



Published in final edited form as:

*Neuroimage*. 2020 June ; 213: 116687. doi:10.1016/j.neuroimage.2020.116687.

## Temporal fluctuations in the brain's modular architecture during movie-watching

Richard F. Betzel<sup>1,2,3,4,\*</sup>, Lisa Byrge<sup>1</sup>, Farnaz Zamani Esfahlani<sup>1</sup>, Daniel P. Kennedy<sup>1,2,3</sup>

<sup>1</sup>Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405

<sup>2</sup>Cognitive Science Program, Indiana University, Bloomington, IN 47405

<sup>3</sup>Program in Neuroscience, Indiana University, Bloomington, IN 47405

<sup>4</sup>Network Science Institute, Indiana University, Bloomington, IN 47405

### Abstract

Brain networks are flexible and reconfigure over time to support ongoing cognitive processes. However, tracking statistically meaningful reconfigurations across time has proven difficult. This has to do largely with issues related to sampling variability, making instantaneous estimation of network organization difficult, along with increased reliance on task-free (cognitively unconstrained) experimental paradigms, limiting the ability to interpret the origin of changes in network structure over time. Here, we address these challenges using time-varying network analysis in conjunction with a naturalistic viewing paradigm. Specifically, we developed a measure of inter-subject network similarity and used this measure as a coincidence filter to identify synchronous fluctuations in network organization across individuals. Applied to movie-watching data, we found that periods of high inter-subject similarity coincided with reductions in network modularity and increased connectivity between cognitive systems. In contrast, low inter-subject similarity was associated with increased system segregation and more rest-like architectures. We then used a data-driven approach to uncover clusters of functional connections that follow similar trajectories over time and are more strongly correlated during movie-watching than at rest. Finally, we show that synchronous fluctuations in network architecture over time can be linked to a subset of features in the movie. Our findings link dynamic fluctuations in network integration and segregation to patterns of inter-subject similarity, and suggest that moment-to-moment fluctuations in functional connectivity reflect shared cognitive processing across individuals.

---

\* rbetzel@indiana.edu.

RFB and DPK conceptualized the project. LB and DPK collected data. RFB and DPK wrote the original draft. All authors helped with data analysis. All authors helped review and edit the manuscript.

#### AUTHOR CONTRIBUTIONS

All authors conceived of the study. LB and DPK designed experimental protocols, collected, and preprocessed all fMRI data. RFB and FZE performed all analyses. All authors wrote and approved of the manuscript.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

#### CODE AVAILABILITY

All data and code are available from the authors upon reasonable request.

## INTRODUCTION

Cognitive processes are supported by the transient coupling and uncoupling of activity between distant brain regions [1, 2]. These patterns can be modeled as networks, whose nodes and edges represent regions and their pairwise functional connectivity (FC) [3]. Brain networks can then be investigated using methodology from network science [4, 5], a discipline that provides both a conceptual and mathematical framework for investigating the architecture and dynamics of real-world networks.

Most FC analyses have focused on intrinsic connectivity, which can be reconstructed from brain activity recorded during rest, i.e. task-free conditions [6]. One of the most salient features of resting FC is its modular organization, which refers to the decomposability of the brain into segregated sub-networks called “modules” or “communities” [7, 8]. The boundaries of these modules closely recapitulate patterns of task-evoked activity [9], are highly replicable [10–12], and are thought to engender specialized brain function [13].

Interestingly, the level of segregation in the brain can be modulated. One contributing factor is cognitive state; performing cognitively-demanding tasks requires coordination between modules that at rest are functionally isolated from one another [14]. This constraint results in increased inter-modular FC, which effectively decreases segregation. This phenomenon is general, and has been reported across a wide range of tasks [15–18].

Another factor that can modulate segregation is time. So-called “time-varying” FC can be estimated by partitioning a scan session into a series of windows and estimating FC separately for each window, making it possible to track changes in network architecture over short timescales [19, 20]. Using this approach, many studies have reported temporal fluctuations in the brain’s level of segregation, suggesting that modules transiently couple and uncouple with one another over time [21–23]. This has given rise to the hypothesis that the brain alternates between segregated and integrated states, reflecting periods of local, specialized information processing and inter-modular information transfer, respectively [17, 23–25]

Unlike task FC, where variation in segregation can be attributed to task-related cognitive processes, time-varying FC analysis is usually applied to resting-state datasets in which subjects’ cognitive state is not experimentally controlled. This makes it difficult to ascertain whether fluctuations in segregation are driven by coincident fluctuations in neurophysiological variables like drowsiness or attention [26, 27], internal cognitive processes [28–30], or merely reflect sampling variability around a temporally stationary pattern of FC [31, 32].

Indeed, disambiguating meaningful fluctuations in time-varying FC from statistical noise during rest [33] and tasks [34] remains an open challenge. Here, we propose addressing this question using naturalistic imaging, which involves presenting subjects with short videos (though for consistency with the literature we refer to these as “movies” throughout this manuscript) during the course of a scan session. Movie-watching can be considered a state situated somewhere between that of rest and task, in which subjects are presented with

identical stimuli (akin to task) but are not instructed to respond in any particular way (akin to rest).

Recently, movie-watching paradigms have become popular within the neuroimaging [35–38] and network neuroscience communities [39–41], where FC estimated from movie-watching exhibits greater test-retest reliability [42], benefits from reduced in-scanner head motion [43, 44], and enhances the identifiability of individual subjects [45] compared to rest. Additionally, movies feature correlated categories that better match the statistical properties of our day-to-day experiences. For example, rather than viewing a disembodied human face, faces that appear in a movie are usually accompanied by speech, movement, narrative, and context.

Recent studies have attempted to decode the network correlates of ongoing cognitive processes. Using measures like inter-subject FC to track cohesive, population-level fluctuations in connections [46], these studies have sought to identify patterns of covariation between edges and features present in the movie [47, 48]. While these and related approaches hold promise for identifying shared drivers of time-varying FC, the inter-subject FC measure makes them less-suited for studying the topological properties of FC from individual subjects. Consequently, the principles that govern fluctuations in FC during movie-watching have not been fully elucidated.

To address this question, we estimate time-varying FC by applying sliding-window analysis to movie-watching data. We take advantage of the fact that subjects observe the same time-locked stimuli to focus on moments in time when FC across subjects coalesces into similar patterns. We develop a statistical test for identifying those periods, and discover that high inter-subject similarity is linked to decreases in segregation, as indexed by the modularity measure. Next, we show that time-varying FC during movie-watching can be described in terms of edge clusters – groups of connections that respond similarly over the course of the scan. Finally, we link these patterns of fluctuations to features present in the movie. Collectively, our findings represent a conceptual bridge between network analysis of time-varying connectivity and naturalistic imaging, and strengthens the link between studies of inter-subject similarity with cognition.

## RESULTS

We analyzed fMRI data from 29 subjects each of which underwent eight scans (four rest and four movie-watching) on two separate sessions, totaling approximately two hours worth of data for every subject. For both scan categories, we estimate time-averaged (static) and time-varying whole-brain FC. Complete methodological details can be found in the section entitled Materials and Methods. In this section, we summarize the results of several analyses. First, we compared time-averaged FC between the conditions, identifying differences at the level of individual connections, but also in terms of segregation and integration. Next, we examined time-varying FC and designed a statistical procedure to identify temporal windows during which subjects' networks become highly similar to each other. We investigated network structure within these windows and discovered that high levels of inter-subject similarity corresponded to decreased modularity and increased

dissimilarity with respect to resting FC. Next, we investigate the edge-level correlates of movie-watching, leveraging hypergraph clustering to uncover constellations of connections that follow similar trajectories during movie-watching. Lastly, we demonstrated that fluctuations in time-varying FC were differentially associated with the presence and absence of specific features in the movie, suggesting possible drivers of brain state changes.

### Time-averaged FC during rest and movie-watching conditions

Time-averaged FC is thought to reflect the strength of communication between pairs of brain regions. The magnitude of coupling can be modulated by sensory input, stimulation, or task constraints. In this section we compare rest and movie-watching to assess what effect, if any, movie-watching has on whole-brain patterns of FC.

In Fig. 1a and Fig. 1b, we show the session-, subject-, and time-averaged FC for the resting and movie-watching conditions, respectively. Visually, these two patterns appear highly similar (the correlation of their upper triangle elements confirms as much  $r = 0.69$ ;  $p < 10^{-5}$ ). However, when we compute the average change in each connection's weight, we find systematic differences between the two conditions (Fig. 1c; we show a system-averaged version of the same plot in Fig. S1a)<sup>1</sup>. To assess the statistical significance of differences, we compared the observed differences to those obtained using a permutation-based null model in which we randomly permuted condition labels ("rest" and "movie-watching"; 10,000 permutations; permutations always within subjects). In general, we found that most connections (>60%) exhibited a statistically significant difference between conditions (Fig. 1d; false discovery rate fixed at  $q = 0.01$ ;  $p_{adjusted} = 0.0062$ ).

In general, these differences were distributed across the brain and involved multiple cognitive systems [49]. To help visualize these differences, we generated force-directed embeddings of the task-free and movie-watching networks (Fig. 1e and Fig. 1f). In both plots, we used arrows to draw attention to particularly salient differences. For instance, in task-free conditions we found that sub-components of the salience and default networks appeared segregated from one another, but that during movie-watching FC between those systems increased so that they were more strongly connected and integrated (Fig. 1e,i). Similarly, both sub-components of the visual system appeared highly segregated at rest (Fig. 1e,ii), but became more integrated during movie-watching. Other interesting differences included the integration of default mode with visual sub-components (Fig. 1e,iii) and the dissolution and distribution of default mode components throughout the network (Fig. 1e,iv).

These previous analyses were carried out at the level of individual connections. We also calculated changes in modularity and participation coefficients, measures that index the level of segregation among network modules [50] and the extent to which nodes' connections are distributed across module boundaries [51], respectively. In agreement with past studies showing that network segregation decreases during tasks [15], we found that modularity,  $q^*$  is greater during rest than movie-watching (within-subject  $t$ -test,  $p < 0.01$ ; Fig. 1g). Specifically, we calculated the difference in  $q^*$  for movie-watching and rest conditions

<sup>1</sup>For this analysis we truncated the fMRI BOLD time series for each session so that FC was always estimated using the same number of observations.

separately for each subject and tested the null hypothesis that the mean difference in  $q^*$  was equal zero. We also found that, as a result of changes in FC, nodes have repositioned themselves with respect to cognitive systems, with areas in the default mode and visual systems forming stronger inter-modular connections during movie-watching than at rest, suggesting that they are becoming more “hub-like” (although we note that participation coefficient is one of many measures used to assess a node’s hubness [52] (Fig. 1*h* and Fig. 1*j*).

In summary, the findings presented here suggest that movie-watching drives the brain into a more integrated and less modular network architecture. While visual and somatomotor systems are among those whose position within the network changes, differences between the two conditions are distributed and manifest within and across virtually all brain areas and systems. This is an important observation; during movie-watching subjects receive visual and auditory stimuli. One possibility is that differences in network architecture are localized to brain systems associated with the processing of those sensory modalities. However, the involvement of DMN and higher-order association cortices suggest that movie-watching taps into a constellation of brain systems, likely involved in processing more than low-level sensory stimuli.

### Inter-subject similarity and time-varying FC

The previous analysis was carried out on time-averaged FC and revealed differences in network structure over entire scan sessions. To investigate changes over shorter timescales, we calculated time-varying FC using a sliding window technique. Past analyses of time-varying FC networks have focused mostly on properties of time-varying FC defined at the level of individual subjects. Our focus, on the other hand, is on correlated fluctuations in network structure in which subjects’ networks reconfigure in concert with one another. Correlated fluctuations are more likely to reflect shared network-level responses to stimuli in the movie [38, 53]. To investigate these kinds of fluctuations, we developed a metric of inter-subject similarity (ISS), which assesses at each time point, the mean similarity of subjects’ FC to one another (Fig. 2).

Because FC tends to be similar across individuals, in general ISS will be nonzero, even in the absence of “true” inter-subject synchrony. To distinguish periods of true inter-subject similarity from baseline, we develop a novel statistical testing procedure. Specifically, we compare mean ISS during movie-watching, when subjects are presented with identical time-locked audiovisual stimuli, with mean ISS during resting conditions, when each subject’s FC fluctuates independently over time. In Fig. 3*a* we show the distribution of ISS values across time for one of the movies. We show an analogous plot for the resting condition in Fig. 3*b* (analogous plots for all four movie and resting scans are shown in Fig. S2). Note that the shape of the distribution fluctuates over time, periodically increasing to reflect windows in time when subjects’ FC becomes highly similar. Rest, on the other hand, is characterized by a distribution of ISS values that changes little over time. To identify periods of high synchrony across subjects, we compared mean ISS at each time point during movie-watching with the distribution of mean ISS values during rest. We note that ISS is related to

the concept of inter-subject FC [46] and that, in general, periods of high ISS correspond to periods when inter-subject FC is not equal to zero (Fig. S3).

Using this procedure, we estimated a  $p$ -value at each time point corresponding to the probability that the observed ISS value would occur during rest. This procedure resulted in a time-series of  $p$ -values, which we could use to identify periods when ISS during movie-watching was greater than expected by chance. In Fig. 3c we show negative  $\log p$ -values for the movie-watching condition. Note that there exist periods when the  $p$ -values exceed even stringent statistical thresholds. In contrast, Fig. 3d shows analogous  $p$ -values for one of the task-free scans (the null distribution was estimated from the three other task-free scans to avoid circularity). Note that in this case, the  $p$ -values never exceed the statistical threshold. This observation suggests that, not only does this test distinguish periods of time when subjects exhibited high synchrony during movie-watching, but also demonstrates high specificity by accurately detecting no ISS during rest.

In more detail, when we applied this test to time-varying FC estimated during movie-watching and identified periods of time when ISS was greater than expected by chance (false discovery rate controlled at  $q = 0.01$ ;  $p_{adjusted} = 6.3 \times 10^{-4}$ ). On average across all movies, this procedure detected  $22.8 \pm 14.2$  periods of duration  $3.6 \pm 2.4$  seconds ( $4.5 \pm 2.9$  TRs)<sup>2</sup>. We note that in-scanner head motion (framewise displacement) was not significantly different between periods of high and low ISS (Fig. S4) and that ISS was not sensitive to global signal regression (Fig. S5).

### Inter-subject FC fluctuations are associated with fluctuations in modularity and static resting FC

Periods of high ISS are, by definition, separated by periods of low ISS. What features of brain networks distinguish these periods of time from one another? To address this question, we extracted representative networks for each subject and for every period of low/high ISS. We defined a representative network to be the network within a given low/high ISS period with the greatest average similarity to other networks within the same window, ensuring that representative networks were all estimated using the same number of samples and not biased by the duration of a given low/high ISS time period.

For each representative network we measured two quantities: first, its modularity,  $q^*$  [54], a measure of the level of segregation among a network's communities, and also its similarity to time-averaged FC. In terms of modularity, we found that periods of high ISS were associated with decreased modularity (Fig. 4a;  $p = 0.0008$ ), suggesting that cognitive systems become increasingly integrated with one another during periods when high ISS. At the same time, we found that during periods of high ISS were associated with reduced similarity to time-averaged FC (Fig. 4b,c).

To better visualize changes in network organization during high and low ISS, we projected subjects' FC patterns in two-dimensional space using multidimensional scaling, which reduces the dimensionality of these data while approximately preserving distance structure.

<sup>2</sup>We note that the precise number and duration of these periods will depend upon the choice of statistical threshold.



In Fig. 4*d* and Fig. 4*e* we show those low-dimension embeddings with points colored according to two different criteria. In Fig. 4*d*, each point is colored according to its similarity with respect to the time-averaged resting FC. In Fig. 4*e*, points are colored according to their modularity (z-scored within subjects to remove subject-specific differences in baseline modularity).

Note that the brightest points in both plots largely coincide with one another, indicating that modular networks are also more rest-like. We further quantified this relationship by computing the correlation of modularity with the similarity to time-averaged, task-free FC and found that the two were strongly related to one another (Fig. 4*f*;  $r = 0.67$ ;  $p < 0.01$ ).

We investigated this relationship further by categorizing each point in Fig. 4*e,f* according to whether it corresponded to a period of high or low ISS. We found that periods of low ISS overlapped closely with one another and were concentrated near the middle of the embedding diagram (Fig. 4*g*). Periods of high ISS, on the other hand, were distributed around the border of the embedding (Fig. 4*h*). We show the difference in these density plots in Fig. 4*i*.

In summary, we found that subjects oscillate between periods of high and low similarity, high and low modularity, and network architectures that are rest-like and dissimilar from rest. Surprisingly, all of these changes were correlated with one another. One implication of these findings concerns the link between modularity and cognitive processing. Past studies have reported that when subjects perform cognitively-demanding tasks, their network modularity decreases [15, 16, 18]. Here, we observe that same network-level signature, but because movie-watching is unconstrained, we lack a direct link to cognition. On the other hand, the decreases in modularity took place when network architecture was least similar to rest and subjects were most similar to one another, suggesting that subjects were responding in concert to some aspect of the movie stimulus.

### **Movie-watching organizes functional connections into temporal communities**

In the previous sections, we showed that movie-watching induces changes in network structure compared to rest and that, when FC is resolved at finer temporal scales, FC oscillates between distinct brain states. It remains unclear, however, which functional connections drive this effect and whether those connections exhibit consistent organization across movies. To address this question, we used a data-driven clustering method [50] to uncover groups of connections that not only exhibit similar temporal trajectories across subjects, but also follow similar trajectories across time [55, 56]. This approach involved calculating subject-averaged time-varying FC, yielding a time series for every edge. We then estimated the edge-by-edge correlation matrix, which we clustered using modularity maximization [57, 58] (we chose the resolution parameter to maximize the difference of minimum within-module density minus the maximum between-module density; see Fig. S6). This procedure resulted in the detection of twelve large communities comprised of edges that responded similarly across subjects and were also correlated with each other. We show the results of this procedure in Fig. 5.

Fig. 5a depicts the edge correlation matrix that we used as input to the modularity maximization algorithm ordered according to detected communities. Fig. 5b depicts the same matrix after averaging connection weights within and between pairs of community. We show the edge trajectories in Fig. 5c. Similar trajectories for all movies are shown in Fig. S7.

To further characterize each community, we mapped edges to the nodes upon which they were incident. Because some of a node's edges can be affiliated with one community and other edges affiliated with another, this means that brain regions can participate in multiple communities simultaneously. We show the five largest edge communities mapped back to the level of nodes and onto brain regions in Fig. 5d-h (the remaining communities and community overlap indices are shown Fig. S8, Fig. S9, and Fig. S10). The first community was the largest, and was made up of 2175 edges (34% of all possible) and included mostly higher-order association cortex while excluding sub-components of visual and somatomotor systems (Fig. 5d). The second community (1473 edges or 23%) involved similar brain areas and systems, but a different collection of edges (Fig. 5e). This observation suggests that these higher order association areas are not unifunctional in the context of movie-watching, but play multiple, dissociable roles. Figs. 5f-h depict communities three, four, and five. These communities were comprised of 558, 490, and 435 edges (8.7%, 7.6%, and 6.8% of all connections), and involved default mode and visual systems, default mode and limbic, and the visual system, respectively. We present matrix visualizations of temporal communities in Figs. 5i-m and, for each community, the proportion of edges associated with canonical brain systems that make up each community (Fig. 5n-r).

Collectively, these observations suggest that movie-watching does not induce changes in the weights of single, independent connections. Instead, movie-watching leads to correlated fluctuations among groups of edges. These groups are not simple recapitulations of known cognitive systems, but involve multiple systems, suggesting that movie-watching demands from a fixed neural substrate a multitude of functions. Similar to other studies that reported high levels of variability in somatomotor cortex [59–61], we find that somatomotor cortex is among the systems with greatest participation across communities (Fig. S9). This agrees with other studies that have reported activations in motor cortex under diverse contexts, including observing actions of others [62], planning and execution of imagined movements [63], and social cognition [64]. From a methodological perspective, these findings complement other approaches for identifying patterns in brain network organization [7, 8, 11, 49]. Unlike these past efforts, which imposed non-overlapping partitions over brain regions, the approach used here maintains an advantage in that it allows for brain regions to be affiliated with more than one community.

### Linking inter-subject similarity to movie features

In the previous section, we showed that functional connections followed clustered trajectories across time during movie-watching. However, these results do not directly identify the movie features to which the subjects were responding. In this section, we investigate the relationship of features from the movies to time-varying fluctuations in FC. We use a multi-linear modeling approach to identify statistical associations between connection weight and the following hand-coded features: presence of human (in any



context), human faces and voices (specifically), inter-personal interactions, mean luminance of projected images, and transitions between films (blank screens) (Fig. 6a). This analysis returns a regression coefficient for each edge in the network, indicating the sign and magnitude of correspondence between the feature time-series and edge weight fluctuations.

In general, we found that each feature was associated with a distinct template of connections (Fig. 6b). Most dissimilar was the constellation of edges associated with the black screens that were interspersed between movies, which we label “blank”. Though we found evidence of brain-wide associations, the strongest regression coefficients were concentrated within the peripheral visual system (Fig. 6c, i). This is in contrast to the homocentric features (presence of “human”, “human face”, “human voice”, “human interaction”), whose regression weights were all modestly correlated with one another (reflecting, in part, the fact that those regressors were, themselves, correlated; see Fig. S12). These features were associated with spatially-distributed yet feature-specific patterns of connections. For instance, in the case of “human voice”, sub-components of the default mode (DMNb) strongly decouple from the control network and a sub-component of the salience/ventral attention network, but increase connection strength to other default components along with the dorsal attention network (Fig. 6f, l). The appearance of human faces, on the other hand, were accompanied by increased coupling between visual and dorsal attention networks and lacked the strong decoupling of DMNb with the salience/ventral attention network. We note that these observations are consistent irrespective of whether our linear models predict FC using each feature independently or whether features are combined into a larger model (see Fig. S11 for comparison of regression coefficients from both models and Fig. S12 for details of how each model deals with correlated regressors).

Not surprisingly, we found that the total variance of time-varying FC explained by the fit model was small ( $R^2 \approx 1.6\%$ ). This is likely a consequence of the relatively small number of predictors and the fact that they were, themselves, correlated. In future studies, two steps could be taken to improve these outcomes. We can derive more comprehensive annotations of movie features for input into models. These features could be manually coded (as in [65]), extracted automatically using powerful machine learning tools (e.g. with <https://www.clarifai.com>), or by designing synthetic virtual environments with preprogrammed sets of features accessible to the experimenter [66]. Secondly, we can improve the models by using multivariate methods like partial least squares and canonical correlation analysis to uncover patterns of edges and movie features that maximally covary with one another.

These observations, along with supplementary analyses that related time-varying changes in participation coefficient with movie features (Fig. S13), elaborate on results from the previous section and suggest that not only do groups of connections follow similar trajectories across time, but system-specific sets of connections fluctuate in tandem with the appearance and disappearance of specific features in the movie. This observation, then, serves as a link between time-varying changes in network architecture to features present in the movie.

## DISCUSSION

In this report we aimed to investigate the functional network organization during movie-watching. We found that over long timescales (static FC) and compared to rest, naturalistic stimuli induced increasingly integrated networks and systematic changes in hub structure. Over short timescales, we focused on periods when time-varying network structure appeared highly similar across subjects, which we argue corresponds to periods when subjects may be engaged in similar cognitive processes. We developed a statistical framework for identifying such periods, and applied it to movie-watching and resting state data, where it correctly identified periods of high inter-subject similarity during movie-watching but none during rest. In the case of movie-watching, we found that periods of high inter-subject similarity correspond to highly integrative brain states, in which low inter-subject similarity corresponds to network architecture that was increasingly rest-like and more segregated. We also showed that during movie-watching individual functional connections follow similar trajectories and can be clustered accordingly. Finally, we show that time-varying FC coincides with the presence and absence of particular features in the movie, suggesting possible drivers of time-varying FC.

### Naturalistic stimuli as a filter for detecting network events in time-varying FC

Many studies have investigated the time-varying architecture of functional brain networks [19], characterizing the persistence and variability of individual connections [67], uncovering flexible brain regions that change modular assignments across time [59, 68], clustering networks into “states” based on their recurrence structure over one or many scan sessions [60, 69], or relating changes in functional network structure back to anatomical connectivity [25, 70]. Implicitly, these and similar studies have regarded observed time-varying fluctuations in FC as meaningful in some way, reflecting either changes in cognitive state, underlying neurophysiological processes [71], or some combination of the two. However, it remains unclear whether these assumptions are, in fact, warranted. Recent studies have shown that sampling from time-invariant correlation structure produces variability consistent with patterns of observed time-varying FC [31, 32], while the observation that FC is stable both during sleep [72] and under anesthesia raises the question to what extent an individual’s cognitive state is manifest in FC [73], leading to debate over the verisimilitude of observed time-varying FC [33].

Here, we develop a novel statistical framework to help us better understand which fluctuations are more or less likely to be spurious. To this end, we use a measure of inter-subject similarity as a filter for deciding whether time-varying changes in FC are more or less synchronous across subjects than we would expect. This measure, though it may be conservative and likely fails to identify meaningful but individually variable and idiosyncratic fluctuations in time-varying FC, nonetheless presents a statistical argument that there exist co-fluctuations in time-varying FC not easily explainable were time-varying FC simply reflecting sampling variability from a stationary correlation structure.

These findings have implications for our interpretation of time-varying FC not just during passive movie-watching [35], but also during task-free conditions. Most importantly, we show that the fMRI BOLD response is capable of resolving meaningful fluctuations in time-

varying FC (albeit only in the statistical sense). This is an important observation, as previous studies have suggested that, combined with the windowing procedure, the slow hemodynamic response is not sufficiently temporally resolved to detect moment-to-moment fluctuations in FC [32]. This suggests that, at least in principle, similar fluctuations should be detectable at rest, which would provide critical validation of studies claiming that some fraction of observed fluctuations in FC at rest are cognitively relevant [26, 29, 30].

We note that, in general, there may be multiple strategies for “filtering” time-varying FC to identify synchronous network responses across subjects. Here, we developed a statistical framework based on time-varying FC. A complementary approach that has been investigated in previous studies [46–48] involves estimating inter-subject FC and tracking its changes across time. Although these approaches identify inter-subject synchrony at similar points in time (Fig. S3), they differ along several dimensions. For instance, our approach allows us to easily recover single-subject FC matrices at different points in time, e.g. when our model indicates high or low levels of synchrony, to compare the properties of those networks. On the other hand, assessing single-subject contributions to ISFC is challenging, as every ISFC matrix is a joint assessment of connection-level synchrony between *subject pairs*. Similar challenges exist in terms of interpreting network attributes of ISFC. For instance, two subject’s brain networks can both appear highly modular, while the ISFC matrix computed for the same pair of subjects may exhibit low levels of modularity. Future work should investigate the relative advantages of these techniques.

### **Integrated and segregated brain states during movie-watching**

One of the emerging themes in network neuroscience is that the human brain negotiates a careful balance between segregated and integrated states of information processing [74]. This balance is believed to be a critical ingredient for complex, adaptive behavior – a fully integrated system may not be able to support specialized function, whereas a system composed of completely autonomous sub-units may lack the flexibility to perform complex behaviors [75].

This balance gets reflected in the organization of brain networks [7, 16, 17, 76], which can be decomposed into modules that reflect known functional systems. Connections within and between these systems get refined, strengthened, and weakened to support ongoing task demands [15]. These refinements to connections’ weights are far from random, and follow a simple rule that influences the extent to which sub-systems are more or less segregated from one another: when an individual is tasked with performing a complex cognitive task, within-system connections weaken while between-system connections become stronger. That is, with increased cognitive load, systems become more integrated and less segregated [77].

Here, we used the modularity metric to assay the extent to which systems are segregated from one another [50], and observed that modularity was closely related to the level of inter-subject similarity. Specifically, when subjects networks were more similar to one another, their networks were less modular (less segregated) compared to periods of time when subjects’ networks appeared dissimilar. This observation is interesting for several reasons. First, it partially corroborates the implicit assumption that network inter-subject similarity indexes periods of shared or similar cognitive processing. We make this claim based on past

studies that reported similar decreases in modularity while subjects were explicitly instructed to engage in cognitively complex tasks [15, 16].

Second, this observation forces us to reinterpret other studies that have documented time-varying changes in modularity at rest [22, 23, 67]. In those studies, it was speculated that changes in modular structure across time might be driven by corresponding changes in cognitive state. Validating this hypothesis, however, was impossible due to the unconstrained nature of resting-state. Here, we observe similar patterns during movie-watching, which we *can* partially validate by showing that they occur in synchrony across individuals, suggesting that those fluctuations have similar etiological origins.

In summary, we show that brain network segregation, as measured by the modularity metric, tracks the similarity of network structure between different individuals, a property it shares with task-evoked FC. This observation suggests that periods of high inter-subject similarity may correspond to periods when subjects are attending or responding to stimuli in the movie in similar ways. Nonetheless, we also observed reduced segregation during periods of low inter-subject similarity, whose origins remain unclear. Future experimental work should investigate idiosyncratic components to time-varying FC and concurrent mental operations.

### Temporal segregation through edge clustering

While segregation and integration are usually treated as properties of a network's nodes, we also studied segregation from the perspective of edges in the network. Specifically, we focused on temporal segregation. That is, we partitioned connections into clusters according to the similarity of their trajectories over time, an approach adopted by several similar studies [55, 56]. Because each node in the network maintains many connections and because clusters were defined at the level of these connections, it was possible for a node to be associated with multiple clusters.

Some of these clusters were largely unimodal and resembled the known system-level architecture of brain networks [8, 10, 11] (see, for example, the visual cluster depicted in Fig. 5h and the somatomotor clusters shown in Fig. S8). On the other hand, we found clusters comprised of multiple brain systems, sometimes dominated by association cortex (for example, the clusters shown in Fig. 5e,g), but other times involving mixtures of association and sensory cortices (the clusters in Fig. 5d,f).

Collectively, these observations suggest that brain network segregation is not simply a static, connectional property, but one that is also encoded across time. This observation has been made before, with many studies findings that the level of segregation in brain networks changes over from moment to moment [22, 23, 67]. These studies, however, have several key limitations. Notably, they have focused on segregation during rest, making it difficult to understand what factors might be responsible for driving fluctuations in segregation. Though these studies have modeled network organization across time, the network structure at each instant was encoded through node-node interactions, implying that the community structure was non-overlapping. Here, we focus on time-varying fluctuations during movie-watching, where changes in network structure are presumably driven by stimuli in the movie, and we model encode network structure through edge-edge interactions. This approach relaxes the

definition of what it means for brain regions (or connections) to form modules, allowing brain regions to simultaneously participate in multiple clusters. This overlapping cluster organization agrees with our intuitions that brain regions can play multiple functional roles depending upon context, rather than the sometimes brittle definition of clusters as non-overlapping, which can (falsely) reinforce the notion that there is a one-to-one mapping of regions to function.

Neuroscience aside, overlapping community structure should be investigated further. Recent mathematical results have indicated that the mapping of clusters to edge weights is many to one; the implication is that unless we know the true process by which our network was generated (and its relationship with its cluster structure), we cannot unambiguously claim that the detected clusters are “correct” [78]. This fact motivates the exploration of other approaches for grouping a network into clusters [18, 79]. Future work should focus on detailed comparisons of different clustering methods and their implications for understanding brain function.

### Future directions

Here, we developed a time-varying FC framework for studying movie-watching and other naturalistic stimuli. The primary contribution of this framework is that it enables us to partition time into periods when subjects exhibit similar or dissimilar network architectures. Beyond the analyses presented here, this framework opens up several avenues for future research.

One particularly simple extension involves shifting away from periods of time when subjects are more similar than expected to periods of time when they are more *dissimilar* than expected. This extension could be useful for identifying idiosyncratic and individualized responses to stimuli. In other words, rather than focusing on fluctuations in FC that are shared across individuals, focus on periods of time when those fluctuations diverge, which could prove useful for fingerprinting [80].

We used a statistic – inter-subject similarity – that measured, on average, how similar all pairs of subjects were to one another according to their whole-brain patterns of FC. This statistic could be made more useful by, instead of averaging over all subjects, clustering the inter-subject similarity matrix, revealing groups of subjects that may be more internally similar than other groups [81]. This approach could be used for differentiating behavioral phenotypes or revealing sub-structure within a broader disorder [82, 83].

### Limitations

This study also has a number of important limitations. One such limitation is the use of a group parcellation to define nodes in the network. Recent studies have shown that subtle misalignments of subjects to such parcellations can induce biases in FC [84]. Other studies have shown that, with enough data, it is possible to generate more accurate subject-specific parcels that improve parcel homogeneity [85, 86]. Future work should investigate this possibility.

Another limitation concerns the inter-subject similarity measure, itself. In computing this measure, we consider the similarity across *all* possible connections. That is, we calculated similarity based on the whole brain. Future work should be directed to investigate system- or area-specific similarity [46], which could resolve in greater anatomical detail the drivers of inter-subject synchronization.

Yet another limitation concerns the choice of preprocessing pipeline. In particular, we opted to regress out the global gray matter signal from regional time series. This procedure has proven effective in reducing the contribution of in-scanner head motion and physiological noise on FC [87, 88] and for centering connection weights [89]. However, the global signal has also been linked to neurophysiological processes [90] and also serves as an index of arousal [71]. This suggests that, to whatever extent the global signal contributes to synchronous fluctuations in time-varying FC across subjects, our analysis may fail to characterize those fluctuations. We note, however, that the precise contribution of the global signal to FC (and the verisimilitude of those contributions) remains an active and highly debated area of research [91, 92] with no clear resolution [93–95]. Future work on this topic, including comparisons of the fMRI BOLD signal with other imaging modalities [96], will help clarify this debate and lead to more refined processing pipelines for estimating both time-invariant and time-varying FC.

A final and important limitation concerns the nature of FC itself. In this study, we reasoned that if we were to observe temporally correlated fluctuations in FC across subjects, then those fluctuations are more likely to have been driven by movie features and stimuli. Following this reasoning, we performed a series of statistical analyses to demonstrate that FC fluctuations across subjects were, indeed, correlated. It is important to note, however, that FC estimated from observations may not reflect true functional coupling (even if it is correlated across subjects). It remains possible that, due to sampling variability, slow and heterogeneous hemodynamic responses, or other confounding factors, the patterns of FC we estimated using the windowing procedure do not reflect true interregional coupling [97]. Uncovering the drivers of time-varying FC (especially at rest; [98]) has proven challenging, and at present it remains unclear to what extent observed fluctuations in FC are driven by confounding factors or reflect underlying changes in mental processes and cognitive state [31, 32]. This area of research remains highly active and future work promises to foster innovative solutions [33].

## MATERIALS AND METHODS

### Demographics

We analyzed MRI data collected from  $N_s = 29$  subjects (5 female, 24 male; 25 were right-handed). This cohort was male-dominant, as subjects were intended to serve as controls for a study in autism spectrum disorder, which is more common in men than women. At the time of their first scan, the average subject age was  $24.9 \pm 4.7$  years.



## MRI acquisition and processing

MRI images were acquired using a 3T whole-body MRI system (Magnetom Tim Trio, Siemens Medical Solutions, Natick, MA) with a 32-channel head receive array. Both raw and prescan-normalized images were acquired; raw images were used at all preprocessing stages and in all analyses unless specifically noted. During functional scans, T2\*-weighted multiband echo planar imaging (EPI) data were acquired using the following parameters: TR/TE = 813/28 ms; 1200 vol; flip angle = 60°; 3.4 mm isotropic voxels; 42 slices acquired with interleaved order covering the whole brain; multi-band acceleration factor of 3. Preceding the first functional scan, gradient-echo EPI images were acquired in opposite phase-encoding directions (10 images each with P-A and A-P phase encoding) with identical geometry to the EPI data (TR/TE = 1175/39.2 ms, flip angle = 60°) to be used to generate a fieldmap to correct EPI distortions, similar to the approach used by the Human Connectome Project [99]. High-resolution T1-weighted images of the whole brain (MPRAGE, 0.7 mm isotropic voxel size; TR/TE/TI = 2499/2.3/1000 ms) were acquired as anatomical references.

All functional data were processed according to an in-house pipeline using FEAT (v6.00) and MELODIC (v3.14) within FSL (v. 5.0.9; FMRIB's Software Library, [www.fmrib.ox.ac.uk/fsl](http://www.fmrib.ox.ac.uk/fsl)), Advanced Normalization Tools (ANTs; v2.1.0) [100], and Matlab R2014b. This pipeline was identical to the **GLM + MGTR** procedure described in [88].

In more detail, individual anatomical images were bias-corrected and skull-stripped using ANTs, and segmented into gray matter, white matter, and CSF partial volume estimates using FSL FAST. A midspace template was constructed using ANTs' *buildtemplateparallel* and subsequently skull-stripped. Composite (affine and diffeomorphic) transforms warping each individual anatomical image to this midspace template, and warping the midspace template to the Montreal Neurological Institute MNI152 1mm reference template, were obtained using ANTs.

For each functional run, the first five volumes ( $\approx 4$  seconds) were discarded to minimize magnetization equilibration effects. Framewise displacement traces for this raw (trimmed) data were computed using *fsl\_motion\_outliers*. Following [88, 101], we performed FIX followed by mean cortical signal regression. This procedure included rigid-body motion correction, fieldmap-based geometric distortion correction, and non-brain removal (but not slice-timing correction due to fast TR [99]). Preprocessing included weak highpass temporal filtering ( $>2000$  s FWHM) to remove slow drifts [99] and no spatial smoothing. Off-resonance geometric distortions in EPI data were corrected using a fieldmap derived from two gradient-echo EPI images collected in opposite phase-encoding directions (posterior-anterior and anterior-posterior) using FSL topup.

We then used FSL-FIX [102] to regress out independent components classified as noise using a classifier trained on independent but similar data and validated on hand-classified functional runs. The residuals were regarded as "cleaned" data. Finally, we regressed out the mean cortical signal (mean BOLD signal across gray matter partial volume estimate obtained from FSL FAST). All analyses were carried out on these data, which were registered to subjects' skull-stripped T1-weighted anatomical imaging using Boundary-Based Registration (BBR) with *epi reg* within FSL. Subjects' functional images were then

transformed to the MNI152 reference in a single step, using ANTS to apply a concatenation of the affine transformation matrix with the composite (affine + diffeomorphic) transforms between a subject's anatomical image, the midspace template, and the MNI152 reference. Prior to network analysis, we extracted mean regional time series from regions of interest defined as sub-divisions of the 17-system parcellation reported in [11] and used previously [83, 103, 104]. Wakefulness during movie and rest scans was monitored in real-time using an eye tracking camera (Eyelink 1000).

### Naturalistic stimuli

All movies were obtained from Vimeo (<https://vimeo.com>). They were selected based on multiple criteria. First, to ensure that movies represented novel stimuli, we excluded any movie that had a wide theatrical release. Secondly, we excluded movies with potentially objectionable content including nudity, swearing, drug use, etc. Lastly, we excluded movies with intentionally startling events that could lead to excessive in-scanner movement.

Each movie lasted between 45 and 285 seconds (approximately 1 to 5 minutes). Each movie scan comprised between four and six movies with genres that included documentaries, dramas, comedies, sports, mystery, and adventure. See Table. S1 for more details.

### Connectivity measures

**Time-averaged FC**—Let  $\mathbf{x}_i = [x_i(1), \dots, x_i(T)]$  be the vector of activity for region  $i$ . We calculated time-averaged FC as the Pearson correlation of activity recorded in region  $i$  and  $j$ :

$$r_{ij} = \frac{1}{T-1} \sum_{t=1}^T \frac{(x_i(t) - \mu_{x_i})(x_j(t) - \mu_{x_j})}{\sigma_{x_i}\sigma_{x_j}} \quad (1)$$

Where  $\mu_{x_i}$  and  $\sigma_{x_i}$  are the time-averaged mean and standard deviation of activity in region  $i$ . All correlation coefficients were subsequently Fisher transformed.

**Time-varying FC**—In addition to time-averaged FC, we also calculated time-varying FC using a sliding window approach. This involves defining a window length of  $L$  (in units of TRs) and computing FC using samples within that window only. The time-varying FC at time  $t$  was calculated as:

$$r_{ij}(t) = \frac{1}{L-1} \sum_{\tau=t}^{t+L-1} \frac{(x_i(\tau) - \mu_{x_i}(t))(x_j(\tau) - \mu_{x_j}(t))}{\sigma_{x_i}(t)\sigma_{x_j}(t)} \quad (2)$$

where:

$$\mu_{x_i}(t) = \frac{1}{L} \sum_{\tau=t}^{t+L-1} x_i(\tau) \quad (3)$$

and

$$\sigma_{x_i}(t) = \frac{1}{L-1} \sum_{\tau=t}^{t+L-1} (x_i(\tau) - \mu_{x_i}(t)) \quad (4)$$

are the mean and standard deviation of region  $i$ 's activity recording in the  $L$ -length window starting at time  $t$ .

An important decision to make in computing time-varying FC using sliding windows is the choice of window length,  $L$ . Here, we set  $L = 10$  TRs (for a window duration 8.13 s). We note that this length is considerably shorter than the durations suggested elsewhere, which have argued that the minimum window duration should be inversely proportional to the slowest frequency component of the fMRI signal. For instance, if filtered to 0.01 – 0.1 Hz, then the shortest window should be  $\frac{1}{0.01} = 100$  seconds [105, 106]. Our rationale for disregarding this rule of thumb is that the inter-subject similarity measure (see the next section) effectively acts as a filter to reduce the rate of false positives. That is, if fluctuations in time-varying FC within a subject are spurious, it is unlikely that similarly spurious fluctuations in another subject occur at precisely the same instants in time.

### Inter-subject similarity

Let  $r^u(t)$  and  $r^v(t)$  be the network structure estimated for subjects  $u$  and  $v$  at time  $t$ . We calculate the similarity of these two networks by vectorizing their upper triangle elements and computing the correlation of those vectors:  $r_{u,v}(t)$ . In practice, we compute this measurement for all pairs of subjects and for every window, generating a distribution of inter-subject similarity scores that evolves over time.

### Statistical framework for evaluating inter-subject similarity

In general, our aim was to develop a procedure for testing whether the inter-subject similarity (ISS) of time-varying FC was stronger during movie-watching than chance. This required that we complete two interrelated tasks. First, we needed to develop a test statistic (an estimate of ISS at each point in time,  $t$ ). Then, we needed a null model to generate a null distribution against which we could compare our empirical test statistic and estimate a p-value.

Addressing the first point was simple. For each movie-watching scan, we estimated subjects' time-varying FC and computed the mean inter-subject similarity at every instant. That is, we calculated:

$$\mu_r(t) = \frac{2}{N_{sub}(N_{sub} - 1)} \sum_{u,v > u} r_{u,v}(t) \quad (5)$$

This procedure generated a time-series that indexed how similar subjects' FC patterns were at each point in time.

Addressing the second point - null model construction - required more thought. We realized, however, that the resting-scans could serve as a convenient null model. Because subjects'

recorded activity is spontaneous and temporally asynchronous with respect to one another, any measured similarity between subjects at rest must be driven by similarity in their intrinsic (task-free) architectures or by noise.

So, we concatenated the vectors of mean inter-subject similarity for all four resting scans and pooled their values to generate a null distribution. For all four movie-watching scans, we proceeded frame by frame and calculated the percent of the null distribution that was greater than the inter-subject similarity at that frame in the movie. We also repeated this procedure for resting scans to estimate framewise p-values. To avoid circularity, the null distribution we used to generate p-values for scan Rest1 consisted only of data from scans Rest2, Rest3, and Rest4.

### Modularity maximization

Modularity maximization is a heuristic for detecting communities in networks [50]. Intuitively, it attempts to decompose a network into non-overlapping sub-networks such that the observed density of connections within sub-networks maximally exceeds what would be expected by chance, where chance is determined by the user. The actual process of detecting communities is accomplished by choosing community assignments that maximize a modularity quality function,  $Q$ , defined as:

$$Q = \sum_{ij} B_{ij} \delta(g_i, g_j) \quad (6)$$

where  $B_{ij} = A_{ij} - P_{ij}$  is the  $\{i, j\}$  element of the modularity matrix, which represents the observed weight of the connection between nodes  $i$  and  $j$  minus the expected weight. The variable  $g_i$  is the community assignment of node  $i$  and  $\delta(x, y)$  is the Kronecker delta function, whose value is 1 when  $g_i = g_j$  and 0 otherwise. The modularity,  $Q$ , is effectively a sum over all edges that fall within communities and is optimized when the the observed weights of connections is maximally greater than the expected. In general, larger values of  $Q$  are thought to reflect superior community partitions.

**Signed and correlation matrices**—In this manuscript, we use two variations of modularity. First, we use the modularity quality function as a means of assessing the level of segregation in a network. Specifically, we estimate a variant of modularity,  $Q^*$ , which has been shown to be especially well-suited for use with correlation matrices [54]:

$$Q^* = Q^+ + \frac{v^-}{v^+v^-} Q^- \quad (7)$$

where  $Q^\pm = \frac{1}{v^\pm} \sum_{ij} (r_{ij}^\pm - \frac{k_i^\pm k_j^\pm}{v^\pm}) \delta(g_i, g_j)$ . In this expression,  $r_{ij}^\pm$  represents either the positive or negative elements of the correlation matrix,  $k_i^\pm = \sum_j r_{ij}^\pm$ , and  $v^\pm = \sum_i k_i^\pm$ . In all instances where we report  $Q^*$ , it is with respect to the seventeen systems reported in [49].

**Edge communities**—We also apply a second variant of modularity maximization to edge correlation matrices. As a product of calculating time-varying FC, we obtain a time series

for each connection in the network, i.e.  $\mathbf{r}_{ij} = [r_{ij}(t=1), \dots, r_{ij}(t=T-L+1)]$ . From these edge-level time series, we can calculate the edge-by-edge correlation matrix,  $\Omega$ , whose element  $\Omega\{i, j\}$ ,  $\{k, l\}$  is equal to the correlation of  $\mathbf{r}_{ij}$  and  $\mathbf{r}_{kl}$ .

Here, we generate a modularity matrix wherein we treat the edge correlations from the movie as our observed network and the edge correlations from rest as our expected or chance network. This matrix is calculated as:  $B(\gamma) = [\Omega^{movie} - \Omega^{rest}] - \gamma$ , where  $\gamma$  is a structural resolution parameter that can be tuned to detect different numbers and sizes of communities [107]. We can then optimize the corresponding modularity to identify edges whose trajectories across time are more correlated during movie-watching than during rest. Because this procedure assigns edges to communities rather than nodes, it is possible for a fraction of a nodes' edges to be associated with multiple distinct communities, so that each node has an overlapping community structure. We note that this approach is similar to the hypergraph clustering reported in [55, 56].

An important open question concerns choosing the optimal value of  $\gamma$ . Here, we tested 51 different  $\gamma$  values, linearly-spaced between from  $-0.005 - 0.36$ . We focused on the value that maximized the minimum average within-community correlation minus the maximum average between-community correlation.

We used a generalization of the popular Louvain algorithm to optimize  $Q$  [108, 109]. This algorithm is nondeterministic, meaning that different initializations lead to slightly different estimates of community structure (we run the algorithm 100 times). To resolve variability in these estimates we use consensus clustering [110]. The variant of consensus clustering used here is slightly different from that of [110]. Specifically, we calculate a co-assignment matrix from the 100 estimated partitions, whose elements indicate the fraction of times that pairs of nodes were assigned to the same community. In [110], this matrix is iteratively thresholded and clustered until convergence. Here, we estimate the probability that two nodes would be assigned to the same community were community labels randomly permuted. Then, we construct a new modularity matrix by subtracting this probability from the co-assignment matrix, which we then cluster using modularity maximization [67, 111, 112].

**Community overlap entropy**—The edge community detection procedure assigned every pair of nodes,  $\{i, j\}$  to a community. To better understand how these edge-level labels were related to individual brain areas (nodes), we calculated each node's community overlap entropy [113]. Let  $\Gamma_i = \{g_{i1}, \dots, g_{iN}\}$  be the set of community assignments for all edges involves node  $i$ . Each element,  $g_{ij} \in \{1, \dots, K\}$  indicates to which of the  $K$  communities edge  $\{i, j\}$  was assigned. We then calculate the fraction of all edges assigned to each community,  $c$ , as  $Pr_c$ , and subsequently calculate the entropy over this distribution as:

$H = - \sum_{c=1}^K Pr_c \log(Pr_c)$ . We normalize this entropy to the interval  $[0, 1]$  as

$H_{norm} = \frac{H}{\log(K)}$ . We repeat this procedure separately for every node  $i \in \{1, \dots, N\}$ , resulting in node-defined entropies,  $H_i$ .

**Participation coefficient**—Knowing a network’s community structure allows us to classify nodes based on their functional roles. One popular measure for doing so is the so-called *participation coefficient* [51], which measures the extent to which a node’s connections are distributed within or across modules. The participation coefficient is calculated as:

$$P_i = 1 - \sum_{m=1}^M \left( \frac{\kappa_{im}}{k_i} \right)^2 \quad (8)$$

where  $k_i$  and  $\kappa_{im}$  are the total number (or weight) of connections made by node  $i$  overall and to module  $m$ .

### Multi-linear models

To map the association of time-varying FC with features present in the movie, we build edge-level multi-linear regression models. Let  $\mathbf{r}_{ij} = [r_{ij}(1), \dots, r_{ij}(T)]$  denote the time series of connection weights between nodes  $i$  and  $j$ , and let  $\mathbf{f}_l = [f_l(1), \dots, f_l(T)]$  be the z-scored time series of feature  $l$ . All features were coded as binary variables; the exception was luminance, which was coded continuously.

To model the response of connection  $\{i, j\}$  to feature,  $l$ , we constructed the simple linear model:

$$\hat{\mathbf{r}}_{ij} = \beta_0 + \beta_l \mathbf{f}_l + \epsilon. \quad (9)$$

In this equation,  $\hat{\mathbf{r}}_{ij}$  is the predicted connection time series,  $\beta_0$  is the intercept term, and  $\beta_l$  is the regression coefficient for feature  $l$ .

We extended this simple linear regression to include multiple terms – one for each of the six coded features:

$$\hat{\mathbf{r}}_{ij} = \beta_0 + \sum_{l=1}^6 \beta_l \mathbf{f}_l + \epsilon. \quad (10)$$

We fit the model to time series data from every connection by minimizing the least square error. The  $p$ -value associated with each  $\beta$  regression coefficient was calculated using a one-sample  $t$ -test.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### ACKNOWLEDGMENTS

RFB and FZE acknowledge support from Indiana University Office of the Vice President for Research Emerging Area of Research Initiative, Learning: Brains, Machines and Children. This work was supported by the NIH (R01MH110630 and R00MH094409 to DPK and T32HD007475 Postdoctoral Traineeship to LB). For supercomputing resources, this work was supported in part by Lilly Endowment, Inc., through its support for the



Indiana University Pervasive Technology Institute, and in part by the Indiana METACyt Initiative. The Indiana METACyt Initiative at IU was also supported in part by Lilly Endowment, Inc. We thank Hu Cheng for MRI protocol development, Soyoung Park for training the FIX classifier, and Brad Caron and Susannah Burkholder for data collection.

## References

- [1]. McIntosh AR, *Neural Networks* 13, 861 (2000). [PubMed: 11156197]
- [2]. Bressler SL and Menon V, *Trends in cognitive sciences* 14, 277 (2010). [PubMed: 20493761]
- [3]. Park H-J and Friston K, *Science* 342, 1238411 (2013). [PubMed: 24179229]
- [4]. Barabási A-L et al., *Network science* (Cambridge university press, 2016).
- [5]. Bullmore E and Sporns O, *Nature Reviews Neuroscience* 10, 186 (2009). [PubMed: 19190637]
- [6]. Wang J, Zuo X, and He Y, *Frontiers in systems neuroscience* 4, 16 (2010). [PubMed: 20589099]
- [7]. Sporns O and Betzel RF, *Annual review of psychology* 67, 613 (2016).
- [8]. Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, et al., *Neuron* 72, 665 (2011). [PubMed: 22099467]
- [9]. Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, et al., *Proceedings of the National Academy of Sciences* 106, 13040 (2009).
- [10]. Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, and Petersen SE, *Cerebral cortex* 26, 288 (2014). [PubMed: 25316338]
- [11]. Thomas Yeo B, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, et al., *Journal of neurophysiology* 106, 1125 (2011). [PubMed: 21653723]
- [12]. Shen X, Tokoglu F, Papademetris X, and Constable RT, *Neuroimage* 82, 403 (2013). [PubMed: 23747961]
- [13]. Clune J, Mouret J-B, and Lipson H, *Proceedings of the Royal Society b: Biological sciences* 280, 20122863 (2013).
- [14]. Wig GS, *Trends in Cognitive Sciences* 21, 981 (2017). [PubMed: 29100737]
- [15]. Cole MW, Bassett DS, Power JD, Braver TS, and Petersen SE, *Neuron* 83, 238 (2014). [PubMed: 24991964]
- [16]. Cohen JR and D'Esposito M, *Journal of Neuroscience* 36, 12083 (2016). [PubMed: 27903719]
- [17]. Shine JM, Bissett PG, Bell PT, Koyejo O, Balsters JH, Gorgolewski KJ, Moodie CA, and Poldrack RA, *Neuron* 92, 544 (2016). [PubMed: 27693256]
- [18]. Betzel RF, Bertolero MA, and Bassett DS, *bioRxiv*, 355016 (2018).
- [19]. Hutchison RM, Womelsdorf T, Allen EA, Bandettini PA, Calhoun VD, Corbetta M, Della Penna S, Duyn JH, Glover GH, Gonzalez-Castillo J, et al., *Neuroimage* 80, 360 (2013). [PubMed: 23707587]
- [20]. Handwerker DA, Roopchansingh V, Gonzalez-Castillo J, and Bandettini PA, *Neuroimage* 63, 1712 (2012). [PubMed: 22796990]
- [21]. Jones DT, Vemuri P, Murphy MC, Gunter JL, Senjem ML, Machulda MM, Przybelski SA, Gregg BE, Kantarci K, Knopman DS, et al., *PloS one* 7, e39731 (2012). [PubMed: 22761880]
- [22]. Liégeois R, Ziegler E, Phillips C, Geurts P, Gómez F, Bahri MA, Yeo BT, Soddu A, Vanhaudenhuyse A, Laureys S, et al., *Brain structure and function* 221, 2985 (2016). [PubMed: 26197763]
- [23]. Fukushima M, Betzel RF, He Y, de Reus MA, van den Heuvel MP, Zuo X-N, and Sporns O, *NeuroImage* 180, 406 (2018). [PubMed: 28823827]
- [24]. Zalesky A, Fornito A, Cocchi L, Gollo LL, and Breakspear M, *Proceedings of the National Academy of Sciences* 111, 10341 (2014).
- [25]. Fukushima M, Betzel RF, He Y, van den Heuvel MP, Zuo X-N, and Sporns O, *Brain Structure and Function* 223, 1091 (2018). [PubMed: 29090337]
- [26]. Kucyi A, Hove MJ, Esterman M, Hutchison RM, and Valera EM, *Cerebral cortex* 27, 1831 (2017). [PubMed: 26874182]

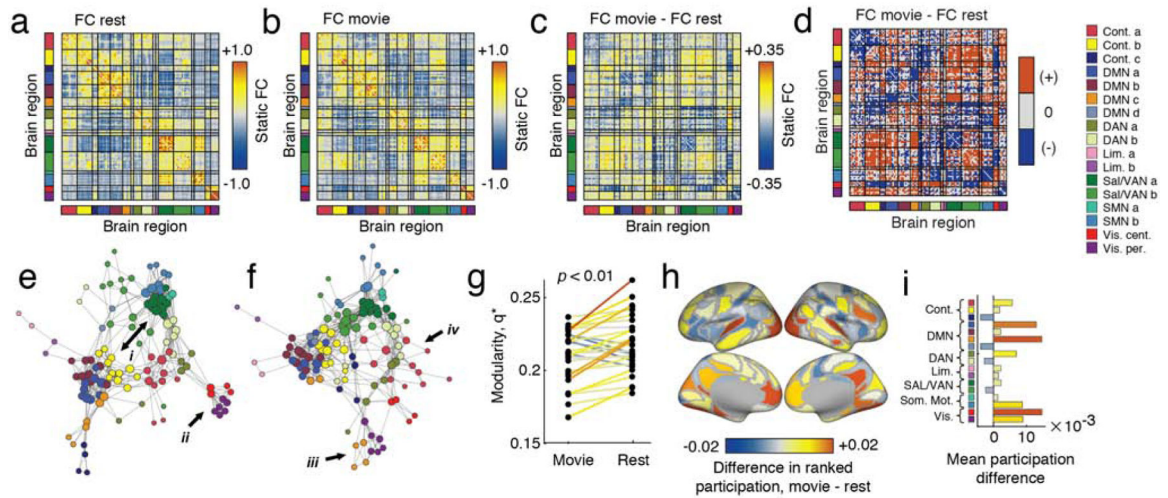
- [27]. Haimovici A, Tagliazucchi E, Balenzuela P, and Laufs H, *Scientific reports* 7, 5908 (2017). [PubMed: 28724928]
- [28]. Kucyi A and Davis KD, *Neuroimage* 100, 471 (2014). [PubMed: 24973603]
- [29]. Barttfeld P, Uhrig L, Sitt JD, Sigman M, Jarraya B, and Dehaene S, *Proceedings of the National Academy of Sciences* 112, 887 (2015).
- [30]. Sadaghiani S, Poline J-B, Kleinschmidt A, and DEsposito M, *Proceedings of the National Academy of Sciences* 112, 8463 (2015).
- [31]. Liegeois R, Laumann TO, Snyder AZ, Zhou J, and Yeo BT, *Neuroimage* 163, 437 (2017). [PubMed: 28916180]
- [32]. Laumann TO, Snyder AZ, Mitra A, Gordon EM, Gratton C, Adeyemo B, Gilmore AW, Nelson SM, Berg JJ, Greene DJ, et al., *Cerebral cortex* 27, 4719 (2016).
- [33]. Lurie D, Kessler D, Bassett D, Betzel RF, Breakspear M, Keilholz S, Kucyi A, Liégeois R, Lindquist MA, McIntosh AR, et al., (2018).
- [34]. Gonzalez-Castillo J and Bandettini PA, *Neuroimage* 180, 526 (2018). [PubMed: 28780401]
- [35]. Sonkusare S, Breakspear M, and Guo C, *Trends in Cognitive Sciences* (2019).
- [36]. Huth AG, Nishimoto S, Vu AT, and Gallant JL, *Neuron* 76, 1210 (2012). [PubMed: 23259955]
- [37]. Hasson U, Malach R, and Heeger DJ, *Trends in cognitive sciences* 14, 40 (2010). [PubMed: 20004608]
- [38]. Hasson U, Furman O, Clark D, Dudai Y, and Davachi L, *Neuron* 57, 452 (2008). [PubMed: 18255037]
- [39]. Ren Y, Nguyen VT, Sonkusare S, Lv J, Pang T, Guo L, Eickhoff SB, Breakspear M, and Guo CC, *Nature communications* 9, 4875 (2018).
- [40]. Gilson M, Deco G, Friston KJ, Hagmann P, Mantini D, Betti V, Romani GL, and Corbetta M, *Neuroimage* 180, 534 (2018). [PubMed: 29024792]
- [41]. Kim D, Kay K, Shulman GL, and Corbetta M, *Cerebral Cortex* 28, 3065 (2017).
- [42]. Wang J, Ren Y, Hu X, Nguyen VT, Guo L, Han J, and Guo CC, *Human brain mapping* 38, 2226 (2017). [PubMed: 28094464]
- [43]. Vanderwal T, Kelly C, Eilbott J, Mayes LC, and Castellanos FX, *Neuroimage* 122, 222 (2015). [PubMed: 26241683]
- [44]. Greene DJ, Koller JM, Hampton JM, Wesevich V, Van AN, Nguyen AL, Hoyt CR, McIntyre L, Earl EA, Klein RL, et al., *Neuroimage* 171, 234 (2018). [PubMed: 29337280]
- [45]. Vanderwal T, Eilbott J, Finn ES, Craddock RC, Turnbull A, and Castellanos FX, *Neuroimage* 157, 521 (2017). [PubMed: 28625875]
- [46]. Simony E, Honey CJ, Chen J, Lositsky O, Yeshurun Y, Wiesel A, and Hasson U, *Nature communications* 7, 12141 (2016).
- [47]. Bolton TA, Jochaut D, Giraud A-L, and Van De Ville D, *JoVE (Journal of Visualized Experiments)*, e59083 (2019).
- [48]. Bolton TA, Freitas L, Jochaut D, Giraud A-L, and Van De Ville D, *NeuroImage*, 116571 (2020). [PubMed: 31987996]
- [49]. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, Eickhoff SB, and Yeo BT, *Cerebral Cortex*, 1 (2017). [PubMed: 28365777]
- [50]. Newman ME and Girvan M, *Physical review E* 69, 026113 (2004).
- [51]. Guimera R and Amaral LAN, *nature* 433, 895 (2005). [PubMed: 15729348]
- [52]. van den Heuvel MP and Sporns O, *Trends in cognitive sciences* 17, 683 (2013). [PubMed: 24231140]
- [53]. Hasson U and Honey CJ, *NeuroImage* 62, 1272 (2012). [PubMed: 22348879]
- [54]. Rubinov M and Sporns O, *Neuroimage* 56, 2068 (2011). [PubMed: 21459148]
- [55]. Bassett DS, Wymbs NF, Porter MA, Mucha PJ, and Grafton ST, *Chaos: An Interdisciplinary Journal of Nonlinear Science* 24, 013112 (2014).
- [56]. Davison EN, Schlesinger KJ, Bassett DS, Lynall M-E, Miller MB, Grafton ST, and Carlson JM, *PLoS computational biology* 11, e1004029 (2015). [PubMed: 25569227]
- [57]. Traag VA, Van Dooren P, and Nesterov Y, *Physical Review E* 84, 016114 (2011).

- [58]. Bazzi M, Porter MA, Williams S, McDonald M, Fenn DJ, and Howison SD, *Multiscale Modeling & Simulation* 14, 1 (2016).
- [59]. Betzel RF, Satterthwaite TD, Gold JI, and Bassett DS, *Scientific Reports* 7, 520 (2017). [PubMed: 28364117]
- [60]. Shine JM, Koyejo O, and Poldrack RA, *Proceedings of the National Academy of Sciences* 113, 9888 (2016).
- [61]. Esfahlani FZ, Bertolero M, Bassett D, and Betzel R, *BioRxiv*, 590935 (2019).
- [62]. Fadiga L, Fogassi L, Pavesi G, and Rizzolatti G, *Journal of neurophysiology* 73, 2608 (1995). [PubMed: 7666169]
- [63]. Lotze M, Montoya P, Erb M, Hülsmann E, Flor H, Klose U, Birbaumer N, and Grodd W, *Journal of cognitive neuroscience* 11, 491 (1999). [PubMed: 10511638]
- [64]. Uddin LQ, Iacoboni M, Lange C, and Keenan JP, *Trends in cognitive sciences* 11, 153 (2007). [PubMed: 17300981]
- [65]. Huth AG, De Heer WA, Griffiths TL, The-unissen FE, and Gallant JL, *Nature* 532, 453 (2016). [PubMed: 27121839]
- [66]. Mueller C, Luehrs M, Baecke S, Adolf D, Luetzkendorf R, Luchtman M, and Bernarding J, *Journal of neuroscience methods* 209, 290 (2012). [PubMed: 22759716]
- [67]. Betzel RF, Fukushima M, He Y, Zuo X-N, and Sporns O, *NeuroImage* 127, 287 (2016). [PubMed: 26687667]
- [68]. Bassett DS, Wymbs NF, Porter MA, Mucha PJ, Carlson JM, and Grafton ST, *Proceedings of the National Academy of Sciences* 108, 7641 (2011).
- [69]. Allen EA, Damaraju E, Plis SM, Erhardt EB, Eichele T, and Calhoun VD, *Cerebral cortex* 24, 663 (2014). [PubMed: 23146964]
- [70]. Shen K, Mišić B, Cipollini BN, Bezgin G, Buschkuehl, Hutchison RM, Jaeggi SM, Kross E, Peltier SJ, Everling S, et al., *Proceedings of the National Academy of Sciences* 112, 6473 (2015).
- [71]. Chang C, Leopold DA, Schölvinc ML, Mandelkow H, Picchioni D, Liu X, Frank QY, Turchi JN, and Duyn JH, *Proceedings of the National Academy of Sciences* 113, 4518 (2016).
- [72]. Sämann PG, Wehrle R, Hoehn D, Spormaker VI, Peters H, Tully C, Holsboer F, and Czisch M, *Cerebral cortex* 21, 2082 (2011). [PubMed: 21330468]
- [73]. Mhuircheartaigh RN, Rosenorn-Lanng D, Wise R, Jbabdi S, Rogers R, and Tracey I, *Journal of Neuroscience* 30, 9095 (2010). [PubMed: 20610743]
- [74]. Tononi G, Sporns O, and Edelman GM, *Proceedings of the National Academy of Sciences* 91, 5033 (1994).
- [75]. Edelman G and Tononi G, *The lifespan development of individuals: Behavioral, neurobiological, and psychosocial perspectives: A synthesis*, 179 (1997).
- [76]. Sporns O, *Current opinion in neurobiology* 23, 162 (2013). [PubMed: 23294553]
- [77]. Gallen CL and D'Esposito M, *Trends in cognitive sciences* (2019).
- [78]. Peel L, Larremore DB, and Clauset A, *Science advances* 3, e1602548 (2017). [PubMed: 28508065]
- [79]. Faskowitz J, Yan X, Zuo X-N, and Sporns O, *Scientific reports* 8, 12997 (2018). [PubMed: 30158553]
- [80]. Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, and Constable RT, *Nature neuroscience* 18, 1664 (2015). [PubMed: 26457551]
- [81]. Rasero J, Pellicoro M, Angelini L, Cortes JM, Marinazzo D, and Stramaglia S, *Network Neuroscience* 1, 242 (2017). [PubMed: 29601048]
- [82]. Yang Z, Xu Y, Xu T, Hoy CW, Handwerker DA, Chen G, Northoff G, Zuo X-N, and Bandettini PA, *Scientific reports* 4, 5549 (2014). [PubMed: 24989351]
- [83]. Byrge L, Dubois J, Tyszka JM, Adolphs R, and Kennedy DP, *Journal of Neuroscience* 35, 5837 (2015). [PubMed: 25855192]
- [84]. Bijsterbosch JD, Woolrich MW, Glasser MF, Robinson EC, Beckmann CF, Van Essen DC, Harrison SJ, and Smith SM, *Elife* 7, e32992 (2018). [PubMed: 29451491]

- [85]. Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, Berg JJ, Ortega M, Hoyt-Drazen C, Grattton C, Sun H, et al., *Neuron* 95, 791 (2017). [PubMed: 28757305]
- [86]. Laumann TO, Gordon EM, Adeyemo B, Snyder AZ, Joo SJ, Chen M-Y, Gilmore AW, McDermott KB, Nelson SM, Dosenbach NU, et al., *Neuron* 87, 657 (2015). [PubMed: 26212711]
- [87]. Ciric R, Wolf DH, Power JD, Roalf DR, Baum GL, Ruparel K, Shinohara RT, Elliott MA, Eickhoff SB, Davatzikos C, et al., *Neuroimage* 154, 174 (2017). [PubMed: 28302591]
- [88]. Byrge L and Kennedy DP, *NeuroImage* 171, 376 (2018). [PubMed: 29288128]
- [89]. Fox MD, Zhang D, Snyder AZ, and Raichle ME, *Journal of neurophysiology* 101, 3270 (2009). [PubMed: 19339462]
- [90]. Schölvinck ML, Maier A, Frank QY, Duyn JH, and Leopold DA, *Proceedings of the National Academy of Sciences* 107, 10238 (2010).
- [91]. Power JD, Laumann TO, Plitt M, Martin A, and Petersen SE, *Trends in cognitive sciences* 21, 911 (2017). [PubMed: 28939332]
- [92]. Power JD, Plitt M, Gotts SJ, Kundu P, Voon V, Bandettini PA, and Martin A, *Proceedings of the National Academy of Sciences* 115, E2105 (2018).
- [93]. Glasser MF, Coalson TS, Bijsterbosch JD, Harrison SJ, Harms MP, Anticevic A, Van Essen DC, and Smith SM, *NeuroImage* 181, 692 (2018). [PubMed: 29753843]
- [94]. Power JD, “Temporal ica has not properly separated global fmri signals: A comment on glasser et al.(2018).” (2019).
- [95]. Glasser MF, Coalson TS, Bijsterbosch JD, Harrison SJ, Harms MP, Anticevic A, Van Essen DC, and Smith SM, *NeuroImage* 197, 435 (2019). [PubMed: 31026516]
- [96]. Lake EM, Ge X, Shen X, Herman P, Hyder F, Cardin JA, Higley MJ, Scheinost D, Papademetris X, Crair MC, et al., *bioRxiv*, 464305 (2019).
- [97]. Hindriks R, Adhikari MH, Murayama Y, Ganzetti M, Mantini D, Logothetis NK, and Deco G, *Neuroimage* 127, 242 (2016). [PubMed: 26631813]
- [98]. Betzel RF, Faskowitz J, and Sporns O, *bioRxiv*, 800045 (2019).
- [99]. Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, Duff E, Feinberg DA, Griffanti L, Harms MP, et al., *Neuroimage* 80, 144 (2013). [PubMed: 23702415]
- [100]. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, and Gee JC, *Neuroimage* 54, 2033 (2011). [PubMed: 20851191]
- [101]. Burgess GC, Kandala S, Nolan D, Laumann TO, Power JD, Adeyemo B, Harms MP, Petersen SE, and Barch DM, *Brain connectivity* 6, 669 (2016). [PubMed: 27571276]
- [102]. Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, and Smith SM, *Neuroimage* 90, 449 (2014). [PubMed: 24389422]
- [103]. Betzel RF, Byrge L, He Y, Goñi J, Zuo X-N, and Sporns O, *Neuroimage* 102, 345 (2014). [PubMed: 25109530]
- [104]. Betzel RF, Miši B, He Y, Rumschlag J, Zuo X-N, and Sporns O, *arXiv preprint arXiv:1510.08045* (2015).
- [105]. Zalesky A and Breakspear M, *Neuroimage* 114, 466 (2015). [PubMed: 25818688]
- [106]. Leonardi N and Van De Ville D, *Neuroimage* 104, 430 (2015). [PubMed: 25234118]
- [107]. Reichardt J and Bornholdt S, *Physical Review E* 74, 016110 (2006).
- [108]. Blondel VD, Guillaume J-L, Lambiotte R, and Lefebvre E, *Journal of statistical mechanics: theory and experiment* 2008, P10008 (2008).
- [109]. Jutla IS, Jeub LG, and Mucha PJ, URL <http://netwiki.amath.unc.edu/GenLouvain> (2011).
- [110]. Lancichinetti A and Fortunato S, *Scientific reports* 2, 336 (2012). [PubMed: 22468223]
- [111]. Betzel RF, Medaglia JD, Papadopoulos L, Baum GL, Gur R, Gur R, Roalf D, Satterthwaite TD, and Bassett DS, *Network Neuroscience* 1, 42 (2017). [PubMed: 30793069]
- [112]. Betzel RF and Bassett DS, *Proceedings of the National Academy of Sciences*, 201720186 (2018).
- [113]. Faskowitz J, Esfahlani FZ, Jo Y, Sporns O, and Betzel RF, *bioRxiv*, 799924 (2019).

**SIGNIFICANCE STATEMENT**

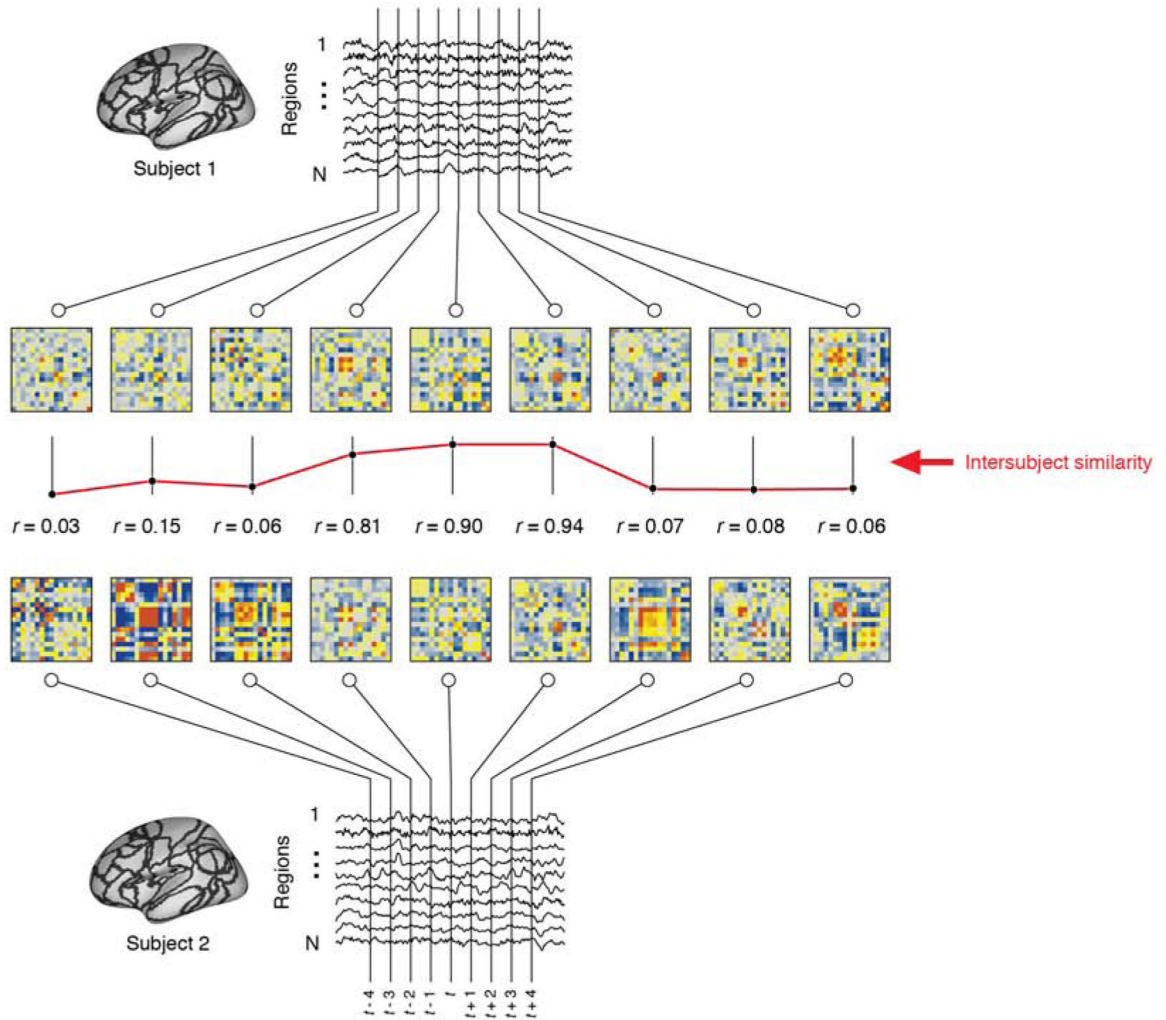
Cognition fluctuates over short timescales. To understand the network-level correlates, many studies have embraced “dynamic” network analysis. This approach is prone to false positives and sampling variability can artifactually create the illusion of change. Moreover, several groups have argued that fMRI is incapable of detecting true changes in network structure over time, even if they exist. Here, we address these challenges directly using naturalistic stimuli. We create dynamic networks during rest and movie-watching and show that subjects’ networks periodically become synchronized during movie-watching but never at rest, confirming that dynamic network analysis is capable of detecting cognitively meaningful changes in brain network state. We further explore synchronization effects using network modularity analyses.



**FIG. 1. Time-averaged FC at rest versus movie-watching.**

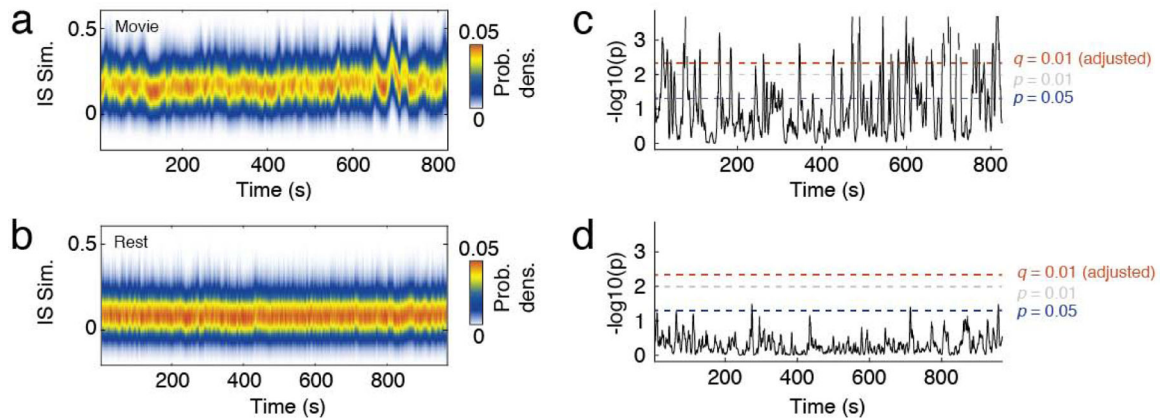
(a) Time-averaged FC matrix at rest, ordered by brain systems. (b) Time-averaged FC matrix during movie-watching, ordered by brain systems (c) Unthresholded mean differences in FC between time-averaged movie-watching and rest. (d) Result of statistical analysis comparing movie-watching and rest. Red cells indicate connections that were significantly stronger during movie-watching than rest, whereas blue indicates the opposite. Grey cells were not statistically different between conditions. Panels e and f depict thresholded versions of the matrices from a and b, with node locations determined by a force-directed embedding algorithm. We draw attention to particularly salient differences using Roman numerals. (g) Differences in between movie-watching and rest. (h) Node-level differences in ranked participation coefficient. (i) Same information as in panel h, but with differences in participation coefficient averaged by cognitive system.





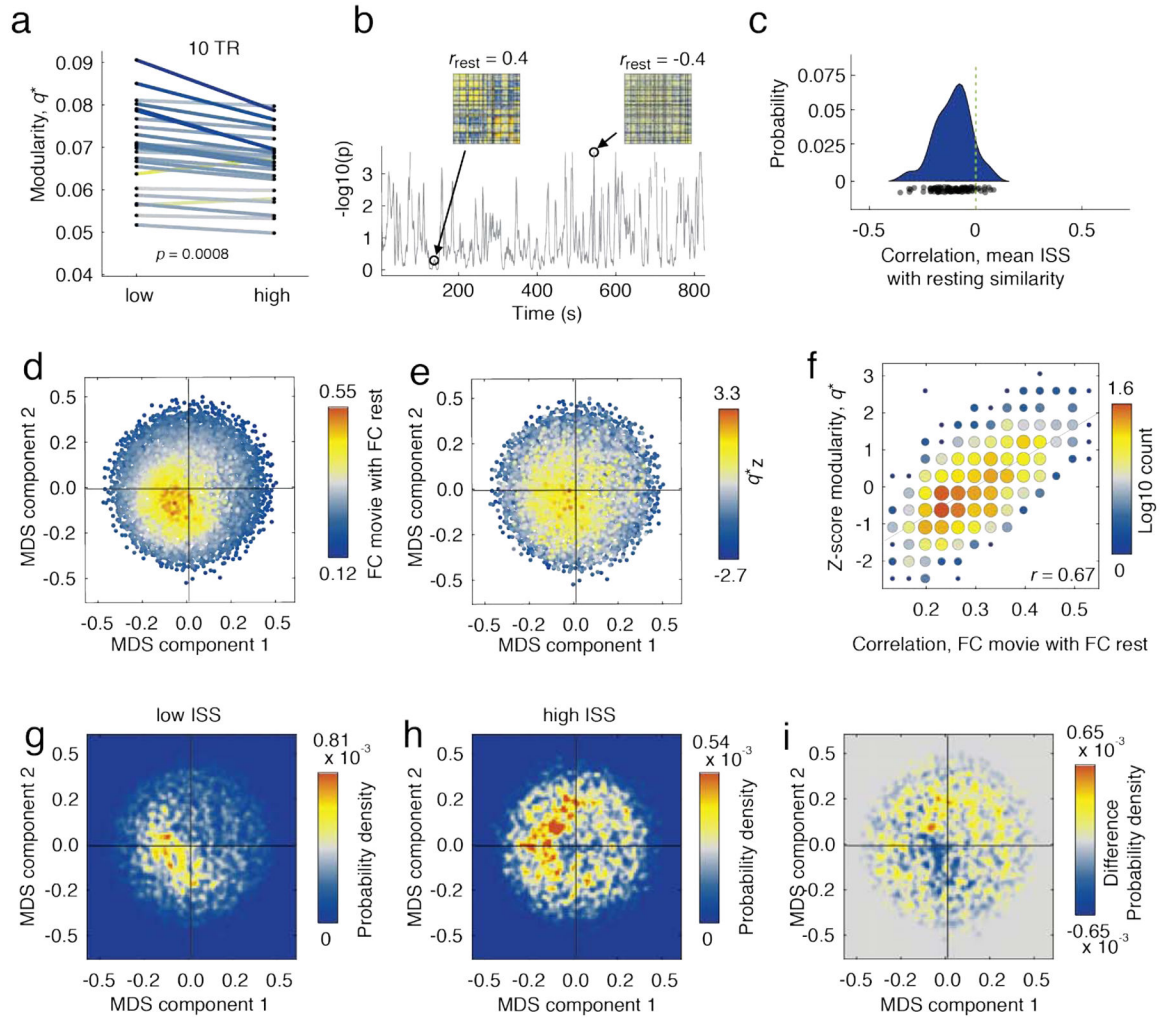
**FIG. 2. Schematic illustration of inter-subject similarity.**

We compare network structure between pairs of subject and across time using the inter-subject similarity (ISS) metric. To calculate ISS, we first divide fMRI BOLD time series into overlapping windows and, using only the samples that fall within a window, estimate each window's FC. Given networks estimated for two subjects at time  $t$ , we calculate ISS by vectorizing the upper triangle elements of each network and computing the Pearson correlation of those elements. We repeat this procedure for each window in time and for all pairs of subjects. This procedure generates a time-varying estimate of the similarity between subjects' networks.



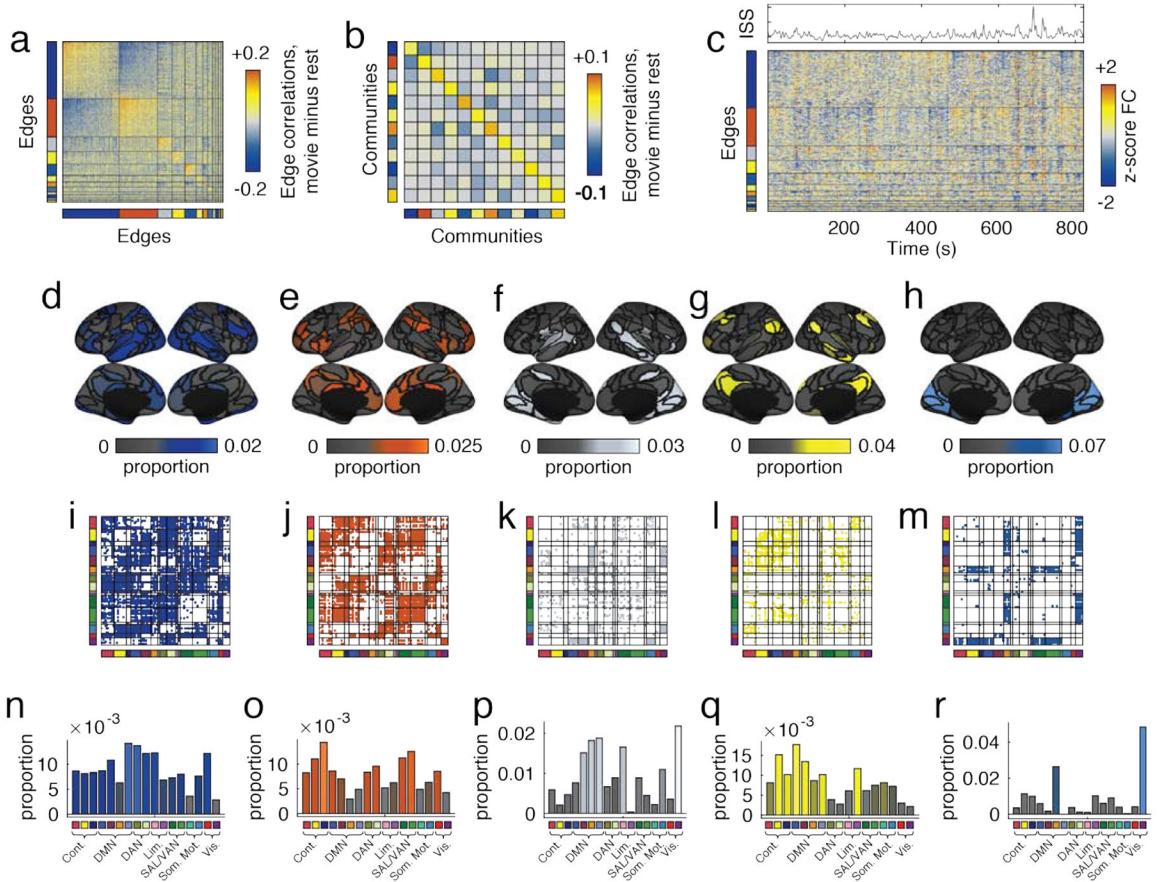
**FIG. 3. Statistical analysis of inter-subject similarity.**

(a) Distribution of inter-subject similarity (ISS) across time during a representative movie-watching scan. (b) Distribution of inter-subject similarity (ISS) across time during rest. (c) We developed a statistical test to quantify the probability of observing a mean ISS value during movie-watching by chance, where we defined chance based on ISS values at rest averaged across time. This procedure associated each moment in time with a  $p$ -value, against which we performed statistical testing. The three dotted lines represent different critical values, ranging from  $p = 0.05$  (uncorrected for multiple comparisons) to an adjusted  $p$ -value after fixing false discovery rate at  $q = 0.01$  (1%). When we applied these critical values to movie-watching data, we consistently identify periods of time that exceed the criterion for statistical significance. (d) Applying the same criteria to resting data (using ISS values from 3/4 scans to estimate  $p$ -values for the remaining scan), we find no points in time that exceed the critical value.



**FIG. 4. Differences in modular architecture during periods of high and low inter-subject similarity.**

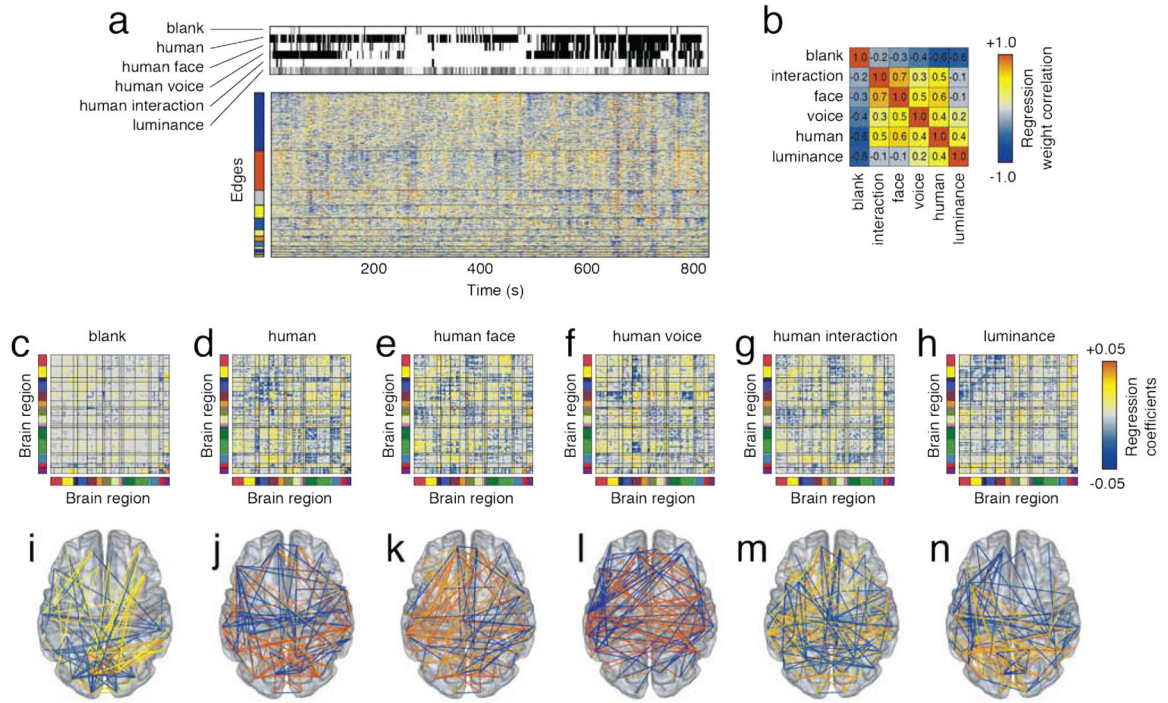
(a) Differences in modularity during low *versus* high inter-subject similarity. (b) We calculated the correlation of subjects' functional connectivity at rest with their time-varying functional connectivity for every instant in time. (c) Distribution of correlation coefficients from previous analysis. We used multi-dimensional scaling to embed subjects' functional connectivity matrices in two dimensions. In panels *d* and *e* we show the same set of points, each corresponding to functional network and colored according to their similarity with resting functional connectivity and modularity. In panel *d*, for instance, brighter points indicate greater similarity to resting FC. Similarly, in panel *e*, brighter points indicate greater levels of modularity. (f) Scatterplot of resting correlations with modularity. (g) The density of points corresponding to low inter-subject similarity coincides with functional networks that are both modular and correlated with static rest. (h) Points corresponding with high inter-subject similarity are distributed along the perimeter. (i) Difference in distributions.



**FIG. 5. Detecting edge communities.**

Panels *a* and *b* depict edge correlation matrix and the community-averaged edge correlation matrix. Rows and columns are ordered by community. Here, we exclude small and singleton communities. (*b*) Edge time series organized by community label. To better understand which parts of the brain were associated with each of the five communities, we calculated the fraction of all edges assigned to community that were incident upon each brain region. Panels *d-h* show these proportions projected onto cortical surfaces. We also visualize the edge communities by labeling edges in the connectivity matrix by community (panels *i-m*) and by aggregating node proportions according to cognitive system (panels *n-r*).





**FIG. 6. Relating movie features to time-varying CC.**

We hand-coded and tracked six features across all movies. These included blank screens, presence of a human, a human face, human voice, interactions between humans, and the luminance of the projected image. (a) Subject-averaged time-varying FC alongside feature time series. We used linear models to identify combinations of features that predicted subject-averaged time-varying FC. In (b), we show the correlation of regression weights for each of the six features. (c - h) Regression coefficients plotted in matrix form. Bright red and yellow elements indicated that the presence of a feature resulted in increased FC, while dark blue indicated that the presence of a feature resulted in decreased FC. (i - n) The top 2% elements (by absolute value) from each of the six matrices.