Cerebral Cortex, 2021;00: 1–15

https://doi.org/10.1093/cercor/bhab175
Original Article

ORIGINAL ARTICLE

Dynamic Network Analysis Demonstrates the Formation of Stable Functional Networks During Rule Learning

Thomas M. Morin[1,2], Allen E. Chang[2,3], Weida Ma[2,4], Joseph T. McGuire[1,2,3] and Chantal E. Stern[1,2,3]

[1]Graduate Program for Neuroscience, Boston University, Boston, MA 02215, USA, [2]Cognitive Neuroimaging Center, Boston University, Boston, MA 02215, USA, [3]Department of Psychological and Brain Sciences, Boston University, Boston, MA 02215, USA and [4]Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA

Address correspondence to Chantal E. Stern, Department of Psychological and Brain Sciences, Boston University, Boston, MA 02215, USA.
Email: chantal@bu.edu

Abstract

Variations in the functional connectivity of large-scale cortical brain networks may explain individual differences in learning ability. We used a dynamic network analysis of fMRI data to identify changes in functional brain networks that are associated with context-dependent rule learning. During fMRI scanning, naïve subjects performed a cognitive task designed to test their ability to learn context-dependent rules. Notably, subjects were given minimal instructions about the task prior to scanning. We identified several key network characteristics associated with fast and accurate rule learning. First, consistent with the formation of stable functional networks, a dynamic community detection analysis revealed regionally specific reductions in flexible switching between different functional communities in successful learners. Second, successful rule learners showed decreased centrality of ventral attention regions and increased assortative mixing of cognitive control regions as the rules were learned. Finally, successful subjects showed greater decoupling of default and attention communities throughout the entire task, whereas ventral attention and cognitive control regions became more connected during learning. Overall, the results support a framework by which a stable ventral attention community and more flexible cognitive control community support sustained attention and the formation of rule representations in successful learners.

Key words: fMRI, learning, memory, networks

Introduction

We often encounter situations where we must learn and apply rules. For example, imagine searching for parking on an unfamiliar road. Depending on the context of the situation—the location, time of day, and status of your parking permit—a "no-parking" rule may or may not apply. Here, we examine the changes in functional network connectivity associated with learning a set of context-dependent rules.

Studies of rule learning have described a coordinated system of prefrontal cortical regions essential for representing and utilizing rules (for review, see Mansouri et al. 2020). Electrophysiological recordings in non-human primates have identified neurons in orbitofrontal cortex, ventrolateral prefrontal cortex (vlPFC), and dorsolateral prefrontal cortex (dlPFC) that selectively fire for different abstract rules (Hoshi et al. 2000; Wallis et al. 2001). Recordings from dlPFC neurons demonstrate that as rules become learned, the onset of neuronal firing shifts earlier in the trial (Pasupathy and Miller 2005; Cromer et al. 2011).

In humans, network neuroscience offers an array of methods to study the dynamic network changes associated with learning

(for review, see Bassett and Mattar 2017). With these methods, the brain is conceptualized as a network of nodes (brain regions), connected by weighted edges (functional connectivity strength). Nodes can be assigned to different cortical networks, which we call communities. Canonical communities (e.g., default, dorsal/ventral attention, and cognitive control communities) have been defined previously from large samples of resting state fMRI data (Power et al. 2011; Yeo et al. 2011). Dynamic network analyses have been extensively employed in studies of implicit motor sequence learning, showing that during a serial reaction time (SRT) task sensory systems become functionally segregated, and PFC becomes functionally integrated (Bassett et al. 2015). Moreover, the degree of community switching (flexibility) in frontoparietal cortical regions was predictive of subsequent performance on the task (Bassett et al. 2011). Task-specific network reconfiguration has been observed when comparing difficulty levels on an n-back working memory task (Braun et al. 2015), on an abstract reasoning task (Hearne et al. 2017), and during reinforcement learning (Gerraty et al. 2018). Studies of associative memory have shown that the functional relationships within and between communities can be indicative of learning rate and task-performance (Fatima et al. 2016; Gerraty et al. 2018). Broadly, decoupling of default and dorsal attention communities is a strong marker of cognitive states (Dixon et al. 2017) and is a predictor of performance on working memory (Keller et al. 2015) and fluid intelligence tasks (Cole et al. 2012). The cognitive control community regulates this decoupling by supporting switching between different mental representations (Cole and Schneider 2007; Spreng et al. 2013) and becomes active during memory-guided attention (Rosen et al. 2016).

Here, we applied network neuroscience techniques to a context-dependent rule learning task. During the task, naïve participants learned a novel set of context-dependent rules in order to correctly identify paired associates (Zhu et al. 2020). In contrast with previous studies, which have primarily examined rule retrieval, we designed the task to include minimal instruction and no training before scanning to specifically investigate rule learning. Overall accuracy on the task varied significantly over time and across individuals, allowing us to compare the brain network organization of successful and unsuccessful learners. Through a dynamic network analysis, we characterized shifts in network metrics associated with successful learning.

The primary goals of our network analysis were to 1) assess how network dynamics differed between successful and unsuccessful learners and 2) characterize how network measures shifted as subject performance improved on the task. We predicted that successful learners would form stable functional representations of the rules, which would correspond to lower levels of community switching (network flexibility). This hypothesis is supported by previous literature showing that performance on language, reasoning, and working memory tasks is associated with increased stability of functional networks (Schultz and Cole 2016; Ferguson et al. 2017). As a general framework, we predicted that the task would become more automatic for successful learners as their performance improved, corresponding to 1) greater functional segregation (network assortativity) and decreased centrality of cognitive control regions as learning occurred, and 2) decoupling of the default and attention communities. Our results provide a new perspective, identifying a key role for brain-wide functional networks that are shown to decouple and stabilize during rule learning.

Materials and Methods

Participants

A total of 32 participants were scanned for the study (age, mean = 22.78, SD = 3.97; 14 female) and 29 were included in the final analysis. Participants were recruited from the Boston area and were English language speakers between the ages of 18 and 35. All subjects had normal or corrected-to-normal vision and reported no past or current neurological or psychiatric disorders. All procedures were approved by the Boston University Institutional Review Board and all subjects provided written informed consent. For the three participants not included in the final analysis, one withdrew during scanning, one was excluded for undisclosed prior knowledge about the task, and one was excluded for excess head motion (>3 mm framewise displacement during two or more scanning runs).

Cognitive Task

Before fMRI scanning, participants were not pre-trained or provided with task instructions. During scanning, subjects performed a cognitive task designed to test their ability to learn context-dependent associative rules (Fig. 1). Our lab has previously used this task for computational modeling and behavioral testing of context-dependent rule learning (Zhu et al. 2020). During the task, a cue image appeared on the screen (2.0 s), followed by a delay (2.0 s) and then an associate image (2.0 s). Subjects were then allowed 2.0 s to decide whether or not the cue image was correctly paired with the associate image. Subjects indicated their response (match or mismatch) by pressing the left or right button on a response-box, and were provided with feedback ("Correct," "Incorrect," or "No Response") for 0.5 s. Between each trial, a fixation cross was displayed for 4.0 s. There were four context-dependent cue-associate pairings that subjects encountered during the scan, and they are described in Figure 1. Scanning consisted of nine runs, with 32 trials per run.

Behavioral Data Analysis

Accuracy was calculated for each subject across each of the nine scanning runs. Each scanning run consisted of 32 trials of the context-dependent rule learning task, and subject performance was used as a measure of when (and if) subjects shifted from "learning" to "learned" phases. With the task set up in a 2-alternative forced choice format, we assumed that subjects who had not yet learned the context-dependent rules would achieve about 50% accuracy (16 correct responses per run) purely by chance. Using the binomial distribution, we determined the observed level of accuracy per run that would have a 0.1% likelihood of occurring by chance, and set the threshold for above-chance performance to 26 correct responses out of 32 trials per scanning run (or 81.25% correct). If a subject responded correctly to at least 26 trials during one run, then that run was classified as "learned." Subjects were considered "successful" learners if they achieved at least one "learned" scanning run. Subject performance is presented in Figure 3.

MRI Data Acquisition

Scanning was conducted using a 3 Tesla Siemens TIM Trio magnetic resonance scanner and a 32-channel head coil at the Center for Brain Science at Harvard University in Cambridge, Massachusetts. High-resolution T1-weighted multi-planar rapidly

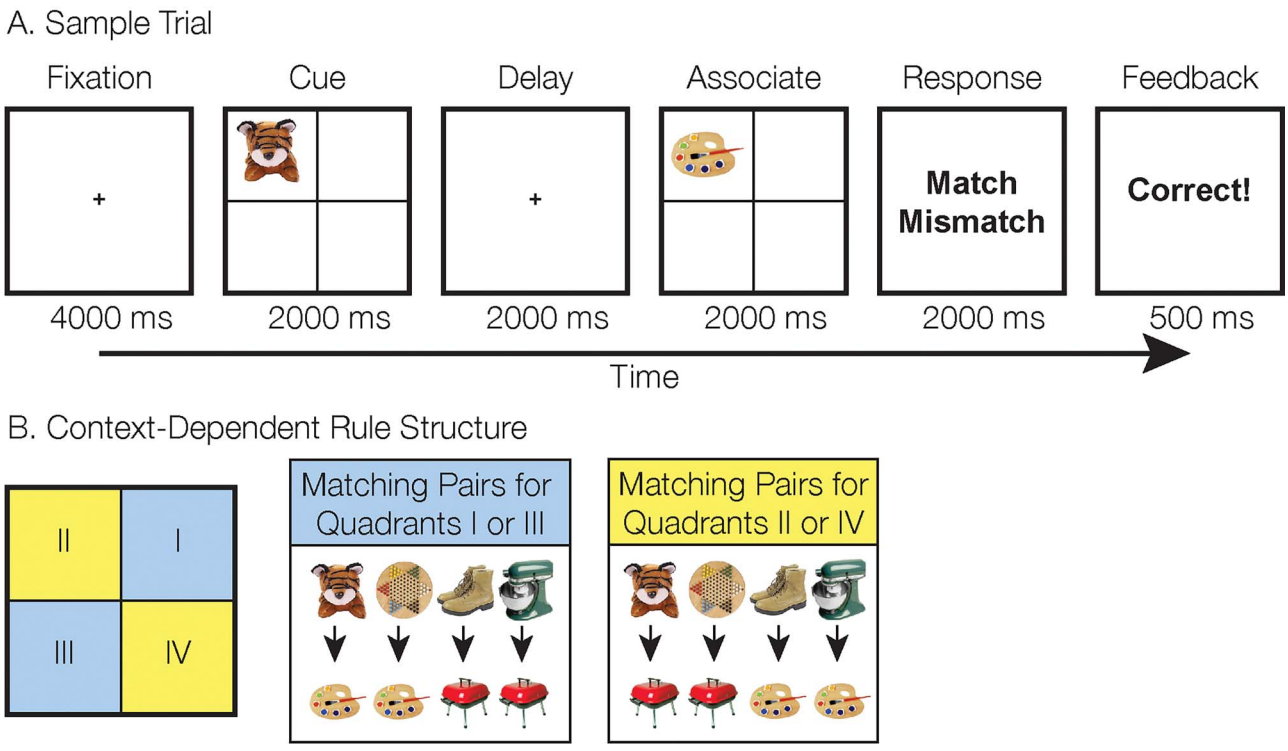

Figure 1. Context-dependent rule learning task. (A) During each trial, the subject saw a fixation (4000 ms), followed by a cue stimulus (2000 ms), delay (2000 ms), and an associate stimulus (2000 ms). Then, the subject was given 2000 ms to decide whether or not the cue image was correctly paired with the associate image. Subjects indicated their response by pressing the left or right button on a response box. The trial ended with 500 ms of feedback ("Correct," "Incorrect," or "No Response"). (B) Subjects were naïve to the fact that a context-dependent rule structure dictated whether cue and associate stimuli were correctly paired. If the cue and associate appeared in Quadrant I or III (highlighted here in blue for emphasis), they obeyed one set of rules (e.g., the tiger and paint-palette are a pair). If instead the cue and associate appeared in Quadrant II or IV (yellow), they obeyed an alternate set of rules (e.g., the tiger and grill are now a match). Subjects were tasked with learning these context-dependent pairings over the course of nine scanning runs. Each scanning run consisted of 32 trials of the task.

acquired gradient echo (MP-RAGE) scans were acquired for each subject (TR = 2530 ms; TE = 1.64 ms; flip angle = 7°; slices = 176; resolution = 1 mm isotropic). T2*-weighted EPI (BOLD) images were acquired using a slice-accelerated EPI sequence that permitted simultaneous multi-slice acquisitions using the blipped-CAIPI technique (TR 2 s; TE 26 ms; flip angle 80°; 6/8 partial-Fourier acquisition) (Setsompop et al. 2012). A total of 69 slices were acquired with a slice acceleration factor of 3 and no skip, covering the whole brain. Images were acquired at a nominal 2-mm isotropic spatial resolution (matrix size: 108 × 108).

fMRI Preprocessing

Functional data were analyzed using the Freesurfer FS-FAST software package (version 6.0, Charlestown, MA; http://surfer.nmr.mgh.harvard.edu/) (Fischl 2012). The following preprocessing steps were performed: slice-time correction, motion correction, spatial smoothing (3 mm FWHM), boundary-based registration with 12-degrees of freedom to co-register each subject's functional and anatomical data, and spatial normalization of each subject's reconstructed surface to Freesurfer's standard template (fsaverage) using spherical surface registration. Functional data were further preprocessed in MATLAB (The MathWorks). Head-motion regression (6 motion parameters and their 6 temporal derivatives), whole-brain signal regression, and ventricular and white matter signal regression were performed (Van Dijk et al. 2010). Additionally, our matrix of nuisance regressors included two event-related regressors corresponding to the Response and Feedback events in the task. For the Response events, regressors were created by convolving a standard hemodynamic response function with a duration-modulated boxcar, where boxcars were scaled by the subject's response time on a per-trial basis. Similarly, the Feedback regressors were created by convolving the same standard hemodynamic response function with a boxcar enveloping the length of the feedback period (0.5 s). We included these event-related regressors to ensure that observed changes in network connectivity were associated with rule learning, rather than differences in response times or feedback presentation between successful and unsuccessful learners. We then calculated framewise displacement by taking the sum of the absolute derivatives of the 6 motion parameters for each time point (Power et al. 2012). A threshold of 0.5 mm was set to identify time points with excessive motion. To avoid artifact spread during wavelet decomposition, high motion timepoints were replaced using linear interpolation (Carp 2013).

Network Construction

From the 2018 Schaefer-parcellation, we defined 400 cortical regions of interest across the two brain hemispheres (200 per hemisphere, see Fig. 2; Schaefer et al. 2018). For each subject and each run of data, these 400 brain regions formed the nodes of a network. To define edges between the nodes, the mean BOLD signal time course was extracted from each region and the wavelet correlation between time-courses for each pair of regions (i, j) was calculated. Maximum overlap

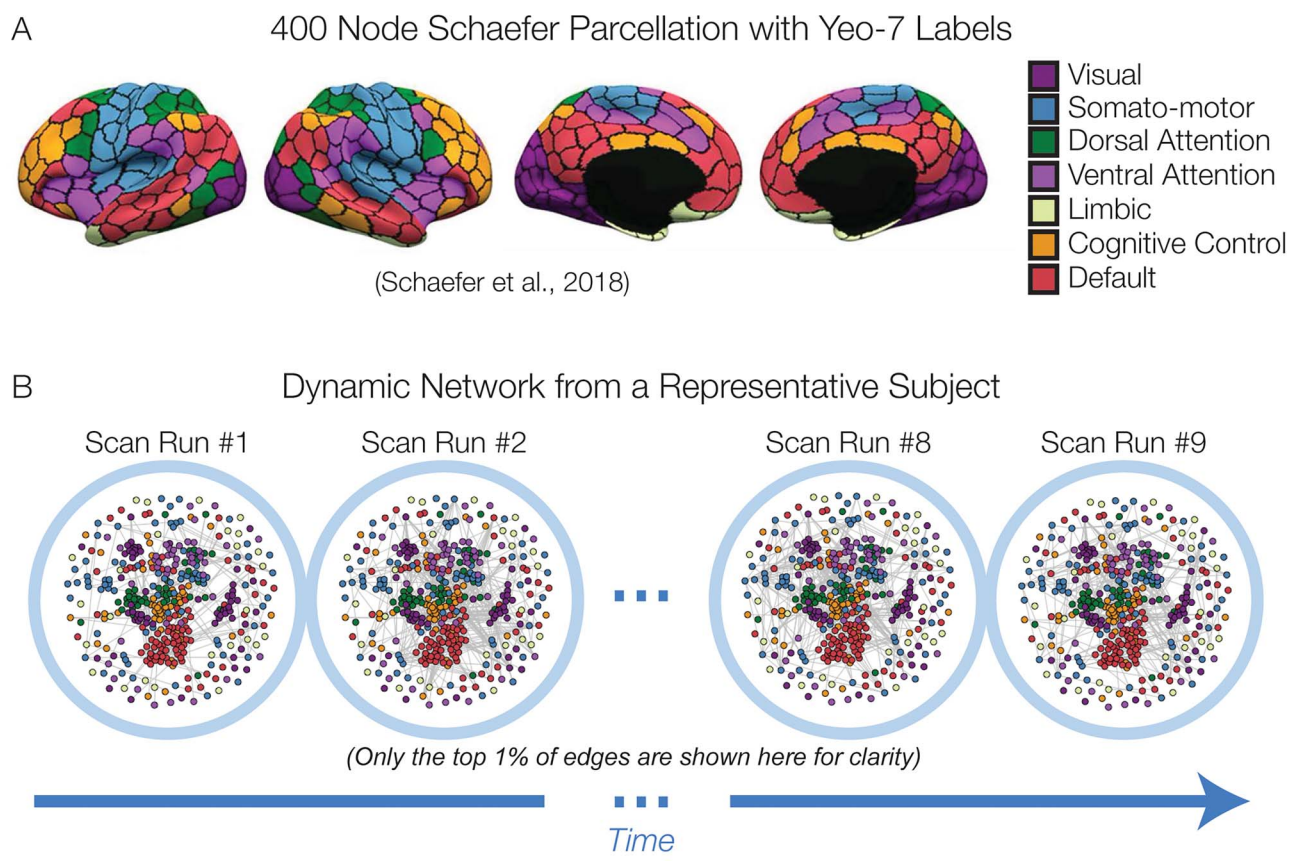

Figure 2. Network construction. (*A*) Inflated representation of cortex divided into 400 regions, each assigned to a Yeo-community (adapted with permission from https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal). (*B*) Multilayer network of the connectivity between regions in a single subject. Each brain region from (*A*) is represented by a node. Edges represent functional connectivity between regions. For clarity, only the top 1% of strongest edges in each layer are shown here.

discrete wavelet transform was used to decompose a time-course into scales corresponding to frequency bands (This is identical to the methods used in Bassett et al. 2011. The software for this method can be found at: http://grinsted.github.io/wavelet-coherence/.) We used scale 2 (0.0625–0.125 Hz) in our analysis because this frequency overlaps most with the canonical hemodynamic response function. In line with previous research using wavelet methods to define network edges, we calculated the Pearson correlation between scale 2 time courses (Bassett et al. 2011; Zhang et al. 2016). Our choice of wavelet correlation over wavelet coherence means that the network does include negative edges. Wavelet correlation was most appropriate for our analysis because we calculated functional connectivity measurements over an entire scanning run (450 s of data, 225 TRs), and the non-stationarity of fMRI time courses makes coherence measures inappropriate for this large time scale (Bullmore et al. 2004; Zhang et al. 2016). Edge weights in the network were subjected to significance testing using the rcorr.adjust function in R. For a pair of nodes i and j, if functional connectivity was deemed statistically different from zero according to a liberal statistical threshold (alpha = 0.05, Holm corrected for multiple comparisons), then the weight of edge (i, j) remained the same. Otherwise, it was set to zero, eliminating the edge from the graph. For each subject, this resulted in nine network layers, spanning the nine runs of fMRI scanning. For all network statistics and clustering methods described below (except for betweenness centrality), we included both positive and negative edge weights. When calculating betweenness centrality, we only included positive edge weights. This is because betweenness centrality calculations depend on deriving shortest paths between pairs of nodes, a problem that is ill-defined for nodes connected by negatively-weighted edges. Although including negative edge weights can complicate the interpretation of network measures, we chose to include negative edge weights when possible. Previous work indicates that negative edges are physiologically meaningful, indicating segregation or competition between regions (Fox et al. 2005; Rubinov and Sporns 2011; Keller et al. 2013). (For additional discussion on the inclusion of negative edges, see: Schwarz and McGonigle 2011; Fornito et al. 2015). Finally, only 3.5% of the within-layer edges were negatively weighted in our subjects' networks, suggesting that their overall influence on network measures is small.

Dynamic Community Detection and Flexibility

For community detection and node flexibility calculations, we combined the nine network layers into a single dynamic network for each subject. In the dynamic network, each layer corresponded to one of the nine scanning runs, ordered in time. Networks were linked across time by connecting each node in one layer to itself in the immediately adjacent layers. Inter-layer edges were equally weighted, $\omega = 1$.

To determine the community membership(s) of each node over time, we used a Louvain community detection algorithm that is generalized for multi-layered networks (Mucha et al.

2010). MATLAB code for the algorithm was produced by Mucha and Porter and is freely available online: http://netwiki.amath.unc.edu/GenLouvain/GenLouvain. The algorithm detects communities by optimizing multilayer modularity, Q_{ml} (originally presented in this format by Bassett et al. 2011):

$$Q_{ml} = \frac{1}{2\mu} \sum_{ijlr} \left\{ \left(A_{ijl} - \gamma_l \frac{k_{il} k_{jl}}{2m_l} \right) \delta_{lr} + \delta_{ij} \omega_{jlr} \right\} \delta \left(g_{il}, g_{jr} \right), \quad (1)$$

where the adjacency matrix of layer l (e.g., scanning run number l) has components A_{ijl}, γ_l is the resolution parameter of layer l, g_{il} gives the community assignment of node i in layer l, g_{jr} gives the community assignment of node j in layer r, the function $\delta(g_{il}, g_{jr})$ is 1 when $g_{il} = g_{jr}$ and 0 otherwise, ω_{jlr} is the connection strength between node j in layer r and node j in layer l, k_{il} is the strength of node i in layer l, m_l is half the sum of all the edge weights in layer l, $2\mu = \sum_{jr} \kappa_{jr}$, $\kappa_{jl} = k_{jl} + \omega_{jl}$, and $\omega_{jl} = \sum_r \omega_{jlr}$. We set $\omega = 1$ for edges connecting the same node in adjacent layers and set the resolution parameter $\gamma_l = 1$.

By this definition, modularity is minimized when strongly connected groups of nodes (both within and between layers) are part of the same community. The between-layer connectivity strength is summarized by the term $\delta_{ij}\omega_{jlr}$, and the within-layer connectivity strength is summarized by the term $(A_{ijl} - \gamma_l \frac{k_{il}k_{jl}}{2m_l})\delta_{lr}$. This algorithm allows a node to change its community allegiance across layers. See Mucha et al. (2010) for details on precise methods and implementations. Note, the community detection algorithm discussed here is a data-driven approach. The communities it assigns are unrelated to the canonical resting state communities (Yeo-communities) that are discussed elsewhere in this paper. The Yeo-communities were determined a priori with a parcellation developed by Schaefer et al. (2018), without any influence from our subjects' fMRI data.

After running the community detection algorithm, we calculated flexibility for each node by counting the number of times a node switched communities and dividing by the possible number of switches (There were eight possible switches, since our network had nine layers.) Because the outcome of the community detection algorithm depends on initial conditions (e.g., which node is first assigned to a community), we ran the algorithm 100 times for each subject, calculating average flexibility for each node across iterations.

Whole brain average flexibility was determined for each subject by taking the mean flexibility across all vertices. The average flexibility was also determined for each of the seven Yeo-communities (Yeo et al. 2011). Average flexibility for the whole brain is plotted against accuracy on the cognitive task for each subject in Figure 4A. To visualize the spatial distribution of flexibly-changing nodes on cortex, and the relationship of flexibility with accuracy, we projected the group mean correlations between the flexibility of an individual node and subject accuracy onto the cortical surface in Figure 4B. Average flexibility for each of the seven Yeo-communities is plotted against task accuracy in Figure 4C. The relationship between flexibility and learning was assessed through Pearson correlation using the cor.test function in R, the results of which are summarized in Table 1 (*P*-values FDR corrected).

Assortative Mixing

To determine whether brain regions connected mostly to other regions in the same Yeo-community and the extent to which this differed over time and across successful and unsuccessful learners, we calculated assortative mixing for each Yeo-community. We performed these calculations within each layer of a subject's network to determine the assortative mixing of Yeo-communities for each scan run. To calculate assortative mixing, we used the assortativity function in the igraph package available for R, which uses Newman's definition of assortativity (Newman 2003):

$$r = \frac{\sum_i e(i,i) - \sum_i a(i) b(i)}{1 - \sum_i a(i) b(i)} \quad (2)$$

where $e(i,j)$ is the fraction of edges connecting nodes labeled i and j, $a(i) = \sum_j e(i,j)$, and $b(j) = \sum_i e(i,j)$. When calculating $e(i,j)$, negative edges subtract from the numerator, but add to the denominator. In this way, negative within-category edges decrease assortativity. Assortativity r is equal to one when there is perfect assortative mixing (all edges are within-category). Similarly, it is equal to zero when the mixing is completely random and r approaches -1 when mixing is disassortative (edges on the graph connect vertices with other vertices outside of their category). We calculated separate assortativity measures for each Yeo-community by re-labeling the nodes of the graph with only two labels: nodes within the community being measured, and those outside of the community being measured. Assortativity across communities, subjects, and scan runs is plotted in Figure 5A and summarized in Table 2 (*P*-values FDR corrected).

Centrality

To determine which brain regions integrated information from many other brain regions, and whether this was altered with time and task performance, we calculated the average betweenness centrality for each of the seven Yeo-communities across subjects and scanning runs. Betweenness centrality was originally defined by Freeman (1978) and is defined by Kolaczyk (2009) as:

$$c_B(v) = \sum_{s \neq t \neq v \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)} \quad (3)$$

where $\sigma(s,t|v)$ is the number of shortest paths from s to t that include v, and $\sigma(s,t)$ is the total number of shortest paths from s to t. Betweenness centrality is high for vertices that are on the shortest paths between many other nodes. Betweenness centrality is ill-defined for networks with negative edge-weights, so we eliminated all negative edges from our networks for this analysis. Additionally, because betweenness centrality relies on a shortest path metric, edges on the graph were transformed from a metric of strength (wavelet correlation) to a metric of distance, using the inverse transform ($1/A_{ij}$). We then calculated betweenness centrality using the betweenness function available in the igraph library for R. Betweenness centrality across communities, subjects, and scan runs is plotted in Figure 5B and displayed in Table 3 (*P*-values FDR corrected).

Inter-Community Edge Strength

To determine how connectivity between the Yeo-communities changed with learning, we calculated the average weight of inter-community edges for each scanning run. The inter-community edge strength between two communities i and j

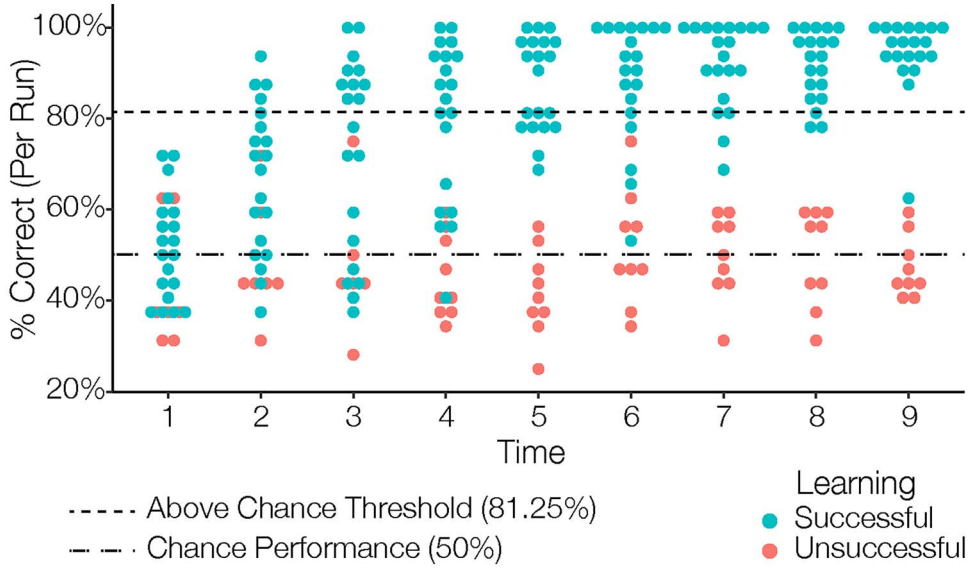

Figure 3. Rule learning ability. Behavioral results. Subject accuracy is plotted here over time. Time points represent each of the 9 scanning runs (32 trials per run). Subjects who achieved accuracy significantly above chance in at least one run are considered successful learners and plotted in blue ($n = 20$). Accuracy measurements for unsuccessful learners are plotted in red ($n = 9$).

is defined as:

$$s(i,j) = \frac{\sum_{i \neq j} w_{ij}}{n_{ij}} \quad (4)$$

where w_{ij} is the weight of an edge connecting a node in community i with a node in community j and n_{ij} is the total number of edges connecting communities i and j. This was calculated using the simplify function available in the igraph library for R. Negative edges decrease the overall inter-community edge strength between two communities. We plotted the time-varying inter-community edge strength for each pair of communities in both successful and unsuccessful learners in Figure 6. Statistics are reported in Table 4 (*P*-values FDR corrected).

Dynamic Network Statistics

To test how assortative mixing, betweenness centrality, and inter-community edge strength differed between learning groups and changed throughout learning, we estimated two linear mixed effects models implemented with the lme4 package in R. With the first model (Model 1) we examined the effects of learning group (successful or unsuccessful) and time on each of the network measures. The model included fixed effects for learning group, time (scan run), and the interaction between learning group and time, as well as a random effect of subject. We report results where the intercept varied across subjects. We also conducted a version of Model 1 allowing slope to vary for each subject, but this model failed to converge and fit poorly in most circumstances, so we do not report it here. A second model (Model 2) was used to test the effect of each subject's unique learning curve (measured as each subject's accuracy on each scan run) on the network metrics. This model included fixed effects for run-to-run accuracy as well as overall task accuracy to control for between-subject differences in overall performance. The model also included a random effect of subject. Both models were fit to data for each of the seven Yeo-communities (or each pair of communities, in the case of inter-community edge strength). We applied a false discovery rate (FDR) correction for multiple comparisons across communities and model parameters (Benjamini and Hochberg 1995). Unless otherwise noted, all reported *P*-values have been corrected using FDR correction.

Results

Behavioral Results

As expected, performance on the cognitive task varied considerably across subjects. Subject accuracy over time is displayed in Figure 3. A few subjects learned the context-dependent rules and achieved ceiling-levels of accuracy in just two or three scanning runs. Other subjects remained at chance-levels of accuracy throughout. Subjects who remained at chance-level performance throughout all nine scanning runs were considered "unsuccessful" learners, and subjects who achieved above-chance performance on at least one scan run were considered "successful" learners. Overall, 20 subjects met criteria as successful learners and 9 subjects were unsuccessful. The variation in learning ability provided us with a natural control group for identifying functional network characteristics associated with successful rule learning.

Network Stability and Community Switching

To examine functional network architecture changes over the course of the task, we can measure the community switching (flexibility) of a given brain region. As a working model, we proposed that during learning, brain regions flexibly adapted their community allegiance as subjects adopted different task strategies. Once the task was learned, the community structure became more rigid (less flexible). Subjects with greater accuracy spent more time in the "learned" phase than the "learning" phase. Therefore, we predicted that overall accuracy would be inversely related to flexibility and hypothesized that sub-

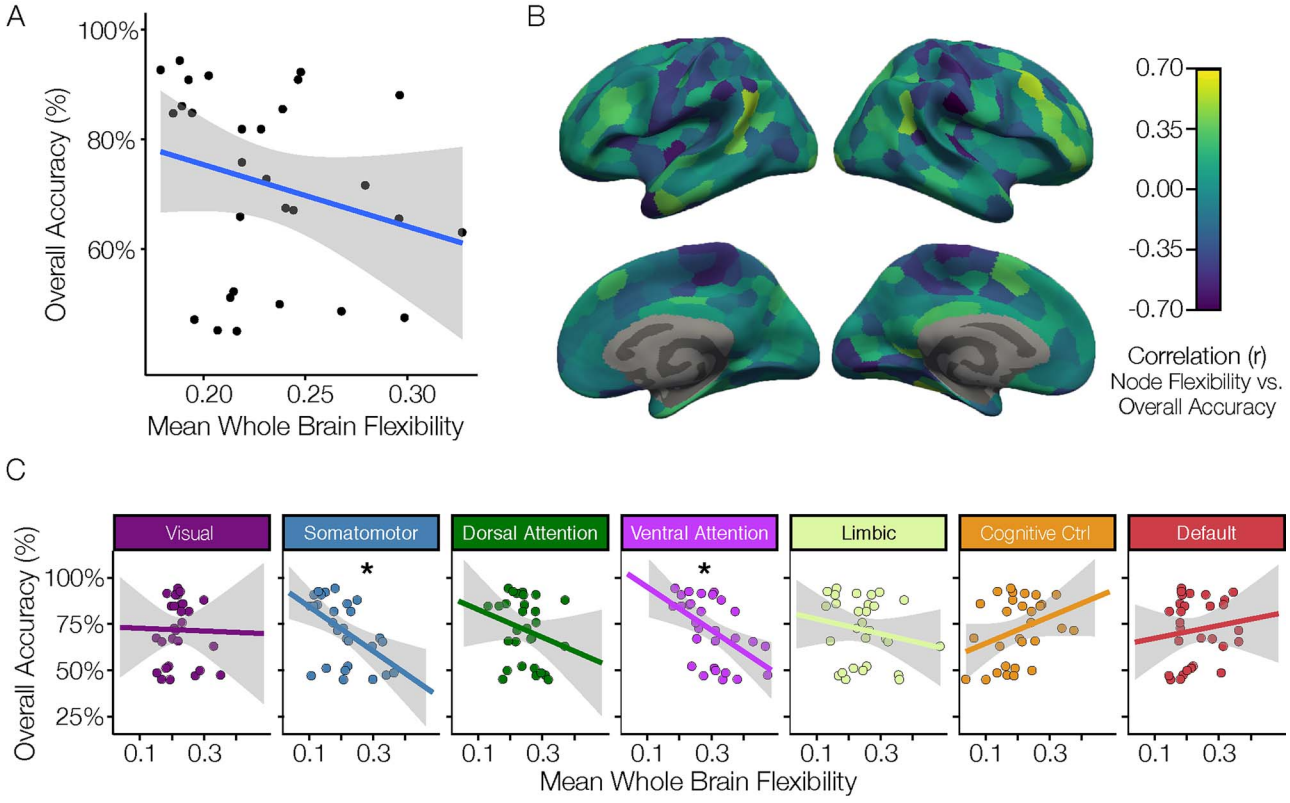

Figure 4. Flexibility. (*A*) Whole-brain average flexibility is plotted against accuracy on the cognitive task for each subject. (*B*) Cortical map of the correlation between flexibility for each of the nodes in the Schaefer400 atlas and overall accuracy. (*C*) Average flexibility within each of the seven Yeo-communities is plotted against task accuracy. Trend lines show a linear regression between flexibility and accuracy. Gray area represents a 95% confidence interval. * denotes statistically significant correlation; $P < 0.05$, FDR corrected. All correlation statistics are presented in Table 1.

Table 1 Relationship between accuracy and flexibility

Community	Correlation of accuracy and flexibility		
	r	T(28)	*P*
Visual	−0.02	−0.11	0.92
Somato-motor	**−0.53**	**−3.23**	**<0.05**
Dorsal attention	−0.25	−1.31	0.32
Ventral attention	**−0.52**	**−3.20**	**<0.05**
Limbic	−0.19	−1.01	0.43
Cognitive control	0.33	1.81	0.22
Default	0.14	0.71	0.55
Whole brain avg.	−0.25	−1.36	0.32

Notes: Pearson correlation was calculated to assess the relationship between accuracy on the cognitive task and flexibility of brain regions; repeated for each Yeo-community and across the whole brain. *r* = Pearson correlation coefficient; T(28) = *T*-statistic with degrees of freedom of 28; *P* = adjusted *P*-value (FDR corrected). Bold results show a statistically significant relationship between accuracy on the task and brain-region flexibility, $P < 0.05$, FDR corrected for multiple comparisons.

jects with greater accuracy on the cognitive task would exhibit decreased brain network flexibility, consistent with the rapid development of stable rule representations.

We plotted mean flexibility across the whole brain against subject accuracy in Figure 4A. We did not observe a significant association between overall accuracy on the task and whole-brain mean flexibility. In Figure 4B, we projected the correlation between accuracy and flexibility for each node in the Schaefer-400 atlas onto the cortical surface. A regional pattern emerged with certain areas showing consistently high or low correlation between accuracy and flexibility. To determine if certain Yeo-communities were driving this result, we plotted the average flexibility for each community against accuracy in Figure 4C. We found a statistically significant Pearson correlation coefficient between flexibility and accuracy in the ventral attention community ($r = -0.524$, $P < 0.05$) as well as in the somatomotor community ($r = -0.527$, $P < 0.05$) (FDR corrected for multiple comparisons). We failed to detect a significant correlation in any of the other Yeo-communities. The correlation coefficients and corresponding *P*-values for all seven Yeo-communities and the whole brain are displayed in Table 1.

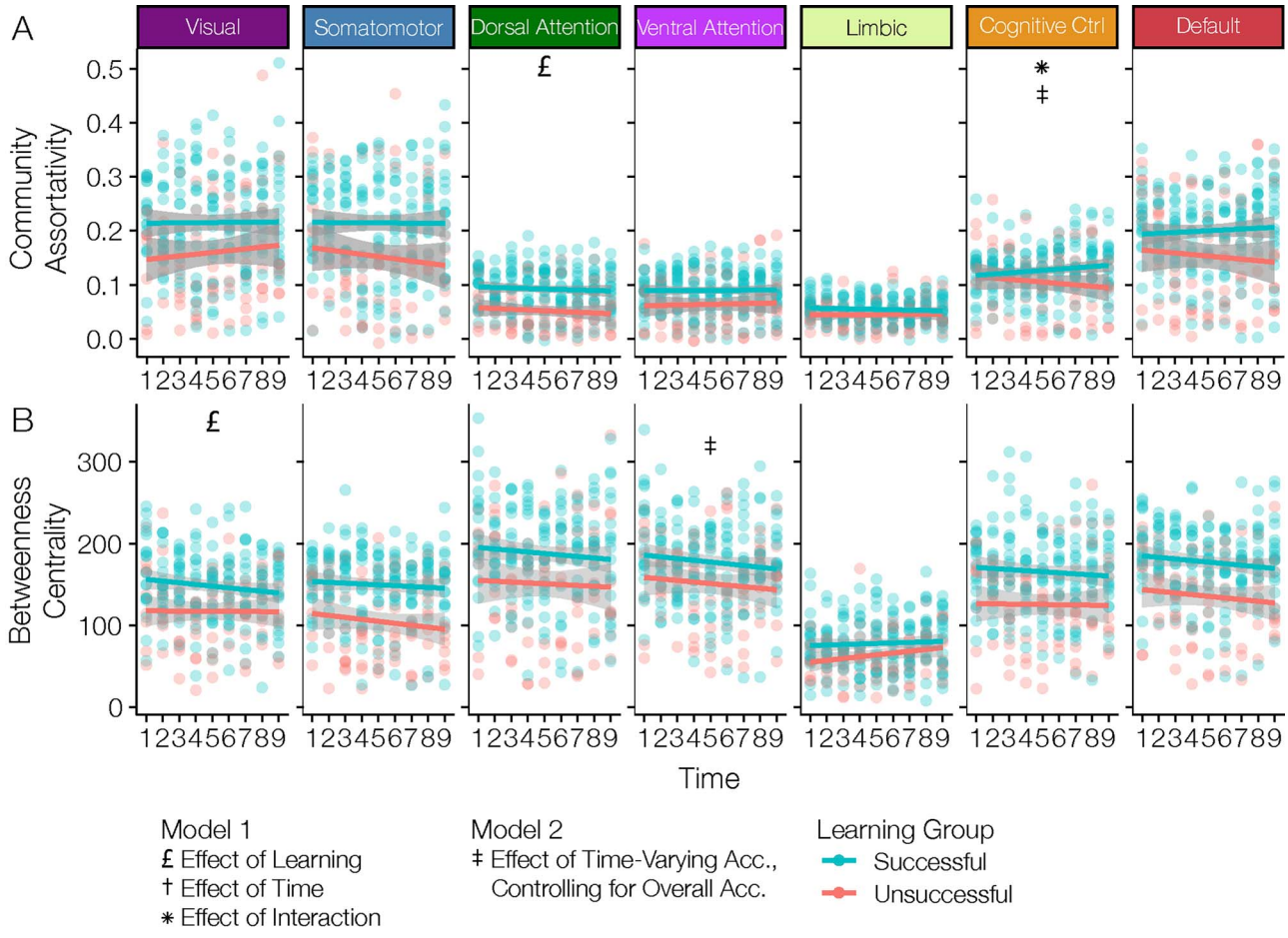

Figure 5. Community assortativity and betweenness centrality. (A) Community assortativity and (B) mean betweenness centrality of the nodes in each Yeo-community were calculated for each of the 9 scanning runs for each subject. Points are colored blue for subjects who successfully learned the task, and red for subjects who did not learn the task. Linear regression of changes in assortativity and betweenness centrality for each group over time are plotted. Gray area surrounding regression lines represents a 95% confidence interval. Two linear mixed effects models were used to assess the effects of learning group and time (Model 1) as well as the effect of run-to-run accuracy (Model 2) on the network measures. Model results are presented in Tables 2 and 3. £ = Effect of Learning Group; † = Effect of Time; * = Effect of Learning Group x Time Interaction; ‡ = Effect of Run-to-Run Accuracy, Controlling for Overall Accuracy ($P < 0.05$, FDR corrected).

Assortative Mixing

Assortative mixing is defined here as the degree to which nodes are highly connected with other nodes in the same Yeo-community. A positive value for assortative mixing indicates that nodes within a community are more strongly connected to each other than they are to nodes outside of their community. We predicted that assortative mixing would be stronger in successful learners, which would correspond to the formation of stable, modular networks and the predicted decrease in node flexibility. For each of the seven Yeo-communities, we calculated assortative mixing during each scanning run. Results are displayed in Figure 5A. Using two linear mixed effects models, we tested the effects of learning, time, and the interaction between learning and time on assortativity (Model 1), as well as the effect of run-to-run accuracy on assortativity while controlling for the effect of overall accuracy (Model 2). Model results are presented in Table 2. Model 1 showed a significant interaction between learning and time on assortativity in the cognitive control community ($T = 2.70$, $P < 0.05$). Post hoc testing with the lstrends command in R showed that the slope of assortativity over time was significantly different between successful learners, who showed an increase in assortativity over time, and unsuccessful learners, who showed a decrease in assortativity over time. However, each group's slope itself did not significantly differ from zero (successful learners: slope = 0.0022, $T = 1.69$, $P = 0.093$; unsuccessful learners: slope = −0.0034, $T = -1.73$, $P = 0.086$). This result indicated that the interconnectivity of cognitive control nodes showed a more positive trend over time in successful learners than in unsuccessful learners, albeit without supporting strong conclusions about the sign of the trend in each group individually. Echoing this finding, Model 2 showed a significant positive effect of run-to-run accuracy on assortativity in the cognitive control community, even when controlling for overall accuracy ($T = 2.60$, $P < 0.05$). This indicates that changes in assortativity of the cognitive control community were significantly associated with individual learning curves across subjects.

Centrality

Centrality measures the degree to which a brain region acts as a mediator between distantly connected brain regions. We

Table 2 Linear mixed effects models for assortativity

Community	Model 1						Model 2			
	Learning		Time		Learning* time		Overall acc.		Acc. per run	
	Effect	P	Effect	P	Effect	P	Effect	P	Effect	P
Visual	0.07	0.12	0.003	0.43	–0.003	0.54	**0.20**	**<0.05**	0.01	0.90
Somatomotor	0.04	0.36	–0.004	0.36	0.004	0.45	**0.22**	**<0.05**	0.01	0.90
Dorsal attn.	**0.04**	**<0.05**	–0.001	0.43	0.000	0.82	**0.12**	**<0.05**	–0.002	0.90
Ventral attn.	0.03	0.15	0.001	0.73	–0.001	0.82	0.07	0.13	0.014	0.55
Limbic	0.01	0.20	0.000	0.90	–0.001	0.62	**0.05**	**<0.05**	–0.017	0.19
Cognitive ctrl	–0.003	0.90	–0.003	0.20	**0.005**	**<0.05**	0.01	0.90	**0.041**	**<0.05**
Default	0.03	0.54	–0.003	0.41	0.004	0.28	**0.20**	**<0.05**	–0.008	0.90

Notes: Results from two linear mixed effects models that were run for each of the seven Yeo-communities. Model 1 treated each subject as a random effect and included fixed effects for learning group (successful or unsuccessful), time, and the interaction between learning group and time. Model 2 treated each subject as a random effect and included fixed effects for overall accuracy and time-varying accuracy (per run). *P* = adjusted *P*-value (FDR corrected). Bold results show a statistically significant effect, $P < 0.05$, FDR corrected for multiple comparisons.

Table 3 Linear mixed effects models for betweenness centrality

Community	Model 1						Model 2			
	Learning		Time		Learning* time		Overall acc.		Acc. per run	
	Effect	P	Effect	P	Effect	P	Effect	P	Effect	P
Visual	**40.60**	**0.04**	−10.122	0.952	−1.98	0.39	**119.87**	**<0.01**	−122.53	0.21
Somatomotor	37.84	0.06	−12.406	0.165	1.37	0.56	**149.08**	**<0.01**	−10.98	0.95
Dorsal attn.	41.07	0.15	−11.049	0.727	−10.80	0.86	**129.02**	**<0.05**	−113.75	0.57
Ventral attn.	27.54	0.25	−11.899	0.391	−10.23	0.95	**145.43**	**<0.01**	**−41.09**	**<0.05**
Limbic	21.95	0.12	2.233	0.097	−11.60	0.35	**59.91**	**<0.05**	7.46	0.57
Cognitive ctrl	45.41	0.06	−10.280	0.952	−11.05	0.73	**150.84**	**<0.01**	−12.14	0.95
Default	41.30	0.06	−12.039	0.286	0.11	0.95	**151.09**	**<0.01**	−111.32	0.57

Notes: Results from two linear mixed effects models that were run for each of the seven Yeo-communities. Model 1 treated each subject as a random effect and included fixed effects for learning group (successful or unsuccessful), time, and the interaction between learning group and time. Assessed the effects of learning, time, and the interaction between learning and time. Model 2 treated each subject as a random effect and included fixed effects for overall accuracy and time-varying accuracy (per run). *P* = adjusted *P*-value (FDR corrected). Bold results show a statistically significant effect, $P < 0.05$, FDR corrected for multiple comparisons.

calculated mean betweenness centrality across nodes in each of the seven Yeo-communities during each scanning run. Results from this analysis are plotted in Figure 5B. Using two linear mixed effects models, we tested the effects of learning, time, and the interaction between learning and time on betweenness centrality (Model 1), as well as the effect of time-varying accuracy on betweenness centrality while controlling for the effect of overall accuracy (Model 2). Model results are presented in Table 3. Model 1 did not reveal any significant interactions between learning and time. Model 2 showed that when controlling for overall accuracy, increased run-to-run accuracy is significantly associated with reduced betweenness centrality in the ventral attention community ($T = -2.32$, $P < 0.05$).

Inter-Community Edge Strength

Inter-community edge strength is a measure of connectivity between a pair of Yeo-communities, and can tell us how communication between two communities changes during learning. We predicted that in successful learners, we would observe a greater decoupling of the default community from the dorsal/ventral attention communities. Previous research has indicated that decoupling of default and attention systems is associated with better working memory and overall fluid intelligence scores (Cole et al. 2012; Keller et al. 2015). Inter-community edge strength is plotted over time for each pair of Yeo-communities and for both successful and unsuccessful learners in Figure 6. Using two linear mixed effects models, we tested the effects of learning, time, and the interaction between learning and time on inter-community edge strength (Model 1), as well as the effect of time-varying accuracy on inter-community edge strength while controlling for the effect of overall accuracy (Model 2). Model results are presented in Table 4. The results support our hypothesis, showing that successful subjects had greater decoupling between default and attention communities. However, the results indicate that this is largely driven by an overall decoupling of the dorsal and ventral attention communities from most other Yeo-communities in successful learners, rather than a specific decoupling of attention systems from the default community. Model 1 showed that compared to unsuccessful learners, the dorsal attention community in successful learners showed significantly reduced inter-community edge strength with visual ($T = -2.71$, $P < 0.05$), somatomotor ($T = -3.67$, $P < 0.01$), ventral attention ($T = -2.92$, $P < 0.05$), and limbic ($T = -3.01$, $P < 0.05$) communities. Reduced inter-community edge strength between the dorsal attention and default communities was only significant before FDR correction for multiple comparisons ($T = -2.40$, $P = 0.06$). Similarly, the ventral attention community

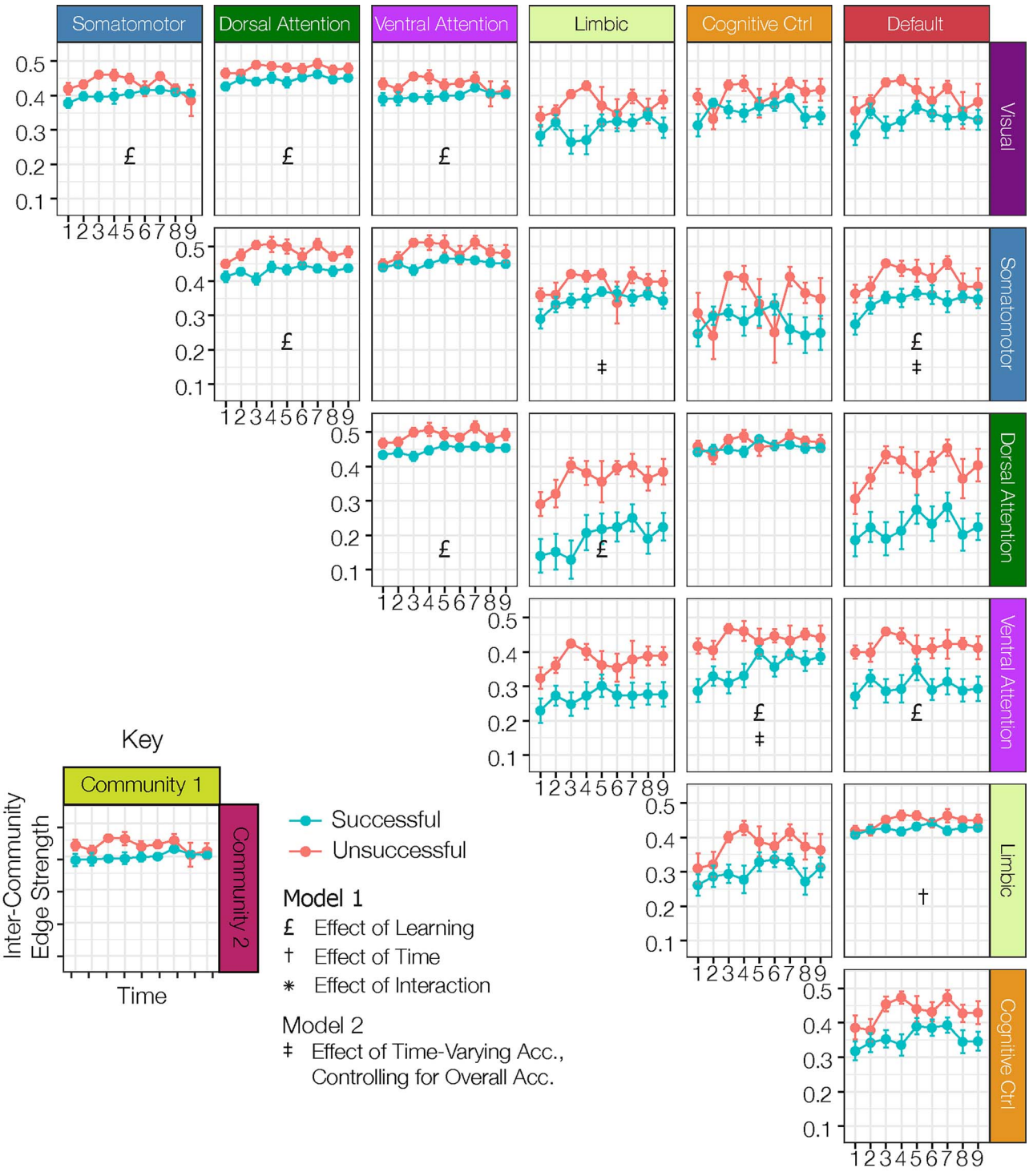

Figure 6. Inter-community edge strength. Inter-community edge strength plotted over time for each pair of Yeo-communities. Points represent mean inter-community strength across subjects in each group: successful learners are colored blue and unsuccessful learners are colored red. Time points represent each of the 9 scanning runs (32 trials per run). Error bars represent standard error. Two linear mixed effects models were used to assess the effects of learning group and time (Model 1) as well as the effect of run-to-run accuracy (Model 2) on the network measures. Model results are presented in Table 4. £ = Effect of Learning Group; † = Effect of Time; ∗ = Effect of Learning Group × Time Interaction; ‡ = Effect of Run-to-Run Accuracy, Controlling for Overall Accuracy ($P < 0.05$, FDR corrected).

in successful learners showed significantly reduced inter-community edge strength with visual ($T = -3.01$, $P < 0.05$), dorsal attention ($T = -2.92$, $P < 0.05$), cognitive control ($T = -3.58$, $P < 0.01$), and default ($T = -2.55$, $P < 0.05$) communities. Reduced inter-community edge strength between the ventral attention and limbic communities was only significant before FDR correction ($T = -2.39$, $P = 0.06$). Interestingly, Model 2 showed several pairs of Yeo-communities in which inter-community

Table 4 Linear mixed effects models for inter-community edge strength

Community 1	Community 2	Model 1						Model 2			
		Learning		Time		Learning* Time		Overall Acc.		Acc. Per Run	
		Effect	P	Effect	P	Effect	P	Effect	P	Effect	P
Visual	Somatomotor	**−0.07**	**<0.01**	−0.004	0.16	**0.007**	**<0.05**	**−0.14**	**<0.01**	0.031	0.21
Visual	Dorsal attn.	**−0.04**	**<0.05**	0.001	0.50	0.001	0.72	**−0.13**	**<0.01**	0.032	0.07
Visual	Ventral attn.	**−0.06**	**<0.05**	−0.002	0.39	0.005	0.09	**−0.13**	**<0.05**	0.031	0.21
Visual	Limbic	−0.09	0.09	0.002	0.76	0.004	0.59	**−0.29**	**<0.01**	0.031	0.51
Visual	Cognitive ctrl	−0.03	0.52	0.005	0.32	−0.004	0.56	**−0.18**	**<0.05**	0.015	0.72
Visual	Default	−0.09	0.09	−0.001	0.86	0.004	0.48	**−0.32**	**<0.01**	0.038	0.38
Somatomotor	Dorsal attn.	**−0.06**	**<0.01**	0.002	0.56	0.001	0.72	**−0.19**	**<0.01**	0.022	0.35
Somatomotor	Ventral attn.	−0.04	0.13	0.002	0.31	0	0.90	**−0.13**	**<0.01**	0.016	0.44
Somatomotor	Limbic	−0.06	0.155	0.003	0.50	0.003	0.63	**−0.25**	**<0.01**	**0.102**	**<0.01**
Somatomotor	Cognitive ctrl	−0.01	0.863	0.006	0.47	−0.01	0.32	−0.08	0.59	−0.08	0.27
Somatomotor	Default	**−0.09**	**<0.05**	0.001	0.86	0.005	0.33	**−0.36**	**<0.01**	**0.114**	**<0.01**
Dorsal attn.	Ventral attn.	**−0.05**	**<0.05**	0.002	0.28	0.001	0.76	**−0.16**	**< 0.01**	0.031	0.10
Dorsal attn.	Limbic	**−0.19**	**<0.05**	0.009	0.28	0.003	0.76	**−0.64**	**<0.01**	0.076	0.29
Dorsal attn.	Cognitive ctrl	−0.01	0.788	0.003	0.25	−0.001	0.76	−0.08	0.08	0.02	0.37
Dorsal attn.	Default	−0.16	0.06	0.007	0.32	−0.002	0.84	**−0.67**	**<0.01**	0.046	0.48
Ventral attn.	Limbic	−0.11	0.061	0.003	0.59	0.001	0.90	**−0.40**	**<0.01**	0.018	0.72
Ventral attn.	Cognitive ctrl	**−0.13**	**<0.01**	0.003	0.63	0.009	0.10	**−0.41**	**<0.01**	**0.106**	**<0.05**
Ventral attn.	Default	**−0.12**	**<0.05**	0	0.95	0	0.97	**−0.46**	**<0.01**	0.003	0.94
Limbic	Cognitive ctrl	−0.07	0.204	0.006	0.32	−0.001	0.90	**−0.28**	**<0.05**	0.023	0.64
Limbic	Default	−0.01	0.495	**0.003**	**<0.05**	−0.002	0.49	−0.07	0.08	0.009	0.59
Cognitive ctrl	Default	−0.07	0.158	0.005	0.26	−0.001	0.86	**−0.32**	**<0.01**	0.05	0.29

Notes: Results from two linear mixed effects models that were run for each pair of the seven Yeo-communities. Model 1 assessed the effects of learning, time, and the interaction between learning and time, on inter-community edge strength. Model 2 assessed the effects of overall accuracy and time-varying accuracy (on a per-run basis) on inter-community edge strength. *P* = adjusted *P*-value (FDR corrected). Bold results show a statistically significant effect, $P < 0.05$, FDR corrected for multiple comparisons.

edge strength was strongly associated with time-varying accuracy on the task. These pairs were somatomotor—default ($T = 3.39$, $P < 0.01$), somatomotor—limbic ($T = 3.06$, $P < 0.01$), and ventral attention—cognitive control ($T = 2.58$, $P < 0.05$).

Discussion

We performed a dynamic network analysis to identify functional brain network changes associated with individual differences in learning a set of context-dependent rules. We propose that a stable ventral attention community supports maintenance of sustained attention in successful learners, and that dynamic changes in assortativity and inter-community edge strength of the cognitive control community are associated with the formation of context-dependent rule representations, and task automaticity. These results highlight key brain network features that are associated with successful learning.

Network Stability

In line with our predictions, we found that participants with greater overall accuracy on the task exhibited decreased average community switching (flexibility) within the ventral attention and somatomotor communities (see Figure 4C, Table 1). We view decreased network flexibility as increased network stability, and in successful subjects this stability correlates with their high accuracy and representation of the context-dependent rules.

Considering the ventral attention community, previous work on rule learning has identified vlPFC and pre-SMA as regions important for representing specific rules, and suppressing activity from a previous task set respectively (Crone, Donohue, et al. 2006a; Crone, Wendelken, et al. 2006b). Moreover, recently published work in non-human primates shows that neuronal population activity in the dorsal anterior cingulate cortex, another prominent node in the ventral attention community, reflects the search for new rules in a categorical rule-learning task (Cohen et al. 2021). For successful learners on our task, it is likely beneficial for these regions to have stable connectivity and reduced community switching as they actively represent learned rules. Considering the somatomotor community, previous work using recordings from motor cortex in non-human primates demonstrates that rules are most easily learned when they are aligned with the intrinsic network activity of this region (Sadtler et al. 2014). Reduced flexibility observed in our successful subjects could reflect a reduced change in somatomotor community activity during rule implementation compared to rest.

Our results complement previous studies of network flexibility where lower overall flexibility was observed in somatomotor cortex compared to other brain regions (Bassett et al. 2011; Betzel et al. 2017; Gerraty et al. 2018). Several studies have also identified a positive association between learning and flexibility in frontoparietal cortical regions (Bassett et al. 2011; Braun et al. 2015; Gerraty et al. 2018). While not statistically significant in our study, we did observe a positive association between flexibility and accuracy in the cognitive control community ($r = 0.22$, $T = 1.81$, $P = 0.22$). In our study, we identified two communities that showed greater stability associated with overall accuracy on the task. This is in line with several large-scale studies that have identified network stability as a key feature contributing

to general intelligence and overall cognitive ability (Schultz and Cole 2016; Ferguson et al. 2017).

Previously, studies of network flexibility and learning have employed homogeneous samples of learners whose learning improved over time (Bassett et al. 2011; Gerraty et al. 2018). Our study expands upon previous work with the inclusion of unsuccessful subjects who never learned the context dependent rules. These subjects comprised a natural control group that allowed us to identify individual differences in brain network structure associated with learning. Additionally, in our study we capture the full learning curves of successful subjects, from the beginning phases of learning, up until performance has plateaued and high accuracy is achieved.

While we highlight the reduced flexibility of the ventral attention community as a key feature of successful learning, we also found changes in the centrality and inter-community edge strength of this community that are associated with learning. It is not necessarily true that a node with low flexibility has unchanging functional connectivity. Rather, a node with low flexibility is part of a stable community. With Louvain community detection in particular, emphasis is put on creating modular networks, where intra-community connections are stronger than inter-community connections (Mucha et al. 2010). A stable community might exhibit low flexibility (the nodes within the community stay a part of that community), but be changing its connectivity with other communities. The community maintains low flexibility by ensuring that all members alter their connectivity patterns together.

Cognitive Control

In the cognitive control community, we observed a significant interaction effect between learning group and time on assortative mixing (Table 2, Model 1). Assortative mixing showed a greater increase over time in successful learners than in unsuccessful learners. Complementing this result, we also found a significant positive association between assortative mixing in the cognitive control community and run-to-run accuracy on the task (Table 2, Model 2).

Brain regions within the cognitive control community are understood to be responsible for higher order executive functioning, including response inhibition, working memory, and planning (Vincent et al. 2008; Nyhus and Barceló 2009; Watson and Chatterjee 2012). A large body of work proposes that rules are represented in prefrontal cortex for the purpose of directing top-down attention (for review, see Miller and Buschman 2014). Additionally, frontoparietal cortex is implicated in strategic access to memory (Badre and Wagner 2007) and memory guided attention (Rosen et al. 2016).

Successful context-dependent rule learning on our task could be conceptualized as the ability to create a task-set structure during uninstructed learning. Modeling work suggests that the creation of task sets could be accomplished through cognitive control and executive functioning (Collins and Frank 2013). A potential mechanism by which the cognitive control community could direct attention is by mediating the decoupling of default and dorsal attention communities, and titrating the balance between inward and outward awareness (Spreng et al. 2013). For successful learners, the context-dependent rule task quickly became automatic, and an increase in assortativity may be reflective of decreased dependence on, or more efficient processing within the cognitive control community for directing top-down attention. This framework is in line with theories suggesting that cognitive control connectivity adaptively reconfigures to meet task demands, primarily through a set of connector hubs (Gratton et al. 2016; Gordon et al. 2018; Gratton, Sun, et al. 2018a; Gratton, Cooper, et al. 2018b). As a working hypothesis, we propose that as learning occurs, executive control over network configuration becomes more efficient and the cognitive control community becomes more strongly connected to itself and less widely integrated in the whole brain network. Similarly, in unsuccessful learners, decreased assortative mixing of the cognitive control community over time is reflective of continual strategy updating as attempts at implementing rules remain unsuccessful.

Interactions between Attention and Cognitive Control

We observed a positive association between run-to-run accuracy on the task and inter-community edge strength of the cognitive control and ventral attention communities (Table 4, Model 2). Additionally, successful learners exhibited significantly lower inter-community edge strength between attention communities and most other communities throughout the brain (Table 4, Model 1).

The vlPFC, pre-SMA, and dorsal anterior cingulate are prominent regions in the ventral attention community that contribute to rule learning (Passingham et al. 2000; Crone, Donohue, et al. 2006a; Crone, Wendelken, et al. 2006b; Cohen et al. 2021). The ventral attention community is often associated with orienting, as it is activated by novel stimuli and stimulus-driven shifts in attention (Corbetta and Shulman 2002; Mantini et al. 2009; Kim 2014). While the Schaefer-400 atlas includes the vlPFC, pre-SMA, and dorsal anterior cingulate as part of the "ventral attention" community, other groups use varying labels and definitions (Power et al. 2011; Yeo et al. 2011). One common framework refers to the ventral attention community as the cingulo-opercular network, and suggests that both the cognitive control community and cingulo-opercular network work together to achieve executive control (Gratton et al. 2016; Gratton, Sun, et al. 2018a).

Research from Crittenden and colleagues showed increased activity in both cingulo-opercular and cognitive control regions during a rule-based associative memory task (Crittenden et al. 2016). Interestingly, they found that specific task rules were most accurately decoded using signals from voxels within cognitive control regions (which they call the frontoparietal network). This led them to suggest that cognitive control regions contain more specific rule-based representations, while cingulo-opercular regions are more important for emphasizing event salience and maintaining focused attention (Crittenden et al. 2016). This interpretation complements other work conceptualizing the ventral attention community as a salience network (Seeley et al. 2007) that plays a top-level role directing changes in functional connectivity between cognitive control and other communities throughout the brain (Sridharan et al. 2008; Menon and Uddin 2010). Our study adds new evidence to this conceptual framework, demonstrating that a stable ventral attention community contributes to steady maintenance of attention throughout the task, and a more flexible cognitive control community demonstrates dynamic changes in assortativity and inter-community edge strength associated with successful rule learning.

We also observed a negative association between the centrality of the ventral attention community and run-to-run accuracy on the task (Table 3, Model 2). This suggests that as

subjects learn the context-dependent rules, ventral attention regions reduce their long-ranging functional connections with other brain regions. Within the frameworks discussed above, decreased centrality of the ventral attention community could be reflective of more targeted salience to specific features of the task. In concert with the increased inter-community edge strength we observe between ventral attention and cognitive control regions, these results suggest more targeted attention to rule representations in successful learners.

Finally, we also observed reduced inter-community edge strength between the dorsal/ventral attention communities and most other communities in successful learners throughout the entire task (Table 4, Model 1). This observation reflects the well-known dissociation between default and attention systems of the brain that is present at rest (Fox et al. 2005) and is modulated by various tasks (Spreng et al. 2013; Keller et al. 2015; Dixon et al. 2017). Miller and Buschman (2014) suggest that local synchrony of neurons in attention and primary sensory areas is essential for synchronizing mental representations with the external world. More broadly, functional segregation of attention communities could be advantageous for maintaining focused attention during learning.

Conclusion

Our work expands upon previous studies of rule learning by using a dynamic network analysis and focusing on brain-wide connectivity. We identified several brain network characteristics associated with individual differences in learning context-dependent rules. These include decreased flexibility of the ventral attention and somatomotor communities, functionally segregated attention systems, and shifts in the cognitive control connectivity associated with successful learning.

Citation Diversity Statement

Recent work in neuroscience has identified a bias in citation practices such that manuscripts written by women and other minorities are under-cited relative to the number of such papers in the field (Dworkin, Zurn, et al. 2020a; Dworkin, Linn, et al. 2020b). Here, we quantify the citation diversity of the present manuscript excluding self-citations of the first and last authors of this manuscript. Our reference list contains 65.0% man–man (first author-last author), 20.0% woman–man, 10.0% man–woman, and 5.0% woman–woman citations. This method is limited in that 1) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identify and 2) it cannot account for intersex, non-binary, or transgender people.

Funding

Office of Naval Research (MURI N00014-16-1-2832, DURIP N00014-17-2304); National Science Foundation (1625552).

Notes

The authors would like to thank Matthew F. Dunne, Stamatios S.P. Liapis, Kylie N. Moore, Caroline Ahn, Dr Rachel K. Nauer, Dr James A. Brissenden, Dr Sean M. Tobyne, Dr Michael Hasselmo, and Dr Eric Kolaczyk for their comments and feedback during the preparation of this manuscript. *Conflict of Interest*: The authors declare no competing financial interests.

References

Badre D, Wagner AD. 2007. Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*. 45:2883–2901.

Bassett DS, Mattar MG. 2017. A network neuroscience of human learning: potential to inform quantitative theories of brain and behavior. *Trends Cogn Sci*. 21:250–264.

Bassett DS, Wymbs NF, Porter MA, Mucha PJ, Carlson JM, Grafton ST. 2011. Dynamic reconfiguration of human brain networks during learning. *Proc Natl Acad Sci USA*. 108:7641–7646.

Bassett DS, Yang M, Wymbs NF, Grafton ST. 2015. Learning-induced autonomy of sensorimotor systems. *Nat Neurosci*. 18:744–751.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B*. 57:289–300.

Betzel RF, Satterthwaite TD, Gold JI, Bassett DS. 2017. Positive affect, surprise, and fatigue are correlates of network flexibility. *Sci Rep*. 7:1–10.

Braun U, Schäfer A, Walter H, Erk S, Romanczuk-Seiferth N, Haddad L, Schweiger JI, Grimm O, Heinz A, Tost H, et al. 2015. Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proc Natl Acad Sci USA*. 112:11678–11683.

Bullmore E, Fadili J, Maxim V, Şendur L, Whitcher B, Suckling J, Brammer M, Breakspear M. 2004. Wavelets and functional magnetic resonance imaging of the human brain. *NeuroImage*. 23:S234–S249.

Carp J. 2013. Optimizing the order of operations for movement scrubbing: comment on Power et al. *Neuroimage*. 76:436–438.

Cohen Y, Schneidman E, Paz R. 2021. The geometry of neuronal representations during rule learning reveals complementary roles of cingulate cortex and putamen. *Neuron*. 109:839–851.e9.

Cole MW, Schneider W. 2007. The cognitive control network: integrated cortical regions with dissociable functions. *Neuroimage*. 37:343–360.

Cole MW, Yarkoni T, Repovš G, Anticevic A, Braver TS. 2012. Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *J Neurosci*. 32:8988–8999.

Collins AGE, Frank MJ. 2013. Cognitive control over learning: creating, clustering, and generalizing task-set structure. *Psychol Rev*. 120:190–229.

Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*. 3:201–215.

Crittenden BM, Mitchell DJ, Duncan XJ. 2016. Task encoding across the multiple demand cortex is consistent with a frontoparietal and cingulo-opercular dual networks distinction. *J Neurosci*. 26:6147–6155.

Cromer JA, Machon M, Miller EK. 2011. Rapid association learning in the primate prefrontal cortex in the absence of behavioral reversals. *J Cogn Neurosci*. 23:1823–1828.

Crone EA, Donohue SE, Honomichl R, Wendelken C, Bunge SA. 2006a. Brain regions mediating flexible rule use during development. *J Neurosci*. 26:11239–11247.

Crone EA, Wendelken C, Donohue SE, Bunge SA. 2006b. Neural evidence for dissociable components of task-switching. *Cereb Cortex*. 16:475–486.

Dixon ML, Andrews-Hanna JR, Spreng RN, Irving ZC, Mills C, Girn M, Christoff K. 2017. Interactions between the default network and dorsal attention network vary across default subsystems, time, and cognitive states. *Neuroimage*. 147:632–649.

Dworkin J, Zurn P, Bassett DS. 2020a. (In)citing action to realize an equitable future. *Neuron*. 106:890–894.

Dworkin JD, Linn KA, Teich EG, Zurn P, Shinohara RT, Bassett DS. 2020b. The extent and drivers of gender imbalance in neuroscience reference lists. *Nat Neurosci*. 23:918–926.

Fatima Z, Kovacevic N, Misic B, McIntosh AR. 2016. Dynamic functional connectivity shapes individual differences in associative learning. *Hum Brain Mapp*. 37:3911–3928.

Ferguson MA, Anderson JS, Spreng RN. 2017. Fluid and flexible minds: intelligence reflects synchrony in the brain's intrinsic network architecture. *Netw Neurosci*. 1:192–207.

Fischl B. 2012. FreeSurfer. *Neuroimage*. 62:774–781.

Fornito A, Zalesky A, Breakspear M. 2015. The connectomics of brain disorders. *Nat Rev Neurosci*. 16:159–172.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *PNAS*. 5:9673–9678.

Freeman LC. 1978. Centrality in social networks conceptual clarification. *Soc Networks*. 1:215–239.

Gerraty RT, Davidow JY, Foerde K, Galvan A, Bassett DS, Shohamy D. 2018. Dynamic flexibility in striatal-cortical circuits supports reinforcement learning. *J Neurosci*. 38: 2442–2453.

Gordon EM, Lynch CJ, Gratton C, Petersen SE, Dosenbach NUF, Nelson Correspondence SM. 2018. Three distinct sets of connector hubs integrate human brain function. *Cell Rep*. 24:1687–1695.

Gratton C, Laumann TO, Gordon EM, Adeyemo B, Petersen SE. 2016. Evidence for two independent factors that modify brain networks to meet task goals. *Cell Rep*. 17:1276–1288.

Gratton C, Sun H, Petersen SE. 2018a. Control networks and hubs. *Psychophysiology*. 55:1–18.

Gratton G, Cooper P, Fabiani M, Carter CS, Karayanidis F. 2018b. Dynamics of cognitive control: theoretical bases, paradigms, and a view for the future. *Psychophysiology*. 55:1–29.

Hearne LJ, Cocchi L, Zalesky A, Mattingley JB. 2017. Reconfiguration of brain network architectures between resting-state and complexity-dependent cognitive reasoning. *J Neurosci*. 37:8399–8411.

Hoshi E, Shima K, Tanji J. 2000. Neuronal activity in the primate prefrontal cortex in the process of motor selection based on two behavioral rules. *J Neurophysiol*. 83:2355–2373.

Keller CJ, Bickel S, Honey CJ, Groppe DM, Entz L, Craddock RC, Lado FA, Kelly C, Milham M, Mehta AD. 2013. Neurophysiological investigation of spontaneous correlated and anticorrelated fluctuations of the BOLD signal. *J Neurosci*. 33:6333–6342.

Keller JB, Hedden T, Thompson TW, Anteraper SA, Gabrieli JDE, Whitfield-Gabrieli S. 2015. Resting-state anticorrelations between medial and lateral prefrontal cortex: association with working memory, aging, and individual differences. *Cortex*. 64:271–280.

Kim H. 2014. Involvement of the dorsal and ventral attention networks in oddball stimulus processing: a meta-analysis. *Hum Brain Mapp*. 35:2265–2284.

Kolaczyk E. 2009. *Statistical analysis of network data—methods and models*. New York (NY): Springer Science+Business Media.

Mansouri FA, Freedman DJ, Buckley MJ. 2020. Emergence of abstract rules in the primate brain. *Nat Rev Neurosci*. 21:595–610.

Mantini D, Corbetta M, Perrucci MG, Romani GL, Del Gratta C. 2009. Large-scale brain networks account for sustained and transient activity during target detection. *Neuroimage*. 44:265–274.

Menon V, Uddin LQ. 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct*. 214:655–667.

Miller EK, Buschman TJ. 2014. Neural mechanisms for the executive control of attention oxford handbooks online neural mechanisms for the executive control of attention. In: *The Oxford handbook of attention*. Oxford (UK): Oxford University Press.

Mucha PJ, Richardson T, Macon K, Porter MA, Onnela JP. 2010. Community structure in time-dependent, multiscale, and multiplex networks. *Science (80-)*. 328:876–878.

Newman MEJ. 2003. Mixing patterns in networks. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*. 67:13.

Nyhus E, Barceló F. 2009. The Wisconsin card sorting test and the cognitive assessment of prefrontal executive functions: a critical update. *Brain Cogn*. 71:437–451.

Passingham RE, Toni I, Rushworth MFS. 2000. Specialisation within the prefrontal cortex: the ventral prefrontal cortex and associative learning. *Exp Brain Res*. 133:103–113.

Pasupathy A, Miller EK. 2005. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature*. 433:873–876.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*. 59:2142–2154.

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, et al. 2011. Functional network Organization of the Human Brain. *Neuron*. 72:665–678.

Rosen ML, Stern CE, Michalka SW, Devaney KJ, Somers DC. 2016. Cognitive control network contributions to memory-guided visual attention. *Cereb Cortex*. 26:2059–2073.

Rubinov M, Sporns O. 2011. Weight-conserving characterization of complex functional brain networks. *Neuroimage*. 56:2068–2079.

Sadtler PT, Quick KM, Golub MD, Chase SM, Ryu SI, Tyler-Kabara EC, Yu BM, Batista AP. 2014. Neural constraints on learning. *Nature*. 512:423–426.

Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, Eickhoff SB, Yeo BTT. 2018. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb Cortex*. 28:3095–3114.

Schultz DH, Cole XW. 2016. Higher intelligence is associated with less task-related brain network reconfiguration. *J Neurosci*. 36:8551–8561.

Schwarz AJ, McGonigle J. 2011. Negative edges and soft thresholding in complex network analysis of resting state functional connectivity data. *Neuroimage*. 55:1132–1146.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD. 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci*. 27:2349–2356.

Setsompop K, Gagoski BA, Polimeni JR, Witzel T, Wedeen VJ, Wald LL. 2012. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magn Reson Med*. 67:1210–1224.

Spreng RN, Sepulcre J, Turner GR, Stevens WD, Schacter DL. 2013. Intrinsic architecture underlying the relations among the default, dorsal attention, and frontoparietal control networks of the human brain. *J Cogn Neurosci*. 25:74–86.

Sridharan D, Levitin DJ, Menon V. 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc Natl Acad Sci USA*. 105:12569–12574.

Van Dijk KRA, Hedden T, Venkataraman A, Evans KC, Lazar SW, Buckner RL. 2010. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J Neurophysiol*. 103:297–321.

Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL. 2008. Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *J Neurophysiol*. 100: 3328–3342.

Wallis JD, Anderson KC, Miller EK. 2001. Single neurons in prefrontal cortex encode abstract rules. *Nature*. 411: 953–956.

Watson CE, Chatterjee A. 2012. A bilateral frontoparietal network underlies visuospatial analogical reasoning. *Neuroimage*. 59:2831–2838.

Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, et al. 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol*. 106:1125–1165.

Zhang Z, Telesford QK, Giusti C, Lim KO, Bassett DS. 2016. Choosing wavelet methods, filters, and lengths for functional brain network construction. *PLoS One*. 11:e0157243.

Zhu H, Paschalidis IC, Chang A, Stern CE, Hasselmo ME. 2020. A neural circuit model for a contextual association task inspired by recommender systems. *Hippocampus*. 30: 384–395.