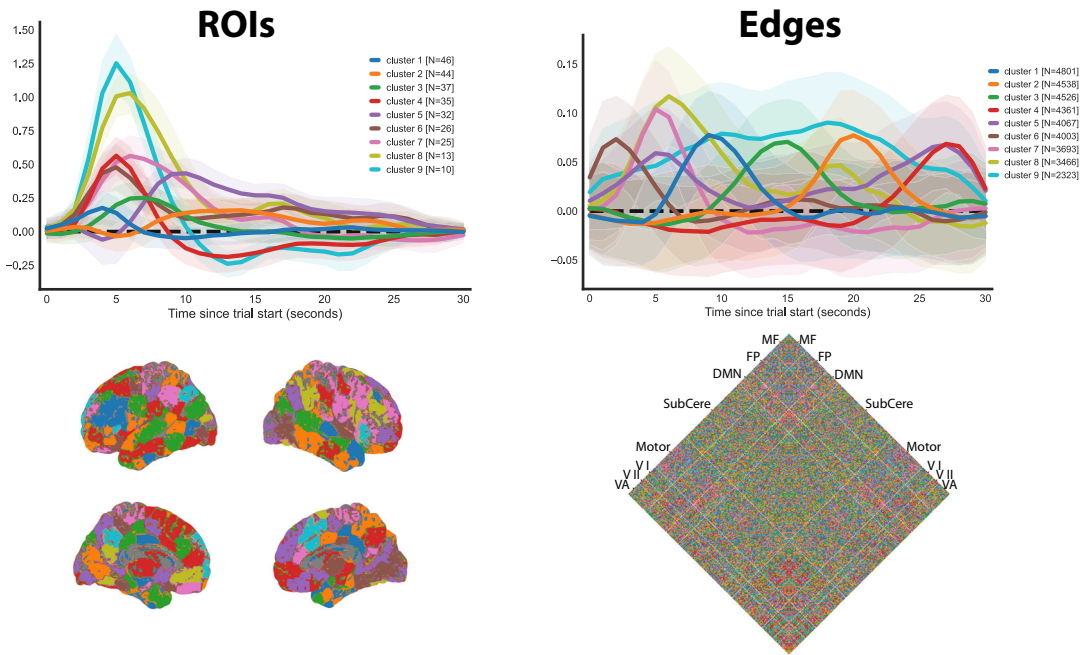
Estimated Hemodynamic Response Functions

Figure 5. Estimated HRFs from univariate ROI activity and edge time series. The responses for the 30 s following rare probe trials were estimated for each ROI and edge using FIR modeling. Fits were performed for the first session of each subject in dataset 1 and averaged across subjects before being clustered using *k*-means. Top, The centroid (mean) of nine clusters are presented for each time series type, ordered by cluster size. Shaded regions reflect the bootstrapped 95% confidence intervals of the means. Bottom, Visualizations of the clusters.

Discussion

We examined whether traditional fMRI GLM analyses can be applied to edge cofluctuation time series to capture moment-to-moment fluctuations in attention. We assessed the replicability of results across two independent datasets in which participants performed the gradCPT (Rosenberg et al., 2016; Yoo et al., 2022), which enables examination of attention both at the event level and via continuous measures across time. In general, activation results replicated previous empirical findings (Esterman et al., 2013; Fortenbaugh et al., 2018). Cofluctuation results extended the activation results by identifying new sets of edges that covary with attention and overlap with networks that predict individual differences in sustained attention. To assess the robustness of previous activation-based results and contextualize our edge-based results, we replicated the univariate activity analyses of Fortenbaugh et al. (2018). Contrasts against baseline all produced results aligning with Fortenbaugh et al. (2018). Correctly withheld responses to rare probe trials (COs), CEs, and both together were associated with increased activity in regions associated with vigilant attention and decreased activity in the DMN. They were also associated with increased lateral visual cortex activity and decreased medial visual cortex activity. This aligns well with a meta-analysis of 11 Go/No-go fMRI studies contrasting No-go and Go trials which implicated the bilateral putamen, middle occipital gyrus, insula, inferior parietal lobule, pre-SMA, middle frontal gyrus, the right superior and inferior frontal gyri, and the left fusiform gyrus (Simmonds et al., 2008). Erroneous withholding on frequent nonprobe trials (OEs) was also associated with regions connected to error processing including the inferior parietal lobule, insula, middle frontal gyrus, and ventromedial prefrontal cortex.

Replicating results of Fortenbaugh et al. (2018), successful withholding of responses (CO − CE) was associated with greater activity in subcortical, cerebellar, visual, and inferior parietal areas and lesser activity in the medial prefrontal cortex, dorsal anterior cingulate cortex, and insula activity in dataset 2 (but not dataset 1). Prior work found increased cuneus, caudate, and inferior frontal activity for successful stopping and increasing anterior cingulate and insula activity with unsuccessful stopping (Chevrier et al., 2007), both of which have previously been associated with error monitoring and adjustments of control (Ramautar et al., 2006; Shenhav et al., 2016; Dali et al., 2023). Our CO − CE contrast then likely captures regions contributing to appropriate motor execution or withholding (subcortical, cerebellar, and inferior parietal areas) during successful inhibition and error processing and control adjustments (medial PFC, dACC, insula) following unsuccessful inhibition.

Results also replicated associations between continuous fluctuations of attention and ROI activity, revealing negative correlations between the VTC and DMN activity and positive correlations between the VTC and task-positive frontal and parietal region activity. Results did not fully replicate pretrial differences between COs and CEs. There were no significant pretrial activity differences in dataset 1, though in dataset 2 the patterns of activity aligned with previous work suggesting that lapses occur when frontoparietal activity is too low or DMN activity is too high (Esterman et al., 2013). Overall, activation contrasts replicated previous work and identified regions implicated in visual processing, motor execution, error processing, control adjustments and top-down attentional modulation.

The only two univariate analyses that did not qualitatively replicate prior results in both datasets were the CO − CE and pre-CO versus pre-CE contrasts. This is perhaps not surprising given the rarity of mountain trials (10%) and the fact that these are difference contrasts. In addition, the trial precursor results have smaller effect sizes than the standard contrast approach

(compare the *t* values of Figs. 4D and 5B in Fortenbaugh et al., 2018). Thus, we may be underpowered for detecting the effects.

While the univariate results can be interpreted in terms of networks, they cannot capture changes in network connectivity that occur without a change in overall activation. To address this limitation, we applied the above fMRI analyses to edge time series to ask whether changes in the cofluctuation of pairs of regions or networks are involved in changes in sustained attention. These edge-based results both aligned with and expanded upon the univariate activation results (Fig. 2). Contrasts with sparse univariate results also produced sparse edge results. In addition, though not captured by this overlap analysis, dataset 1 always produced sparser edge results than dataset 2 (Fig. 3), as it did for the univariate results. These reliable edges were distributed across the brain at the network level, though this distribution was not uniform. Many of the edges connected networks identified by the activation contrasts, specifically the subcortical-cerebellar, frontal, and DMNs. The edges within and between these networks were variably deflected, both matching with predictions (e.g., positive deflections within the medial frontal network for COs vs CEs) and extending beyond the previous literature (e.g., negative deflections between medial frontal and motor networks in the same contrast). Deflections between frontal and subcortical-cerebellar regions and regions in the motor and vision networks emphasize the integration of processing from sensory input all the way to response execution during successfully sustained attention and point to the need to examine network connectivity more directly.

Edges that correlated with VTC produced striking results. First, reliable edges were only found in networks associated with sensory processing and motor production, namely, the visual, motor, and subcortical-cerebellar networks (Fig. 2), rather than association networks such as the frontal networks and the DMN. In addition, those significantly correlated edges in dataset 2 also significantly overlapped with edges identified using CPM in dataset 1 to identify individual differences in gradCPT performance (Fig. 4). A few features of this result are worth noting. First, the edge approach identifies edges that consistently vary with the VTC. In contrast, the CPM approach relies on variance across the group. Although this did not need to be the case, the same edges that track attention fluctuations at the group level also scale with individual differences in overall performance. Second, though the edges which were reliable across datasets were identified in sensory and motor production networks, the edges overlapping between dataset 2 and the CPM edges also included many connections with the medial frontal and frontoparietal networks. This suggests again that dataset 1 may have been underpowered and failed to capture deflections in these frontal networks, previously implicated in sustained attention (Fortenbaugh et al., 2017).

We found that 19–33% of the reliable edges identified via contrast analyses were not predicted from univariate ROI activity, despite ROI activity predicting deflections in more edges for each contrast than were actually identified. Therefore, these edge analyses are extracting reliable and *novel* information about rapid network reconfigurations underlying fluctuations in cognitive processes. Another way to frame this result is that our approach can capture edge deflections driven by task- or event-evoked ROI activity and/or evoked changes in cofluctuation independent of task activation. To many researchers, both of these reflect meaningful neural responses that should be investigated and incorporated into theories relating brain activity to cognition. At the same time, other work has attempted to isolate the interaction between FC and cognitive events, termed psychophysiological interactions (PPIs, Friston et al., 1997). PPI studies how FC changes in a given region by modeling its response to a task, its relationship with a given seed region, and the interaction of that seed region with the task (O'Reilly et al., 2012). While PPI is well powered for detecting interaction in block designs, it is generally underpowered in event designs (Cisler et al., 2014). In contrast, this work isolates edge deflections driven by task events. Future work may compare this edge-based approach for examining event-based interactions to PPI and other methods such as the betas series method (Cisler et al., 2014).

Though this work identified reliable and novel edges while making minimal assumptions about how edge cofluctuations may differ from univariate activity, there are many options for optimizing this workflow and increasing the power of these analyses. First, we found that the HRF is appropriate for only a subset of edges. Other work has also used a canonical HRF function to examine edge time series "events," high amplitude moments that explain a large amount of variance in the connectomes between individuals, and found moderate correlations between a convolved movie-event time series and edge events time series (Tanner et al., 2023). Future work can compare different response functions, including agnostic finite impulse regressors, as well as smoothed time series to see which produces the most reliable results. Using more flexible response functions would also alleviate the issues produced by network-level differences in response functions, which we also saw in this work.

This work also compared max-T and NBS thresholding approaches for identifying significantly deflected edges. Max-T was quite conservative, identifying few edges consistent across datasets. In contrast, NBS identified a reliable set of edges for every contrast where a significant subcomponent was identified in each dataset. However, the overlap found between datasets was relatively small, suggesting that NBS may not be conservative enough. At best, a significant NBS component can be interpreted as containing at least one significant edge (Rosenblatt et al., 2018; Sassenhagen and Draschkow, 2019). We chose a significance threshold of 0.01, in keeping with the creators of NBS (Zalesky et al., 2010; Serin et al., 2021), but this may still produce many false positives. Another work examining cluster correction methods in activation analyses has suggested a threshold of 0.001 (Woo et al., 2014). Methods which reduce the possibility of false negatives *and* false positives when working with edges will be highly valuable.

We circumvented these issues by focusing on edges that were members of the significant NBS components in both datasets. This required that a given edge survive thresholding ($p < 0.01$) and be adjacent to enough other significant edges, in both datasets, effectively producing a final threshold more stringent than $p < 0.0001$. Other research groups could replicate this approach either by using open data which matches the task they are studying or through a split-half approach within their dataset (provided it is large enough). Alternatively, other clustering methods may reduce false positives within a significant component (e.g., the degree based statistic; Yoo et al., 2017).

We have shown that a simple modification to traditional fMRI analyses, converting ROI activity into edge cofluctuation time series, allowed us to extract novel and reliable information in the form of high-frequency edge deflections. One of the major benefits of this approach is its ability to produce windowless estimates of network strength in order to identify rapid network reconfigurations. In some ways, this is comparable to change point detection (Xu and Lindquist, 2015) and latent state

analyses, which identify a set of brain states based on ROI activation and covariance that can be assigned to each timepoint in an fMRI time series. These analyses are in a sense windowless as they can identify shifts between brain states that happen from one moment to the next. Such analyses have identified large-scale network changes tied to changes in attention (Cai et al., 2021; Kondo et al., 2022; Yamashita et al., 2021; Song et al., 2022). We believe this edge-based approach is complementary to latent state analyses in two ways. First, it can easily be combined with temporal regressors to extract information from edge cofluctuations. In contrast, latent state models are typically entirely data driven, and states must be linked to temporal events after the fact. Second, the edge-based approach operates at a higher-dimensional description than whole-brain state changes, which describe coarse, low dimensional changes.

This edge-based approach enables a finer temporal granularity than previous network neuroscientific fMRI analyses. As a result, this method can be applied to any research which seeks to apply the network neuroscience perspective to relatively fast-changing cognitive phenomena. For example, how do networks reconfigure to form "process-specific alliances" when engaging different components of processes during learning (Cabeza and Moscovitch, 2013)? How do networks reconfigure in response to prediction errors during reinforcement learning paradigms? Or, how do networks change in response to increasing working memory loads, especially at supracapacity set sizes? We hope that this work, along with the openly available code and data, will allow researchers to identify rapid network reconfigurations underlying a diverse set of cognitive processes.

This edge-based approach has offered an expanded perspective on relatively rapid network reconfigurations during attentional fluctuations. While the default mode and attentional networks were implicated in lapses and successful attentional control as expected, so were subcortical and cerebellar regions, which are linked to both associative and sensory-motor regions. In addition, there is now the possibility to investigate attentional fluctuations in terms of the relative integration or segregation of a given network (such as the medial frontal network) with respect to the rest of the brain. This expanded perspective can allow us to identify both new connections and new graphical motifs underlying differences in attention across time.

Data Availability

The code for this analysis is available at github.com/henrymj/edge_GLM. Dataset 2 can be accessed at https://nda.nih.gov/edit_collection.html?id=2402. For inquiries about dataset 1, contact the authors of the original manuscript https://doi.org/10.1038/nn.4179.

References

Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124:1–38.

Buckner RL, Dinicola LM (2019) The brain's default network: updated anatomy, physiology and evolving insights. Nat Rev Neurosci 20:593–608.

Cabeza R, Moscovitch M (2013) Memory systems, processing modes, and components: functional neuroimaging evidence. Perspect Psychol Sci 8:49–55.

Cai W, Warren SL, Duberg K, Pennington B, Hinshaw SP, Menon V (2021) Latent brain state dynamics distinguish behavioral variability, impaired decision-making, and inattention. Mol Psychiatry 26:4944–4957.

Chevrier AD, Noseworthy MD, Schachar R (2007) Dissociation of response inhibition and performance monitoring in the stop signal task using event-related fMRI. Hum Brain Mapp 28:1347–1358.

Cisler JM, Bush K, Steele JS (2014) A comparison of statistical methods for detecting context-modulated functional connectivity in fMRI. Neuroimage 84:1042–1052.

Dali G, Brosnan M, Tiego J, Johnson BP, Fornito A, Bellgrove MA, Hester R (2023) Examining the neural correlates of error awareness in a large fMRI study. Cereb Cortex 33:458–468.

Esterman M, Noonan SK, Rosenberg M, Degutis J (2013) In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. Cereb Cortex 23:2712–2723.

Esterman M, Rosenberg MD, Noonan SK (2014) Intrinsic fluctuations in sustained attention and distractor processing. J Neurosci 34:1724–1730.

Faskowitz J, Esfahlani FZ, Jo Y, Sporns O, Betzel RF (2020) Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. Nat Neurosci 23:1644–1654.

Fiebelkorn IC, Kastner S (2019) A rhythmic theory of attention. Trends Cogn Sci 23:87–101.

Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT (2015) Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosci 18:1664–1671.

Fortenbaugh FC, DeGutis J, Esterman M (2017) Recent theoretical, neural, and clinical advances in sustained attention research. Ann N Y Acad Sci 1396:70–91.

Fortenbaugh FC, DeGutis J, Germine L, Wilmer JB, Grosso M, Russo K, Esterman M (2015) Sustained attention across the life span in a sample of 10,000: dissociating ability and strategy. Psychol Sci 26:1497–1510.

Fortenbaugh FC, Rothlein D, McGlinchey R, DeGutis J, Esterman M (2018) Tracking behavioral and neural fluctuations during sustained attention: a robust replication and extension. NeuroImage 171:148–164.

Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997) Psychophysiological and modulatory interactions in neuroimaging. Neuroimage 6:218–229.

Jayakumar M, Balusu C, Aly M (2023) Attentional fluctuations and the temporal organization of memory. Cognition 235:105408.

Joshi A, Scheinost D, Okuda H, Belhachemi D, Murphy I, Staib LH, Papademetris X (2011) Unified framework for development, deployment and robust testing of neuroimaging algorithms. Neuroinformatics 9:69–84.

Jun J, Lee VG (2021) Perceptual and response factors in the gradual onset continuous performance tasks. Atten Percept Psychophys 83:3008–3023.

Kardan O, et al. (2022) Differences in the functional brain architecture of sustained attention and working memory in youth and adults. PLoS Biol 20:e3001938.

Kessler D, Angstadt M, Sripada C (2016) Growth charting of brain connectivity networks and the identification of attention impairment in youth. JAMA Psychiatry 73:481–489.

Kondo HM, Terashima H, Ezaki T, Kochiyama T, Kihara K, Kawahara JI (2022) Dynamic transitions between brain states predict auditory attentional fluctuations. Front Neurosci 16:816735.

Kucyi A, Esterman M, Capella J, Green A, Uchida M, Biederman J, Gabrieli JDE, Valera EM, Whitfield-Gabrieli S (2021) Prediction of stimulus-independent and task-unrelated thought from functional brain networks. Nat Commun 12:1793.

Kucyi A, Esterman M, Riley CS, Valera EM (2016) Spontaneous default network activity reflects behavioral variability independent of mind-wandering. Proc Natl Acad Sci U S A 113:13899–13904.

Kucyi A, Hove MJ, Esterman M, Hutchison RM, Valera EM (2017) Dynamic brain network correlates of spontaneous fluctuations in attention. Cereb Cortex 27:1831–1840.

Langner R, Eickhoff SB (2013) Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. Psychol Bull 139:870.

Mackworth NH (1948) The breakdown of vigilance during prolonged visual search. Q J Exp Psychol 1:6–21.

Mason MF, Norton MI, Van Horn JD, Wegner DM, Grafton ST, Macrae CN (2007) Wandering minds: the default network and stimulus-independent thought. Science 315:393–395.

O'Reilly JX, Woolrich MW, Behrens TE, Smith SM, Johansen-Berg H (2012) Tools of the trade: psychophysiological interactions and functional connectivity. Soc Cogn Affect Neurosci 7:604–609.

Pedregosa F, et al. (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830.

Poole VN, Robinson ME, Singleton O, DeGutis J, Milberg WP, McGlinchey RE, Salat DH, Esterman M (2016) Intrinsic functional connectivity predicts individual differences in distractibility. Neuropsychologia 86:176–182.

Ramautar JR, Slagter HA, Kok A, Ridderinkhof KR (2006) Probability effects in the stop-signal paradigm: the insula and the significance of failed inhibition. Brain Res 1105:143–154.

Robertson IH, Manly T, Andrade J, Baddeley BT, Yiend J (1997) Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia 35:747–758.

Rosenberg MD, Finn ES, Scheinost D, Papademetris X, Shen X, Constable RT, Chun MM (2016) A neuromarker of sustained attention from whole-brain functional connectivity. Nat Neurosci 19:165–171.

Rosenberg M, Noonan S, DeGutis J, Esterman M (2013) Sustaining visual attention in the face of distraction: a novel gradual-onset continuous performance task. Atten Percept Psychophys 75:426–439.

Rosenberg MD, Scheinost D, Greene AS, Avery EW, Kwon YH, Finn ES, Ramani R, Qiu M, Constable RT, Chun MM (2020) Functional connectivity predicts changes in attention observed across minutes, days, and months. Proc Natl Acad Sci U S A 117:3797–3807.

Rosenblatt JD, Finos L, Weeda WD, Solari A, Goeman JJ (2018) All-resolutions inference for brain imaging. Neuroimage 181:786–796.

Sassenhagen J, Draschkow D (2019) Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. Psychophysiology 56:e13335.

Scheinost D, Tokoglu F, Shen X, Finn ES, Noble S, Papademetris X, Constable RT (2016) Fluctuations in global brain activity are associated with changes in whole-brain connectivity of functional networks. IEEE Trans Biomed Eng 63:2540–2549.

Serences JT, Ester EF, Vogel EK, Awh E (2009) Stimulus-specific delay activity in human primary visual cortex. Psychol Sci 20:207–214.

Serin E, Zalesky A, Matory A, Walter H, Kruschwitz JD (2021) NBS-Predict: a prediction-based extension of the network-based statistic. NeuroImage 244:118625.

Shen X, Finn ES, Scheinost D, Rosenberg MD, Chun MM, Papademetris X, Constable RT (2017) Using connectome-based predictive modeling to predict individual behavior from brain connectivity. Nat Protoc 12:506–518.

Shen X, Tokoglu F, Papademetris X, Constable RT (2013) Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. NeuroImage 82:403–415.

Shenhav A, Cohen JD, Botvinick MM (2016) Dorsal anterior cingulate cortex and the value of control. Nat Neurosci 19:1286–1291.

Simmonds DJ, Pekar JJ, Mostofsky SH (2008) Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. Neuropsychologia 46:224–232.

Song H, Shim WM, Rosenberg MD (2022) Large-scale neural dynamics in a shared low-dimensional state space reflect cognitive and attentional dynamics. bioRxiv, 2022-11.

Tanner JC, Faskowitz J, Byrge L, Kennedy DP, Sporns O, Betzel RF (2023) Synchronous high-amplitude co-fluctuations of functional brain networks during movie-watching. Imaging Neuroscience 1:1–21.

Vinh NX, Epps J, Bailey J (2010) Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance. J Mach Learn Res 11:2837–2854.

Waskom ML (2021) Seaborn: statistical data visualization. J Open Source Softw 6:3021.

Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE (2014) Permutation inference for the general linear model. Neuroimage 92:381–397.

Woo CW, Krishnan A, Wager TD (2014) Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. Neuroimage 91:412–419.

Wu EX, Liaw GJ, Goh RZ, Chia TT, Chee AM, Obana T, Rosenberg MD, Yeo BTT, Asplund CL (2020) Overlapping attentional networks yield divergent behavioral predictions across tasks: neuromarkers for diffuse and focused attention? NeuroImage 209:116535.

Xu Y, Lindquist MA (2015) Dynamic connectivity detection: an algorithm for determining functional connectivity change points in fMRI data. Front Neurosci 9:285.

Yamashita A, Rothlein D, Kucyi A, Valera EM, Esterman M (2021) Brain state-based detection of attentional fluctuations and their modulation. Neuroimage 236:118072.

Yoo K, Lee P, Chung MK, Sohn WS, Chung SJ, Na DL, Ju D, Jeong Y (2017) Degree-based statistic and center persistency for brain connectivity analysis. Hum Brain Mapp 38:165–181.

Yoo K, Rosenberg MD, Kwon YH, Lin Q, Avery EW, Sheinost D, Constable RT, Chun MM (2022) A brain-based general measure of attention. Nat Hum Behav 6:782–795.

Zalesky A, Fornito A, Bullmore ET (2010) Network-based statistic: identifying differences in brain networks. Neuroimage 53:1197–1207.

Zamani Esfahlani F, Jo Y, Faskowitz J, Byrge L, Kennedy DP, Sporns O, Betzel RF (2020) High-amplitude cofluctuations in cortical activity drive functional connectivity. Proc Natl Acad Sci U S A 117:28393–28401.