communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-025-09165-7

Feature similarity, a sensitive method to capture the functional interaction of brain regions and networks to support flexible behavior

Xiuyi Wang[1,2,3] ✉, Baihan Lyu[1,2,3], Katya Krieger-Redwood[4], Nicholas E. Souter[5], Golia Shafiei[6], Nan Lin[1,2,3], Jonathan Smallwood[7], Elizabeth Jefferies[4] ✉ & Yi Du[1,2,3] ✉

The brain is a dynamic system where complex behaviours emerge from interactions across distributed regions. Accurately linking brain function to cognition requires methods sensitive to these interactions. We introduce Feature Similarity (FS), which integrates a broad set of interpretable time-series features —such as covariance, temporal dependencies, and entropy —to move beyond traditional single-metric approaches. FS captured functional brain organization: regions within the same network showed greater similarity than those in different networks, and FS identified the principal gradient from unimodal to transmodal cortices. Compared with Pearson correlation-based functional connectivity (FC) and 46 out of 49 statistical pairwise interaction metrics (SPIs), FS demonstrated greater sensitivity to task modulation. Critically, FS revealed a task-dependent double dissociation in the Dorsal Attention Network, interacting more strongly with the Visual network during working memory but with the default mode network during long-term memory. FS thus provides a powerful tool for uncovering task-specific brain network interactions.

A central aim of cognitive neuroscience is to understand how brain regions interact to support flexible behaviour. The increasing availability of large-scale datasets, where participants perform multiple cognitive tasks, provides a valuable opportunity to address this question. However, existing methods typically focus on a single measure of interaction, such as co-variability, failing to capture the multidimensional nature of brain interactions[1]. As a result, these methods struggle to reliably track task-dependent changes in network interactions. Since the choice of interaction metric fundamentally shapes our understanding of the brain's functional organization, there is a critical need for a more comprehensive approach that integrates multiple features to capture context-dependent functional interactions.

Current neuroimaging methods typically measure specific aspects of functional interactions but overlook their complex, multidimensional nature. Many rely on temporal correlation, assessing neural synchrony but with notable limitations. For example, within-subject functional connectivity (FC, here operationalized as Pearson correlation–based coupling between regional time series in this study) cannot distinguish stimulus-driven from intrinsic correlations, reducing task sensitivity[2], while inter-subject FC (ISFC) isolates stimulus-driven effects but requires identical stimuli across participants, limiting its applicability[2]. Psycho-physiological interactions assess task-dependent connectivity but rely on predefined seed regions and assume a linear relationship with task conditions, restricting flexibility[3]. Dynamic causal modelling infers directed interactions but requires a priori model specification, assumes stationarity, and is computationally limited to small networks[4]. Recently, many statistics of pairwise interaction (SPIs) from fields like Earth system and finance have been applied to brain data[1,5], yet they still focus on single interaction features (e.g., causality, co-variability) rather than integrating multiple dimensions. Consequently, these unidimensional metrics often yield inconsistent or contradictory findings. For instance, frontoparietal control network (FPCN) exhibits negative FC with default mode network (DMN)[6], suggesting opposing functions, yet both share similar timescales[7], implying shared functional properties. Such

[1]State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China. [2]Institute of Psychology, Chinese Academy of Sciences, Beijing, China. [3]Department of Psychology, University of Chinese Academy of Sciences, Beijing, China. [4]Department of Psychology, University of York, Heslington, York, UK. [5]School of Psychology, University of Sussex, Brighton, UK. [6]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [7]Department of Psychology, Queens University, Kingston, ON, Canada.
✉e-mail: wangxiuyi@psych.ac.cn; beth.jefferies@york.ac.uk; duyi@psych.ac.cn

inconsistencies underscore the limitations of unidimensional approaches and highlight the need for a multidimensional method sensitive to task-dependent modulations.

Recent advances have introduced multi-metric approaches to functional connectivity, such as similarity network fusion[8], frameworks that integrate connectivity with nodal properties[9], composite FC metrics derived from curated statistics[10], and multi-metric analyses that enhance clinical group discrimination[11]. These studies highlight the value of combining metrics and provide a foundation for the present work.

Feature Similarity (FS) extends this line of work by leveraging a large feature space (>7000 interpretable time-series features from the hctsa library; Fulcher, 2018; Fulcher et al., 2013) (see Fig. 1). Examples include autocorrelation (temporal dependence/intrinsic timescale), entropy (signal complexity and predictability), and variance or skewness (distributional properties). For each region, these features form a profile of local dynamics, and FS between two regions is defined as the Pearson correlation between their feature profiles[12]. In this way, FS is a pairwise metric derived from local dynamics that indexes the similarity of temporal fingerprints across regions. Whereas FC captures synchronous co-fluctuations of raw time series, and SPIs offer diverse but single-property interaction statistics (e.g., causality, co-variability, nonlinear dependence), FS captures resemblance in underlying dynamical profiles, providing a complementary view of brain interactions.

Individual features also correspond to distinct functional roles. For instance, early visual areas and DMN occupy opposing ends of the timescale hierarchy: early visual areas exhibit rapid autocorrelation decay, whereas DMN regions show gradual decay[2,7,13–15]. This timescale difference underlies their functions: early visual representations are minimally influenced by prior knowledge, while DMN activity is strongly shaped by it[16]. Similarly, transmodal regions such as FPCN and DMN, though often negatively correlated in FC[6], share long timescales[7] and process information over extended periods, supporting higher-order functions like maintaining task goals[16,17]. These examples illustrate how multidimensional features reveal functional commonalities beyond those captured by FC or other unidimensional SPIs.

Crucially, features also vary across tasks. For example, neural timescales shorten in transmodal regions during story comprehension but lengthen during motor and working memory tasks[18,19]. Such shifts reflect the brain's adaptive reconfiguration and suggest that multidimensional features are particularly sensitive to task influences. In line with this, FPCN exhibits greater FS with the dorsal attention network (DAN) during working memory tasks and FS with the memory control network during long-term memory tasks[6]. Motivated by this, we apply FS as a hypothesis-free, data-driven approach to test how inter-regional interaction patterns reconfigure across diverse tasks, with the goal of detecting subtle brain interactions that may be overlooked by conventional connectivity metrics. While temporal correlation (FC) and other unidimensional measures (e.g., SPIs) may appear stable, underlying properties such as signal complexity, variability, and autocorrelation may reveal greater task sensitivity. By integrating these dimensions, FS may provides a richer lens on brain interactions than unidimensional approaches.

To evaluate FS's sensitivity in capturing task-dependent network interactions, we examined the functional relationship between DAN and DMN. Traditionally, DAN, which directs external attention, has been considered antagonistic to DMN, which supports internally directed cognition such as memory retrieval and semantic processing[20]. This opposition is reflected in their typical anticorrelation and DAN's strong FC with the Visual network, reinforcing its role in sensory-driven attention. However, recent evidence challenges this strict dichotomy. Using multiple metrics, DAN has been shown to be topographically positioned between the Visual network and DMN, allowing it to flexibly interact with both based on task demands[6]. Moreover, DMN-DAN interactions are integrated into broader network interactions involving the FPCN[21], which shares properties with both DMN and DAN and plays a pivotal role in flexible cognition by bridging external attention and internal processing. FPCN cooperates with DMN via the memory control network during long-term memory tasks but not during working memory tasks (see Fig. 2)[6], suggesting that DAN, too, may align more with DMN in long-term memory tasks, shifting from external sensory attention to internally guided retrieval. Our results demonstrate that FS captures these task-driven shifts more effectively than FC and 46 out of 49 SPIs, revealing a double dissociation – defined here as opposing patterns of interaction across two conditions: DAN interacts more with the Visual network during working memory tasks and with DMN during long-term memory tasks—a distinction that FC and most SPIs fail to detect. These findings highlight FS's unique ability to uncover nuanced, task-dependent brain interaction, establishing it as a powerful tool for studying cognitive flexibility in brain networks.

Results

The results are divided into three sections. (i) First, we took an existing individualized parcellation of the cortex to identify the parcels of each network and examined whether FS is equally capable of capturing the network structure as FC does. As expected, regions belonging to the same intrinsic functional networks had greater FS compared to regions in different networks. (ii) Next, we examined whether FS could capture the organizational principles of the cortex as is seen in FC. We found FS captured the principal intrinsic connectivity gradient that separates sensory-motor regions from transmodal areas, as well as the second component that separates somatomotor and auditory from visual cortex, which have been previously described using decompositions of FC. (iii), Finally, we examined whether FS is more sensitive to task-induced changes in brain interactions than FC and most SPIs by examining how these metrics characterise network interactions across working memory and long-term memory tasks. Specifically, we asked whether FS is superior than FC and SPIs in their capacity to identify differences in the way that DAN interacts with Visual network and DMN during working memory and long-term memory tasks.

The FS method consists of the following steps (see Fig. 1):

Step 1: Extract a diverse set of interpretable features from the time-series data. In this study, FS was implemented using more than 7000 features from the hctsa library[22,23], encompassing distributional properties (e.g., skewness), temporal dependencies (e.g., autocorrelation), and predictability (e.g., entropy).

Step 2: Calculate FS between brain regions by calculating the Pearson correlation coefficients of the extracted features. FS is a pairwise metric derived from local intra-regional dynamics to assess inter-regional interactions, rather than a measure of local dynamics alone.

Step 3: Examine how FS varies across different tasks.

To estimate FC, we computed Pearson correlation coefficients between regional time series for each run of each task and each participant, and applied xDF correction to account for temporal autocorrelation within time series and instantaneous/lagged cross-correlations, thereby yielding more precise and reliable estimates. For both FC and FS, this produced a 400 × 400 matrix per run per participant, with all values Fisher-z transformed prior to averaging. Participant-specific matrices were obtained by averaging across runs within each task, and group-average matrices were derived by averaging across participants when required.

Regions belonging to the same intrinsic functional network have greater FS compared to regions in different networks

We first tested whether parcels within the same network exhibited similar features using participant-specific matrices. We selected Kong et al.' parcellation[24] as it allows for individual-specific parcellations with greater homogeneity[6] while capturing the inherent heterogeneity of large-scale networks like DAN and FPCN. Regions were considered similar in terms of features if the features of two regions were significantly correlated[25]. FS captures regions with similar dynamical properties but not necessarily synchronous activity (Fig. 1). For example, while FPCN-A and DMN both have longer timescales, they often exhibit negative FC[6].

As expected, regions within the same functional network showed higher FC compared to those between networks, both at rest ($t(244) = 136.09$, $p < 0.001$) and during tasks (spatial working memory,

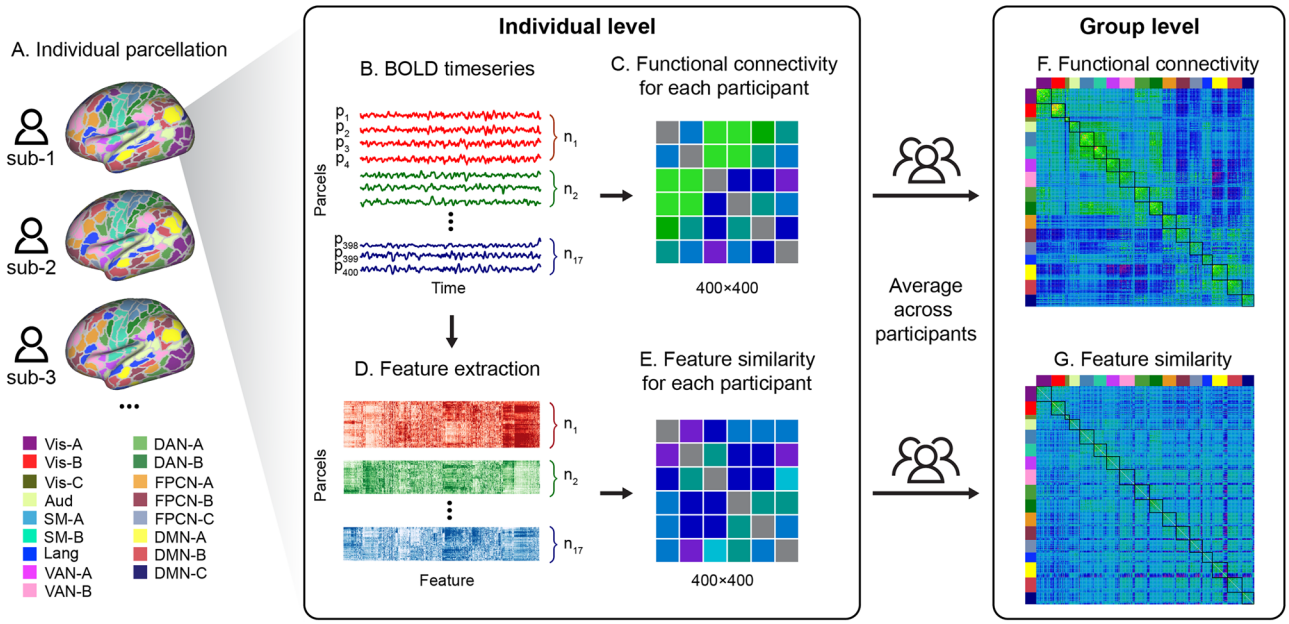

Fig. 1 | The workflow of the FS and FC analysis. A Individual-specific parcellation divided the whole brain into 400 parcels across 17 networks[24]. **B** Average time series of each parcel. **C** Functional connectivity involved calculating Pearson correlation coefficients between the time-series of parcels for each participant. **D** Extraction of features of time series for each parcel. **E** Pearson correlation coefficients of the extracted features represent the pairwise feature similarity between all possible combinations of brain parcels. **F**, **G** Group-level FC and FS matrices were obtained by averaging participant-level matrices across runs and subjects. (Vis = Visual, Aud = Auditory, SM = Sensory-motor, DAN = Dorsal attention network, VAN = Ventral attention network, FPCN = Fronto-parietal control network, Lang = Language, DMN = Default mode network; p1–p400 represent the 400 parcels, and n1–n17 represent the 17 networks in the analysis).

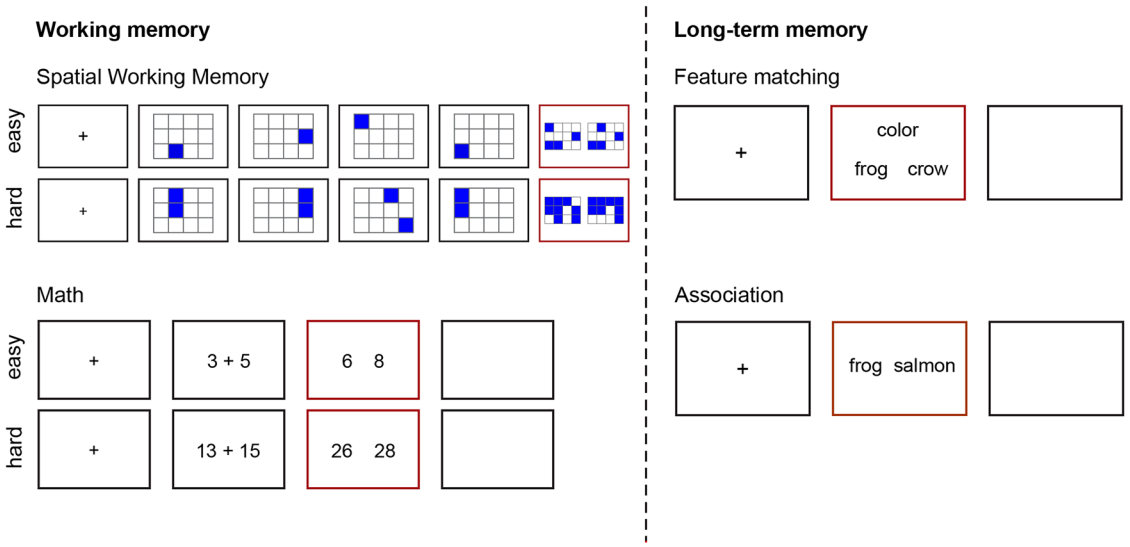

Fig. 2 | The experimental design. To tap working memory, we included two tasks: a spatial working memory task required participants to keep track of sequentially presented locations, while math decisions involved maintaining and manipulating numbers which rely more on working memory. To tap long-term memory, we included two tasks that required controlled retrieval of knowledge; a semantic feature matching task required participants to match probe and target concepts according to a particular semantic feature (colour or shape), while a semantic association task involved deciding if pairs of words were linked in meaning. Response periods are indicated by a red box.

$t(25) = 26.33$, $p < 0.001$; math, $t(25) = 26.36$, $p < 0.001$; semantic feature matching, $t(27) = 70.02$, $p < 0.001$; semantic association, $t(29) = 57.60$, $p < 0.001$) (Fig. 3). Similarly, FS was higher within networks compared to between networks, both at rest ($t(244) = 91.24$, $p < 0.001$) and across tasks (spatial working memory, $t(25) = 18.35$, $p < 0.001$; math, $t(25) = 14.51$, $p < 0.001$; semantic feature matching, $t(27) = 42.27$, $p < 0.001$; semantic association, $t(29) = 38.04$, $p < 0.001$) (Fig. 3). All p-values were FDR-corrected. These findings suggest that FS can capture intrinsic network structure traditionally captured by FC.

We also assessed the similarity between FC and FS by correlating the FC and FS matrices at rest and during tasks. We observed weak but significant positive correlations between FC and FS as indicated by arrows (Fig.3) at rest ($r = 0.29$, $p = 0.0001$) and during tasks (spatial working memory, $r = 0.37$, $p = 0.0001$; math, $r = 0.33$, $p = 0.0001$; semantic feature matching, $r = 0.33$, $p = 0.0001$; semantic association, $r = 0.33$, $p = 0.0001$). All p-values were FWE-corrected using maximum r values (Method 4.5.4). The positive correlations suggest that regions with similar time-series features exhibited coherent spontaneous fluctuations. However, the largest

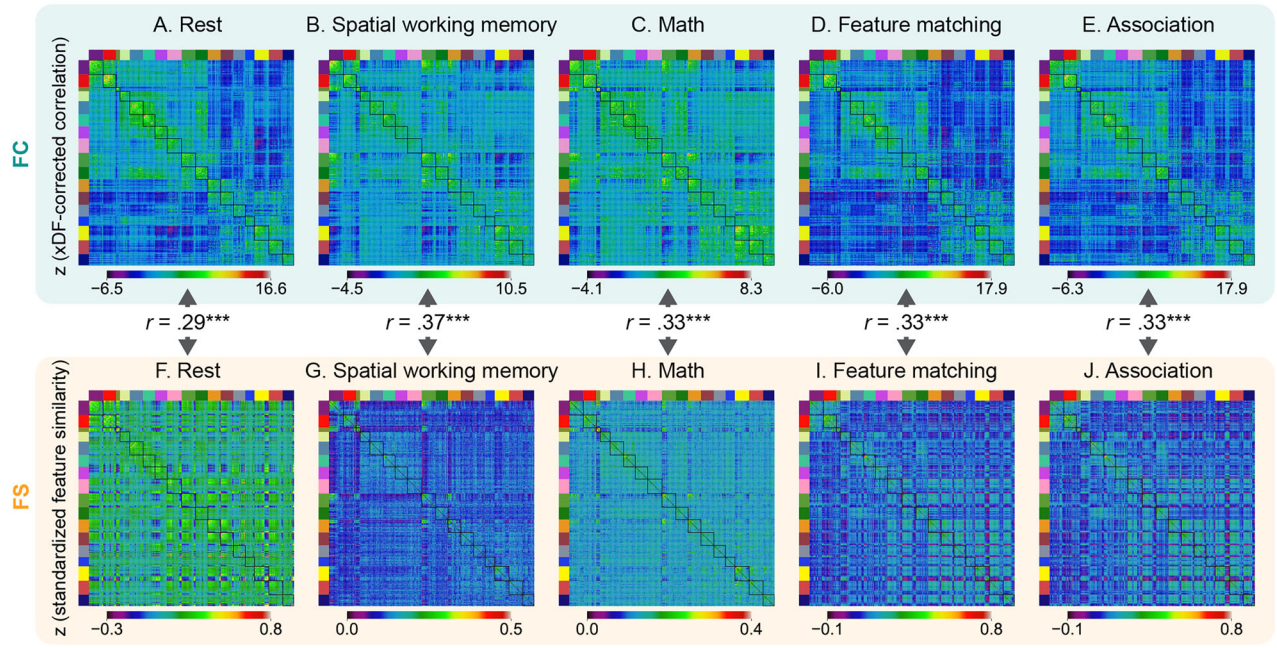

Fig. 3 | FC and FS matrices at rest and during each task. Regions within the same intrinsic functional network showed greater FC and FS than regions in different networks across all tasks and rest. FC and FS were positively but weakly correlated in each condition ($p < 0.0001$). Colorbars indicate z-scored values for each measure: FC panels (**A**–**E**) display z (xDF-corrected correlation) values, reflecting correlation strength corrected for temporal autocorrelation using the xDF method, while FS panels (**F**–**J**) display z (standardized feature similarity) values, reflecting z-scored feature-similarity measures without xDF correction. Differences in value ranges reflect the use of xDF correction for FC only.

correlation was 0.37, indicating that these metrics also capture different information.

FS reveals unique organizational principles beyond FC

We then asked whether FS and FC captured similar or distinct components by examining group-average matrices at rest. Firstly, we performed dimension reduction analysis (diffusion embedding)[26] on the resting state FC matrix for the HCP dataset (Fig. 3A). We focused on the three components with the largest eigenvalues, as they explain 28.02% variance in total and have clear interpretation (see Fig. 4I for scree plot). Consistent with previous studies[26–29], the first component explaining 12.75% of the variance corresponded to the principal gradient described by Margulies et al.[26]. This component separated sensory-motor regions (shown in purple-blue in Fig. 4A) from transmodal areas (shown in red-white in Fig. 4A). The second component explained 11.29% of the variance and separated somatomotor and auditory cortex (shown in purple-blue in Fig. 4B) from visual cortex (shown in red-white in Fig. 4B). The third component explained 3.98% of the variance and separated FPCN regions (shown in purple-blue in Fig. 4C) from DMN regions (shown in red-white in Fig. 4C).

Similarly, we performed dimension reduction analysis on the resting state FS matrix for the HCP dataset (Fig. 3F), using the procedures above. The three components with the largest eigenvalues explained 23.23% of the variance (see Fig. 4I for scree plot). FS captured two components that were similar to the results for FC, and one not uncovered by FC (Fig. 4). The first principal component of FS explained 12.19% of the variance and had a distinct spatial distribution from the first three components of FC: it did not show significant correlations with any of the components of FC ($p > 0.05$; Fig. 4D). The next two components of FS were seen in FC data (Fig. 4). The second component of FS (explaining 7.11% of the variance; Fig. 4E) aligned with the principal gradient of FC (Fig. 4A), separating sensory-motor from transmodal cortex, with a strong correlation between components ($r = 0.89$, $p = 0.0001$, spin permutation corrected). The third component of FS (explaining the 3.94% variance; Fig. 4F) corresponded to the second component of FC (Fig. 4B), separating somatomotor and auditory cortex from visual cortex, with a strong correlation between components ($r = 0.87$, $p = 0.0001$, spin permutation corrected). These findings further confirm that FS captures meaningful information, i.e., organizational principle.

To gain insight into the first component of FS, we investigated its correlation with the map of intrinsic timescale, which reflects the temporal duration of ongoing inputs that the brain can process. The first component resembles the distribution of intrinsic timescale reported by Raut et al.[7],which is longest in transmodal regions[18]. We therefore estimated intrinsic timescale for each parcel by measuring the decay of the temporal autocorrelation function and quantifying the time taken for the autocorrelation function to reach a threshold of $r = 0.5$, which is half of the full width at half maximum. Higher values indicate longer processing times for ongoing inputs. Our results align with Raut et al.[7], showing short timescales in insula and cingulate cortex, and long timescales in angular gyrus, posterior cingulate cortex, and frontal pole. (Fig. 4G). We found that the first principal component, which explained the most variance in FS, was significantly correlated with the intrinsic timescale map (left hemisphere: $r = 0.82$, $p = 0.0001$; right hemisphere: $r = 0.80$, $p = 0.0001$; Fig. 4H, corrected for spatial autocorrelation using spin permutation), indicating that FS can capture meaningful information not captured by traditional FC, such as the intrinsic timescale of the brain's response.

FS showed greater variation across tasks than FC

The above analyses showed that FS captured similar information to FC but also additional dimensions of cortical organisation. Next, we examined whether FS was more sensitive to task modulation than FC. FC measures pairwise co-fluctuation between regional time series (Pearson correlation). FS, by contrast, is a pairwise temporal similarity computed on per-region feature profiles that summarize local dynamics (e.g., entropy, autocorrelation, stationarity, spectral slope). FS therefore bridges local and pairwise levels by comparing local dynamical fingerprints across regions and integrates multiple feature dimensions, rather than targeting a single interaction property as in many SPIs. Given that tasks impose specific cognitive demands and engage distinct neural processes, the brain's functional

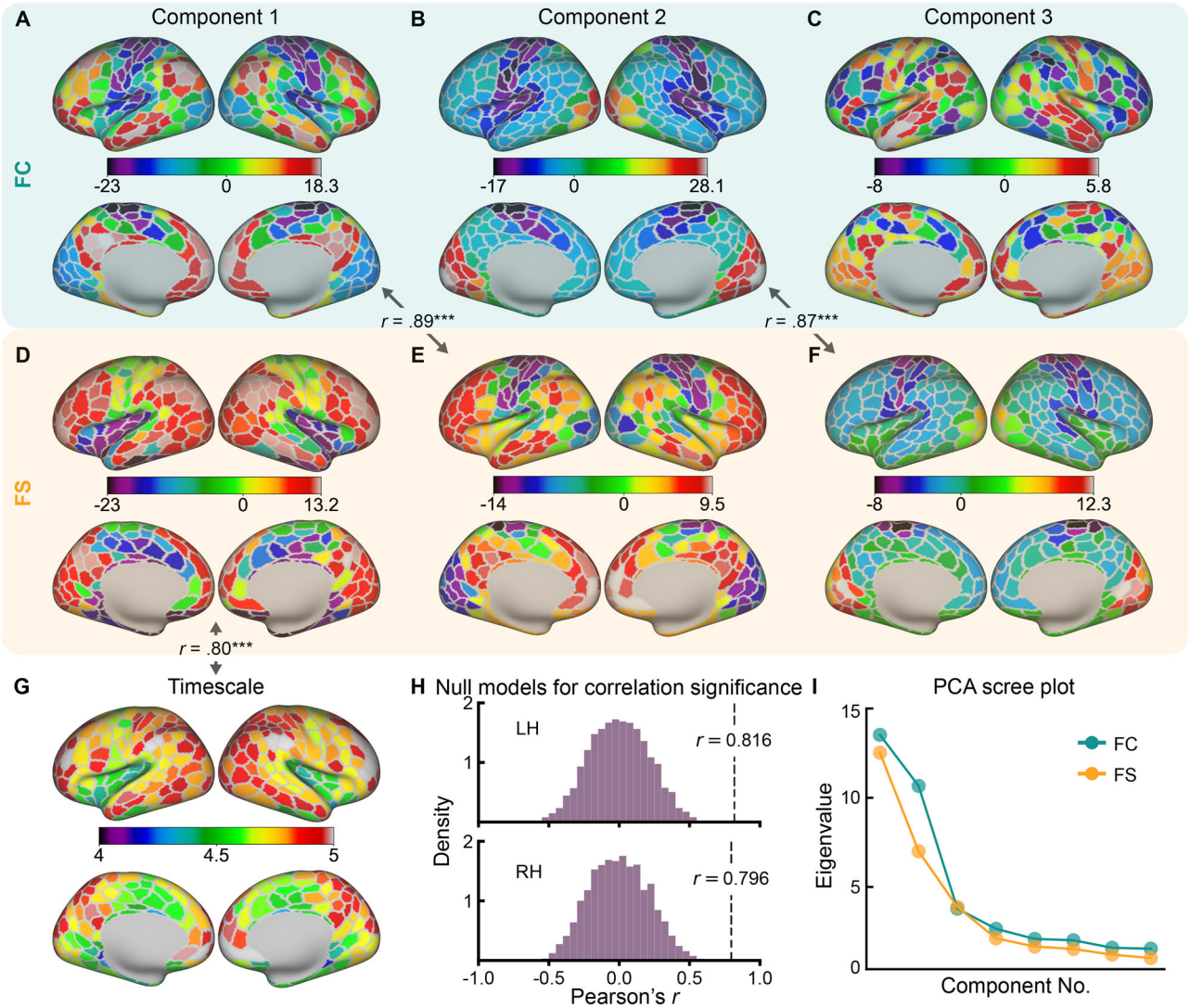

Fig. 4 | The top three principal components of FC and FS revealed two similar components and one distinct component. The first principal component of FC (**A**) corresponds to the second principal component of FS (**E**), as indicated by the arrows denoting a strong correlation ($r = 0.89$, $p = 0.0001$, corrected for spatial autocorrelation using spin permutation), which separates sensory-motor regions from transmodal areas. Similarly, the second principal component of FC (**B**) corresponds to the third principal component of FS (**F**), with a strong correlation ($r = 0.87$, $p = 0.0001$, corrected for spatial autocorrelation using spin permutation), separating somatomotor and auditory cortex from visual cortex. These findings suggest that FS captures similar organizational information to FC. However, the third principal component of FC (**C**), which separates FPCN regions from DMN regions, was not captured by FS. The first principal component of FS (**D**), which corresponds to the intrinsic timescale gradient (**G**), was not captured by FC. **G** The intrinsic timescale map shows short timescales in insula and cingulate cortex, and long timescales in angular gyrus, posterior cingulate cortex, and frontal pole. **H** The first principal component of FS was significantly correlated with the intrinsic timescale gradient map in both the left ($r = 0.82$, $p = 0.0001$) and right ($r = 0.80$, $p = 0.0001$) hemispheres (corrected for spatial autocorrelation using spin permutation). The histograms illustrate the null model distributions. **I** Scree plots showing the eigenvalues for the top eight principal components of FS (orange) and FC (blue).

organization undergoes modulation at multiple levels. Temporal correlation (as captured by FC) may remain stable under certain conditions, but the underlying features of the time-series, such as signal complexity, variability, and temporal autocorrelation, may exhibit more sensitivity to these changes. This multidimensional analysis may allow FS to detect subtle changes in brain interactions during tasks that may be missed by FC.

We computed the Pearson correlation coefficients between the group-average task mean FC matrices (and, analogously, FS matrices) for every possible task pair. For example, we calculated the Pearson correlation coefficients of the FC matrices of the spatial working memory task (Fig. 3B) and math task (Fig. 3C). Similarly, we calculated the Pearson correlation coefficients of the FS matrices of these two tasks (Fig. 3G, H). Then we directly compared these two r values. We found correlations for FC were consistently higher than those for FS across all 10 task pairs (Fig. 5A; $p < 0.05$; see Table 1 for statistics), suggesting that FS may be more sensitive to task differences than FC. Each data point in Fig. 5A reflects one task-pair comparison based on group-averaged FC or FS matrices (rather than individual subjects); the apparent extreme value corresponds to the semantic feature matching versus semantic association pair and is not an outlier.

We further confirmed the task sensitivity of FS at the participant level by comparing FS and FC correlations between two working memory tasks using each participant's task-level matrices. Since the same participants completed both the spatial working memory task and the math task, we calculated the Pearson correlation coefficients between the FC matrix of the spatial working memory task and that of math task for each participant and transformed the resulting r values to z values using Fisher's transformation. Similarly, we also calculated the Fisher's z values between the FS matrices of the two working memory tasks for each participant. Finally, we compared the z values of FC with those of FS by conducting paired t-tests. We found

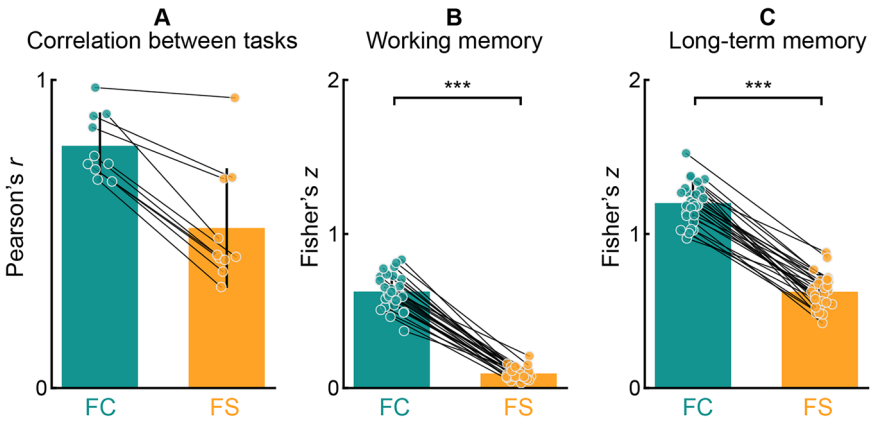

Fig. 5 | FS showed greater variation across tasks than FC. A Group-average FC matrices showed higher correlations across tasks compared to FS matrices. Each dot represents the Pearson correlation coefficient between two group-average task matrices, across the 10 possible task pairs (see Table 1). Because this analysis is based on task pairs rather than participant-level statistics, significance is not indicated with asterisks, although FC correlations were consistently higher than FS correlations ($p < 0.001$). **B** Correlations of FC matrices between the spatial working memory and math tasks were greater than those of FS matrices. Each dot represents the Fisher-z–transformed correlation coefficient between the two task-specific matrices of one participant (n = 27 biologically independent participants). **C** Correlations of FC matrices between the semantic feature matching and semantic association tasks were significantly greater than those of FS matrices. Each dot represents the Fisher-z–transformed correlation coefficient between the two task-specific matrices of one participant ($n = 28$ biologically independent participants).

Table 1 | FS showed greater variation across tasks than FC

Task 1	Task 2	r (FC)	r (FS)	z	p
Rest	Spatial working memory	0.68	0.33	96.31	<0.001
Rest	Math	0.67	0.38	82.64	<0.001
Rest	Feature matching	0.86	0.68	82.48	<0.001
Rest	Association	0.88	0.68	110.49	<0.001
Spatial working memory	Math	0.75	0.43	103.38	<0.001
Spatial working memory	Feature matching	0.73	0.42	94.50	<0.001
Spatial working memory	Association	0.71	0.42	88.63	<0.001
Math	Feature matching	0.75	0.43	103.38	<0.001
Math	Association	0.73	0.43	93.72	<0.001
Feature matching	Association	0.98	0.94	86.35	<0.001

FC correlations across tasks were consistently higher than FS correlations. Across all task pairs, correlations based on FC were significantly higher than those based on FS (all $p < 0.001$), indicating that FS is more sensitive to task differences than FC.

that the correlations of FC were greater than the correlations of FS for the working memory tasks (Fig. 5B; $t = 26.58$, $p < 0.001$). We then conducted the same analysis for the two semantic tasks, semantic feature matching and association, and observed the same pattern (i.e., stronger correlation across tasks for FC than for FS; Fig. 5C, $t = 23.29$, $p < 0.001$). The smaller correlation of FS is not because the data of FS are noisier. By contrast, FS carries meaningful task information, as FS matrices correctly classify task labels for working memory tasks (accuracy = 0.83, $p = 0.0001$) and long-term memory tasks (accuracy = 0.78, $p = 0.0001$), both significantly above chance level (0.5) (see Supplementary Material 2.1 for detailed results). These findings suggest that FS is more sensitive to task modulation than FC, because it incorporates multiple aspects of time-series similarity beyond temporal co-fluctuation. Despite comparable classification accuracy, FC showed greater cross-task correlations than FS. This highlights the sensitivity of FC to common connectivity patterns across tasks, while FS appears to capture finer task-specific variations in interaction patterns, providing a complementary perspective on task modulation.

FS captured network interaction patterns across tasks not captured by FC

After showing that FS was generally more sensitive to task modulation than FC, we examined whether FS could capture the varying network interaction patterns across tasks that missed by FC using the participant-level task matrices, with group effects tested across participants. We tested the interaction difference between DAN and Visual network versus DAN and DMN, to test the hypothesis that DAN is more similar to Visual network in working memory tasks, which rely more on visual features and working memory, and more similar to DMN in long-term memory tasks, which rely more on long-term memory. We expected that this pattern would be seen more readily in FS, since this metric was more sensitive than FC to task demands above.

Since the tasks were presented visually in the current study, we selected DAN-A[24], which is typically engaged together with Visual network and showed greater feature similarity with Visual network than other attention networks (DAN-B, Ventral attention network) in a classification analysis (Supplementary Fig. 1). We compared FC and FS between DAN-A and Visual network versus DAN-A and DMN across tasks, focusing on differences in the strength of these connections. For simplicity, we combined the three visual subnetworks into a single visual network and the three default mode subnetworks into a single DMN. While connectivity differences did not vary across tasks when assessed with FC (Fig. 6, $p > 0.05$, FWE corrected), there was a difference between working memory and long-term memory tasks for FS ($p < 0.001$; FWE corrected). DAN-A always showed greater FC with Visual network than with DMN at rest (Fig. 6,

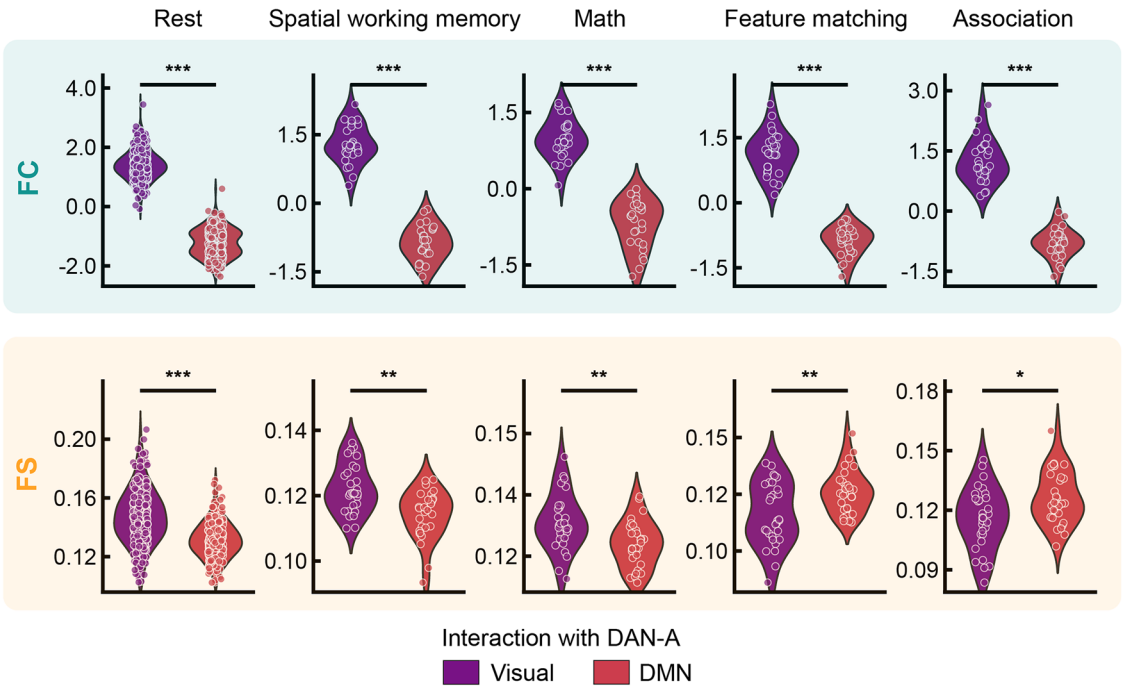

Fig. 6 | FS captured varying interaction patterns for DAN-A across tasks that missed by FC. Top panel: DAN-A always showed greater FC with Visual network than with DMN at rest and for each task and there was no difference across tasks. Bottom panel: The difference between DAN-A and Visual network versus DAN-A and DMN in the working memory tasks was greater than in long-term memory tasks. DAN-A showed greater FS with Visual network than with DMN at rest and for the working memory tasks, but showed the opposite pattern in the long-term memory tasks supported by DMN. Asterisks denote significance levels: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***); $n = 245$ (rest), 27 (spatial working memory), 27 (math), 28 (semantic feature matching), and 30 (semantic association); all represent biologically independent participants.

$t(244) = 53.68$, $p < 0.001$) and for each task (Fig. 6, spatial working memory, $t(25) = 16.19$, $p < 0.001$; math, $t(25) = 11.73$, $p < 0.001$; semantic feature matching, $t(27) = 17.19$, $p < 0.001$; semantic association, $t(29) = 14.95$, $p < 0.001$), consistent with previous studies. However, DAN-A showed greater FS with Visual network than with DMN at rest (Fig. 6, $t(244) = 10.58$, $p < 0.001$) and in working memory tasks (Fig. 6, spatial working memory, $t(25) = 3.55$, $p = 0.004$; math, $t(25) = 3.19$, $p = 0.003$), but showed the opposite pattern in long-term memory tasks (Fig. 5, semantic feature matching, $t(27) = -2.86$, $p = 0.008$; semantic association, $t(29) = -2.64$, $p = 0.01$). These findings suggest that FS is more sensitive to interaction patterns between networks that reflect differing cognitive demands. We obtained consistent interaction patterns when FC was computed either on task-residual time series after FIR regression or directly on parcel time series, confirming that preprocessing differences do not account for the greater task sensitivity of FS.

The task sensitivity of FS was further confirmed when we examined the interaction differences between the domain general control network (FPCN-A) and Visual network versus FPCN-A and DMN. Interaction differences for FPCN-A across tasks were only captured by FS and not by FC (Fig. S2; See Supplementary Material 2.3 for detailed information). There was no difference in FC across tasks for FPCN-A ($p > 0.05$, FWE corrected), with FPCN-A always showing greater FC with Visual network than with DMN at rest ($t = 2.39$, $p = 0.02$) and for each task (spatial working memory, $t = 11.36$, $p < 0.001$; math, $t = 11.44$, $p < 0.001$; semantic feature matching, $t = 6.50$, $p < 0.001$; semantic association, $t = 3.62$, $p = 0.001$). However, FS revealed different patterns across tasks, with a greater difference between FPCN-A and Visual network compared with FPCN-A and DMN for the working memory than long-term memory tasks ($p < 0.001$; FWE corrected). Specifically, FPCN-A showed similar FS with Visual network and DMN at rest ($t = 1.39$, $p = 0.17$) and in the math task ($t = -0.73$, $p = 0.47$) but showed greater FS with DMN than with Visual network in the spatial working memory task ($t = -4.97$, $p < 0.001$) and long-term memory tasks (semantic feature matching, $t = -9.12$, $p < 0.001$; semantic association, $t = -8.25$, $p < 0.001$). These findings support that FS is more sensitive to task modulation than FC.

FS captured network interaction patterns across tasks not captured by most SPIs

While FC (Pearson correlation) remains the dominant method for measuring brain network interactions in fMRI, recent studies have shown that other SPIs, previously applied in different fields, can also be used to measure neural interaction since some of them can classify brain states and conditions in fMRI data[1]. To further assess FS's sensitivity relative to these SPIs, we initially selected all 67 SPIs that significantly classified states in the fMRI film dataset and then refined this set to 49 SPIs with reasonable computational requirements (<5 hours per run per task per subject). These SPIs derived from 20 interaction metrics across six categories: (1) Basic Methods, such as precision, which quantifies pairwise associations while controlling for the effects of other time series; (2) Distance-Based Similarity, which measures statistical similarity or independence based on pairwise distances between bivariate observations; (3) Causal Inference, which aims to infer directed relationships; (4) Information-Theoretic Measures, such as mutual information which quantifies the total dependency between two variables; (5) Spectral Measures, commonly computed in the frequency or time-frequency domain. Examples include spectral coherence magnitude, which quantifies the alignment of frequency components in phase and amplitude; and (6) Miscellaneous Methods: this category includes techniques such as linear model fits, which estimate relationships through regression, and other statistical tools that quantify pairwise interactions but do not fit into the above categories (see Methods and Supplementary Material 1.3.3 for full details). We examined whether the double dissociation observed by FS can be replicated by each SPI by comparing the interaction matrices generated by each SPI across tasks, focusing on key networks (DAN-A, Visual, and DMN) using participant-level task matrices. Specifically, we conducted paired t-tests for each task to assess differences between network pairs to examine whether DAN-A showed greater interaction with Visual network

during the working memory task but with DMN during the long-term memory tasks.

We found that 46 out of the 49 SPIs did not fully replicate the double dissociation observed with FS. Of these, 22 lacked task sensitivity: 14 showed consistently stronger interactions between DAN and Visual networks across all tasks, while 8 showed the opposite pattern (Fig. 7A). Representative examples include covariance from the Basic Statistics category and longest common subsequence from the Distance Similarity category (Fig. 7C).

Nine SPIs demonstrated limited sensitivity, each detecting only a single dissociation. Among these, 9 revealed stronger DAN–Visual interactions during working-memory tasks but failed to capture the reverse pattern during long-term memory tasks. An example is mutual information with a Gaussian kernel from the Information Theory category (Fig. 7D). In contrast, 1 SPI showed the opposite single dissociation, detecting stronger DAN–DMN interactions during long-term memory tasks but not the complementary effect during working-memory tasks.

Only 1 SPI, the Hilbert–Schmidt Independence Criterion from the Distance Similarity category, fully replicated the double dissociation, capturing stronger DAN–Visual in teractions during working-memory and stronger DAN–DMN coupling during long-term memory tasks (Fig. 7B). 2 others—Distance Correlation (Distance Similarity) and time-lagged mutual information (Information Theory)—partially replicated the dissociation, failing to reach significance for DAN–DMN interactions during the feature matching task (FDR-corrected $p = 0.05$ and $p = 0.07$).

The remaining SPIs included 14 methods showing a reverse single dissociation, defined as a reversal of the expected interaction pattern in one condition (e.g., stronger DAN–Visual interaction during long-term memory tasks) without the complementary reversal in the opposite condition (working-memory), and 1 method showing a reverse double dissociation. These patterns were difficult to interpret and the challenges of using SPIs to detect task-specific interactions. Moreover, we observed that different parameterizations of the same SPI could yield inconsistent results (e.g., precision with EmpiricalCovariance revealed a reverse single dissociation, whereas ShrunkCovariance showed no task sensitivity). Full statistical results for all 49 SPIs, including their classification (no sensitivity, single dissociation, double dissociation, reverse patterns), are provided in Supplementary Table 2 and Supplementary Data 1. These findings underscore the challenges of selecting appropriate SPIs and parameter settings.

The failure of most SPIs to replicate the double dissociation observed with FS underscores the superior sensitivity of FS in detecting nuanced brain network interactions. Among the 49 SPIs tested, only a few (3 SPIs) successfully replicated the double dissociation, further validating FS's findings and robustness. This dual observation highlights FS's unique ability to capture task-specific interactions that are often missed by other methods, offering critical insights into cognitive flexibility and functional reconfiguration.

Discussion

In this study, we introduce FS as a robust method for unveiling network structures, brain topography, and task-induced modulations. FS validation rests on two key observations: (1) regions within the same intrinsic network exhibit greater FS than regions across networks, and (2) FS captures the principal gradient from unimodal to transmodal cortices. Building on these validations, FS demonstrated heightened sensitivity to task modulation, revealing interaction patterns not detected by FC or by 46 of 49 SPIs. For example, DAN-A showed greater FS with the Visual network during working memory and with the DMN during long-term memory, indicating a functional reorientation across tasks that FC failed to capture.

The 49 SPIs tested provide a representative benchmark, and there is no evidence that untested measures would alter our conclusions. Of these, only three reproduced aspects of the effect: the Hilbert–Schmidt Independence Criterion (HSIC) fully detected it, while Distance Correlation and time-lagged mutual information recovered it only partially. Their sensitivity to nonlinear or temporal dependencies may explain this partial convergence, but each required careful parameter tuning (e.g., the biased HSIC variant detected the dissociation, whereas the original did not), highlighting their limitations. By contrast, FS identified the effect without parameter adjustment, underscoring its strength as a stable and integrative framework for detecting nuanced brain interactions.

Our use of two independent datasets further strengthens these conclusions. HCP data showed that FS recovers stable large-scale principles such as the unimodal–transmodal hierarchy, and these generalized to the smaller York dataset, as indicated by consistent network organization and highly correlated principal gradients[27]. The York dataset, with independent runs for each task, was best suited for estimating task-specific FS and FC. Although tasks differed in TR, run length, and design, these factors applied equally to FC and FS; if they were the main drivers, both metrics would have shown similar patterns. Instead, FS revealed task-related modulations not detected by FC. We do not claim that FS is always more sensitive than FC, but highlight contexts where it reveals modulations beyond FC. Although HCP also includes task data, its block design with relatively short condition durations is not well suited for condition-specific analyses; resting-state results from HCP were included only for completeness in the task modulation section.

FS extends multi-metric approaches by incorporating over 7000 interpretable time-series features into similarity matrices derived from local dynamics. Unlike prior frameworks that combine multiple connectivity measures[8–11], FS's breadth and grounding in local dynamics enable both recovery of stable organizational principles and sensitivity to diverse task demands. Within this growing line of research, FS offers a generalizable and powerful tool for studying brain network interaction. Importantly, FS is not intended to replace FC or SPIs but to complement them. FC remains valuable for characterizing intrinsic coupling patterns shaped by neuroanatomy, while SPIs capture specific interaction properties[1,5]. FS advances these approaches by integrating a broad and interpretable feature space derived from local dynamics.

Several methodological considerations remain. Although derived from pseudo-rest, our individualized parcellations showed higher within-parcel homogeneity than group-level atlases, supporting their robustness while noting future work should compare across parcellation schemes. We did not explicitly control for inter-parcel distance, but the key double dissociations we report (e.g., DAN-A with Visual versus DMN) cannot be explained by spatial proximity alone, and FC—equally affected by distance—did not show these patterns. Incorporating distance correction in future work would nonetheless strengthen FS's interpretability. The present implementation also assumes stationarity, as features were estimated from the full time series. Future extensions could apply FS to shorter windows or event-related designs to capture non-stationary changes in network interactions, and benchmark it against dynamic FC approaches, which track time-varying fluctuations in connectivity and have proven useful for task-based fMRI. Our analyses further focused on edgewise correlations, which capture task-related variations in pairwise interactions but are not directly sensitive to topological reconfigurations (e.g., fragmentation or hub structure). Combining FS with graph-theoretic measures such as modularity, hubness, or global efficiency could therefore provide complementary insights into task-induced network reorganization. While FS builds on established measures such as covariance, coherence, and temporal scale, the present study does not yet reveal which of the ~7000 extracted features drive the observed network dissociations. Thus, FS offers an interpretable framework in principle, but the neurobiological meaning of the current findings remains limited. Future work should identify the critical feature dimensions underlying these effects to better link FS to neural mechanisms.

Collectively, our findings demonstrate that FS offers a generalizable and powerful framework for studying brain network interactions. By enabling the detection of nuanced task-dependent modulations, FS has the potential to resolve inconsistencies in prior research, inspire reanalysis of existing datasets, and contribute to the development of new theoretical frameworks in cognitive neuroscience.

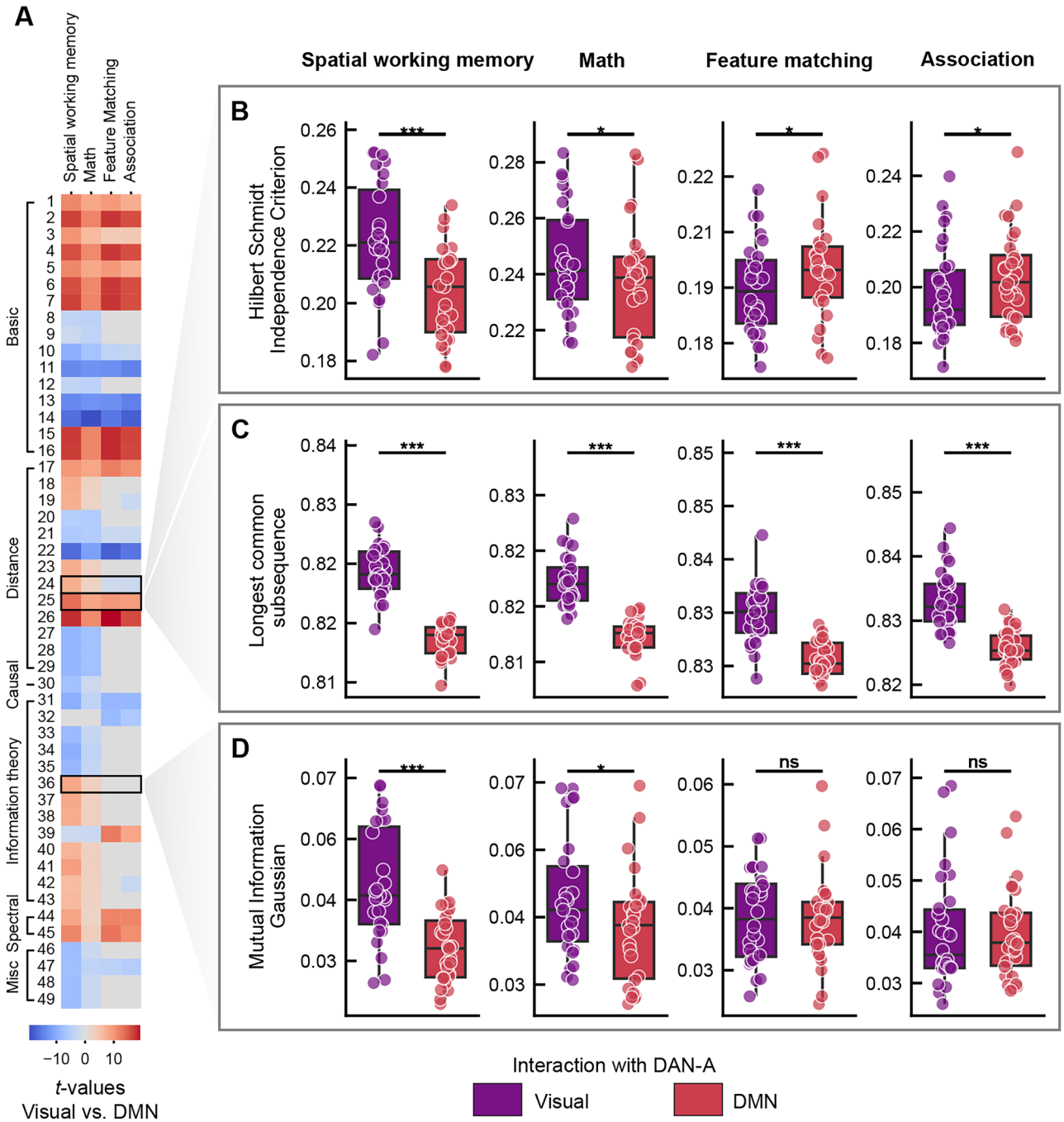

Fig. 7 | Task-dependent interactions of DAN-A with Visual (purple) and DMN (red) networks as measured by 49 SPIs, illustrating varying levels of task sensitivity. A Heatmap showing the task-dependent interactions of DAN-A with Visual and DMN networks across all 49 SPIs. Colours represent t-values comparing interactions between networks, with blue indicating stronger Visual network interactions and red indicating stronger DMN interactions. Each row in panel A is numbered to correspond with the SPI indices in the Supplementary Data, where raw statistical values and classifications of task sensitivity are provided. **B** Hilbert–Schmidt Independence Criterion from the Distance Similarity category, a representative SPI that successfully reveals the double dissociation, showing greater DAN-A-Visual interaction during working memory and greater DAN-A-DMN interaction during long-term memory tasks. **C** Longest common subsequence from the Distance Similarity Category, a representative SPI that lacks task sensitivity, consistently showing greater DAN-A-Visual interaction across tasks. **D** Mutual Information from the Information Theory category, a representative SPI that reveals a single dissociation, showing greater DAN-A-Visual interaction during working memory tasks but failing to reveal the reverse pattern for long-term memory tasks. Significance levels are indicated ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, ns = not significant); $n = 245$ (rest), 27 (spatial working memory), 27 (math), 28 (semantic feature matching), and 30 (semantic association); all represent biologically independent participants. These examples illustrate the variability in task sensitivity across SPIs, underscoring the unique sensitivity of FS to task modulation. Panel C shows an SPI with no task sensitivity, and **D** shows an SPI with weaker task sensitivity (single dissociation). These are included for contrast with **B** to highlight the unique sensitivity of FS to task modulation.

Conclusion

Our results demonstrate that Feature Similarity (FS) represents a methodological advance for understanding brain network interactions. Unlike traditional measures, FS integrates a broad set of interpretable time-series features into similarity matrices derived from local dynamics, allowing it to capture subtle task-dependent modulations while also recovering stable organizational principles. This design provides a generalizable and interpretable tool for reanalysing existing fMRI datasets, resolving

inconsistencies, and refining theories of cognitive flexibility. As neuroimaging methods and experimental designs continue to evolve, FS offers a sensitive framework that can complement these advances and contribute to a deeper understanding of the adaptive nature of brain network interactions.

Methods

This study included three datasets, one publicly available dataset – the Human Connectome Project (HCP)[30] (https://www.humanconnectome.org/), and two task fMRI datasets collected at the University of York, UK[6,29].

We analysed the resting state functional MRI (rsfMRI) data of 245 unrelated participants who completed all four resting state scans from the S900 release of HCP dataset to investigate the functional hierarchy, FC, and FS patterns. In addition, to compare whether FC and FS can capture varying interaction patterns to the same extent, we used two types of tasks (tapping working memory and long-term memory retrieval) which are associated with distinct neurocognitive modes[31]. To tap working memory, participants completed easy and hard spatial working memory and arithmetic tasks designed to localise domain general control regions[32–34], which required information to be maintained and manipulated. They also completed two long-term memory tasks tapping knowledge in long-term memory, a semantic feature matching task that involved linking probe and target concepts (presented as written words) based on colour or shape, and a semantic association task in which participants decided if pairs of words were semantically associated or not[6,29]. These datasets have been previously analysed[6,29].

Participants

All participants were right-handed, native English speakers, had normal or corrected-to-normal vision, and had no history of psychiatric or neurological illness. For the HCP dataset, informed consent was obtained, and the study was approved by the Institutional Review Board of Washington University at St. Louis. For the York working memory and long-term memory dataset, the research was approved by the York Neuroimaging Centre and Department of Psychology ethics committees. All ethical regulations relevant to human research participants were followed.

We analysed the data of 245 neurologically healthy volunteers (130 males, 115 females), aged 23–35 years (mean = 28.21, SD = 3.67), from the HCP dataset[30].

31 neurologically healthy adults (26 females; age: mean ± SD = 20.60 ± 1.68, range: 18–25 years) performed spatial working memory and math tasks in York. One participant with incomplete data (only one of two sessions) was removed. A functional run was excluded if (i) mean relative root mean square (RMS) framewise displacement was higher than 0.2 mm, (ii) it had more than 15% percentage of total frames with motion exceeding 0.25 mm, or (iii) a participant's accuracy on the respective task was three standard deviations below the group mean. If only run of one task was left for a given participant after exclusion, the data from this task were removed for this participant. These exclusion criteria resulted in a final sample of 27 participants for both the spatial working memory task and the math task.

We also analysed a long-term memory task dataset collected at the University of York, UK. 31 healthy adults were recruited from the University of York (25 females; age: mean ± SD = 21.26 ± 2.93, range: 19–34 years). The same exclusion criteria for functional runs were applied as above. For the feature matching task, this left 23 participants with 4 runs, 4 participants with 3 runs, and 1 participant with 2 runs. For the association task, there were 24 participants with 4 runs, 3 participants with 3 runs, and 3 participants with 2 runs. To select experimental materials, we recruited 30 native English speakers who did not participate in the main fMRI experiment as subjects (21 females; age range: 18–24 years). These individuals rated the colour and shape similarity as well as the semantic association strength for each word pair.

Tasks paradigms

Spatial working memory task. Participants were required to maintain four or eight sequentially presented locations in a 3 × 4 grid[33], giving rise to easy and hard spatial working memory conditions (see Fig. 2). Stimuli were presented at the centre of the screen across four steps. Each of these steps lasted for 1 s and highlighted one location on the grid in the easy condition, and two locations in the hard condition. This was followed by a decision phase, which showed two grids side by side (i.e., two-alternative forced choice (2AFC) paradigm). One grid contained the locations shown on the previous four steps, while the other contained one or two locations in the wrong place. Participants indicated their response via a button press and feedback was immediately provided within in 2.5 s. Each run consisted of 12 experimental blocks (6 blocks per condition and 4 trials in a 32 s block) and 4 fixation blocks (each 16 s long), resulting in a total time of 448 s (149 TRs).

Math task. Participants were presented with an addition expression on the screen for 1.45 s and, subsequently made a 2AFC decision indicating their solution within 1 s (see Fig. 2). The easy condition used single-digit numbers while the hard condition used two-digit numbers. Each trial ended with a blank screen lasting for 0.1 s. Each run consisted of 12 experimental blocks (with 4 trials per block) and 4 fixation blocks, resulting in a total time of 316 s (105 TRs).

Semantic feature matching task. Participants were required to make a yes/no decision matching probe and target concepts (presented as words) according to a particular semantic feature (colour or shape), specified at the top of the screen during each trial (see Fig. 2). The feature prompt, probe word, and target words were presented simultaneously. Half of the trials were matching trials in which participants would be expected to identify shared features; while half of the trials were non-matching trials in which participants would not be expected to identify shared features. For example, in a colour matching trial participants would answer 'yes' to the word-pair DALMATIANS—COWS, due to their colour similarity, whereas they would answer 'no' to COAL -TOOTH as they do not share a similar colour.

This task included four runs and two conditions (two features: colour and shape), presented in a mixed design. Each run consisted of four experimental blocks (two 150 s blocks per feature), resulting in a total time of 612 s (408 TRs). In each block, 20 trials were presented in a rapid event-related design. In order to maximize the statistical power of the rapid event-related fMRI data analysis, the stimuli were presented with a temporal jitter randomized from trial to trial[35]. The inter-trial interval varied from 3 to 5 s. Each trial started with a fixation, followed by the feature, probe word, and target word presented centrally on the screen, triggering the onset of the decision-making period. The feature, probe word, and target word remained visible until the participant responded, or for a maximum of 3 s. The condition order was counterbalanced across runs and run order was counterbalanced across participants. Half of the participants pressed a button with their right index finger to indicate a matching trial and responded with their right middle finger to indicate a non-matching trial. Half of the participants pressed the opposite buttons.

Semantic association task. Participants were asked to decide if pairs of words were semantically associated or not (i.e., yes/no decision as above) based on their own experience (see Fig. 2). Overall, there were roughly equal numbers of 'related' and 'unrelated' responses across participants. The same stimuli were used in the semantic feature matching task and semantic association task. For example, DALMATIANS and COWS are semantically related; COAL and TOOTH are not. The feature and association tasks were often separated by one week. This task included four runs, presented in a rapid event-related design. Each run consisted of 80 trials, with about half being related and half being unrelated trials. The procedure was the same as the feature matching task except only two words were presented on the screen.

Image acquisition

Image acquisition for HCP dataset. MRI acquisition protocols of the HCP dataset have been previously described[30,36]. Images were acquired

using a customized 3 T Siemens Connectome scanner having a 100 mT/m SC72 gradient set and using a standard Siemens 32-channel radio-frequency receive head coil. Participants underwent the following scans: structural (at least one T1-weighted (T1w) MPRAGE and one 3D T2-weighted (T2w) SPACE scan at 0.7-mm isotropic resolution), rsfMRI (4 runs ×14 min and 33 s), and task fMRI (7 tasks, 46.6 min in total). Since not all participants completed all scans, we only included 339 unrelated participants from the S900 release. Whole-brain rsfMRI and task fMRI data were acquired using identical multi-band echo planar imaging (EPI) sequence parameters of 2-mm isotropic resolution with a TR = 720 ms. The dMRI data consisted of one 1.25 mm isotropic scan for each participant with 3 shell HARDI type acquisition, including b = 1000, 200, 3000 s/mm^2, total for 270 non-collinear directions. Spin echo phase reversed images were acquired during the fMRI scanning sessions to enable accurate cross-modal registrations of the T2w and fMRI images to the T1w image in each subject and standard dual gradient echo field maps were acquired to correct T1w and T2w images for readout distortion. Additionally, the spin echo field maps acquired during the fMRI session (with matched geometry and echo spacing to the gradient echo fMRI data) were used to compute a more accurate fMRI bias field correction and to segment regions of gradient echo signal loss.

Subjects were considered for data exclusion based on the mean and mean absolute deviation of the relative root-mean-square motion across either four rsfMRI scans, resulting in four summary motion measures. If a subject exceeded 1.5 times the interquartile range (in the adverse direction) of the measurement distribution in two or more of these measures, the subject was excluded. In addition, functional runs were flagged for exclusion if more than 25% of frames exceeded 0.2 mm frame-wise displacement (FD_power). These above exclusion criteria were established before performing the analysis, (for similar implementation, see (for similar implementation, see Faskowitz et al., 2020; Sporns et al., 2021)). The data of 91 participants was excluded because of excessive head motion and the data of another 3 participants was excluded because their resting data did not have all the time points. In total, the data of 245 participants was analysed after exclusions.

Image acquisition for York working memory dataset. MRI acquisition protocols have been described previously[17,37]. Structural and functional data were collected on a Siemens Prisma 3 T MRI scanner at the York Neuroimaging Centre. The scanning protocols included a T1-weighted MPRAGE sequence with whole-brain coverage. The structural scan used: acquisition matrix of 176 × 256 × 256 and voxel size 1 × 1 × 1 mm^3, repetition time (TR) = 2300 ms, and echo time (TE) = 2.26 ms. Functional data were acquired using an EPI sequence with an 800 flip angle and using GRAPPA with an acceleration factor of 2 in 3 ×3 x 4 mm voxels in 64-axial slices. The functional scan used: 55 3-mm-thick slices acquired in an interleaved order (with 33% distance factor), TR = 3000 ms, TE = 15 ms, FoV = 192 mm.

Image acquisition for York long-term memory dataset. MRI acquisition protocols have been described previously[17,37]. Whole brain structural and functional MRI data were acquired using a 3 T Siemens MRI scanner utilising a 64-channel head coil, tuned to 123 MHz at York Neuroimaging Centre, University of York. The functional runs were acquired using a multi-band multi-echo (MBME) EPI sequence, each 11.45 min long (TR = 1.5 s; TEs = 12, 24.83, 37.66 ms; 48 interleaved slices per volume with slice thickness of 3 mm (no slice gap); FoV = 24 cm (resolution matrix = 3 × 3 × 3; 80 × 80); 75° flip angle; 455 volumes per run; 7/8 partial Fourier encoding and GRAPPA (acceleration factor = 3, 36 ref. lines); multi-band acceleration factor = 2). Structural T1-weighted images were acquired using an MPRAGE sequence (TR = 2.3 s, TE = 2.3 s; voxel size = 1 × 1 × 1 isotropic; 176 slices; flip angle = 8°; FoV= 256 mm; interleaved slice ordering). We also collected a high-resolution T2-weighted (T2w) scan using an echo-planar imaging sequence (TR = 3.2 s, TE = 56 ms, flip angle = 120°; 176 slices, voxel size = 1 × 1 × 1 isotropic; Fov = 256 mm).

Image pre-processing

Image pre-processing of HCP dataset. We used HCP's minimal preprocessing pipelines[30]. Briefly, for each subject, structural images (T1w and T2w) were corrected for spatial distortions. FreeSurfer v5.3 was used for accurate extraction of cortical surfaces and segmentation of subcortical structures[38,39]. To align subcortical structures across subjects, structural images were registered using non-linear volume registration to the Montreal Neurological Institute (MNI152) space. Functional images (rest and task) were corrected for spatial distortions, head motion, and mapped from volume to surface space using ribbon-constrained volume to surface mapping.

Subcortical data were also projected to the set of extracted subcortical structure voxels and combined with the surface data to form the standard CIFTI grayordinate space. Data were smoothed by a 2-mm FWHM kernel in the grayordinates space that avoids mixing data across gyral banks for surface data and avoids mixing areal borders for subcortical data. Rest and task fMRI data were additionally identically cleaned for spatially specific noise using spatial ICA + FIX[40] and global structured noise using temporal ICA[41]. For accurate cross-subject registration of cortical surfaces, a multimodal surface matching (MSM) algorithm[42] was used to optimize the alignment of cortical areas based on features from different modalities. MSMSulc ("sulc": cortical folds average convexity) was used to initialize MSMAll, which then utilized myelin, resting state network, and rfMRI visuotopic maps. Myelin maps were computed using the ratio of T1w/T2w images[40]. The HCP's minimally preprocessed data include cortical thickness maps (generated based on the standardized FreeSurfer pipeline with combined T1-/T2-reconstruction). For this study, the standard-resolution cortical thickness maps (32k mesh) were used.

Image pre-processing of York working memory and long-term memory dataset. The York datasets were preprocessed using fMRIPrep 20.2.1 [ref. 43, RRID:SCR_016216], with detailed methods previously described in refs. 6,29. In brief, anatomical preprocessing involved intensity non-uniformity correction, skull stripping, segmentation, and surface reconstruction. Spatial normalization to MNI152 templates was performed through nonlinear registration. For each BOLD run, functional data preprocessing followed standard fMRIPrep procedures. This included generating a reference volume, applying fieldmap correction, and co-registering the BOLD reference to anatomical space. Motion correction and slice-time correction were performed. Resampling was done in the surface space - fsaverage, and final data were output in CIFTI grayordinate space. Post-processing of fMRIPrep outputs was performed using eXtensible Connectivity Pipeline (XCP)[44]. Volumes with framewise displacement greater than 0.3 mm were excluded before nuisance regression. A total of 36 nuisance regressors, including motion parameters, global signal, white matter, and CSF signals, were regressed. Residual time-series were band-pass filtered (0.01–0.08 Hz) and smoothed with a 6.0 mm FWHM Gaussian kernel. Detailed information is provided in Supplementary Materials and Methods, Section 1.1.

The York University dataset was preprocessed using the default parameters of fMRIPrep, while the HCP dataset was analysed in the minimally preprocessed form provided by the consortium. Although different pipelines were applied, we relied on the standard procedures associated with each dataset, and no direct comparisons were made between York and HCP data.

Task fMRI analysis

Individual-specific parcellation. To account for anatomical and functional variability, we used a multi-session hierarchical Bayesian model to estimate individual-specific parcellations, following the methods in refs. [6,24]. This approach defines 400 individualized parcels across 17 networks per participant and the individual-specific parcellation showed greater homogeneity than the parcellation using group atlas[6,24]. Although the specific vertices for each parcel differ across participants, the MS-HBM framework preserves cross-subject correspondence, ensuring

comparability of parcels across individuals. For each parcel in each participant, we extracted the parcel time series by averaging the signals across all vertices belonging to that parcel. Group-level FC and FS matrices were obtained by first calculating parcel-wise matrices for each participant in each task and then averaging them across subjects. In this framework, network labels (e.g., DAN-A, DAN-B) are propagated consistently across individuals by the MS-HBM model, and no post hoc relabelling was performed. Further details are provided in the Supplementary Materials and Methods (Section 1.2.1).

Constructing fMRI FC matrices. To investigate how FC varies across tasks, we computed task-based and resting state FC matrices. We opted not to use the traditional psychophysiological interaction method for measuring task-state functional connectivity due to its potential to inflate activation-induced task-state functional connectivity, which may identify regions that are active rather than interacting during the task[45]. Since task activations can spuriously inflate task-based functional connectivity estimates, it is necessary to correct for task-timing confounds by removing the first-order effect of task-evoked activations (i.e., mean evoked task-related activity, likely active during the task) prior to estimating task-state functional connectivity (likely interacting during the task)[45]. Specifically, we fitted the task timing for each task using a finite impulse response model, a method that has been shown to reduce both false positives and false negatives in FC estimation[46–48]. In long-term memory tasks, approximately 5 timepoints were modelled for each trial. For the working memory task, roughly 2.5 timepoints were modelled for each trial.

Following task regression, we demeaned the residual time series for each parcel and quantified the FC using Pearson correlation for each participant, task and run. The Pearson correlation coefficients might be inflated due to the temporal autocorrelation in task fMRI time series data[49]. To address this, we corrected the Pearson correlation using a correction approach, xDF, which accounts for both autocorrelation within each time series as well as instantaneous and lagged cross-correlations between the time series[50]. This method provides an effective degrees of freedom estimator that addresses cross-correlations, thereby preventing inflation of Pearson correlation coefficients. Our goal was not to remove temporal autocorrelation, but to enhance the precision and reliability of correlation estimates, reducing false functional connectivity between regions. We calculated xDF-adjusted z-scored correlation coefficients to assess the inter-regional relationships in BOLD time series, resulting in a 400 × 400 functional connectivity matrix for each participant, task, and run. Finally, we averaged these functional connectivity estimates within networks, and between pairs of networks, to construct a network-by-network functional connectivity matrix. The same method was used to calculate the resting-state functional connectivity of the HCP dataset and construct a corresponding network-by-network functional connectivity matrix, except without the task regression step. In addition to analyses using task-residual time series after FIR regression, we also computed FC on parcel time series without removing task-evoked responses. Both approaches yielded highly consistent patterns, confirming that preprocessing differences do not account for the observed effects.

Feature extraction of the time-series data. To investigate how the FS patterns vary across tasks, we calculated task-based and resting state FS using the extracted features. To extract the features of the time-series required for this analysis, we used the time-series analysis toolbox (hctsa)[22,23]. With this tool, we transformed each time-series in the dataset into a set of over 7,700 features, which include, but are not limited to, distributional properties, entropy and variability, autocorrelation, time-delay embeddings, and nonlinear properties of a given time-series (Fig. 1)[22,25]. We extracted features from the parcellated fMRI time-series of each participant, each task, and each run separately. After the feature extraction procedure, we removed the outputs of the operations that produced errors and normalized the remaining features (about 6900 features) across parcels using an outlier-robust sigmoidal transform. The resulting normalized feature matrix (400 parcels × ~7000 features × 4 runs) was used to predict networks labels of parcels to identify the DAN subnetwork with varying interaction patterns across tasks (Method 4.5.4) and to construct FS matrix for further analysis (Method 4.5.4).

Constructing fMRI FS matrices. To investigate how FS patterns vary across tasks, we calculated task-based and resting state FS matrices. Specifically, we calculated Pearson correlation coefficients of the extracted features, which represented the pairwise FS between all possible combinations of brain parcels (Fig. 1). This resulted in a 400 by 400 FS matrix for each run each task for each participant. The resulting correlation values were Fisher-z transformed. Finally, to construct a network-by-network FS matrix, we averaged the estimates of FS within networks and between pairs of networks for further analysis. Unlike FC, which required xDF correction to adjust for temporal autocorrelation and cross-correlations, FS did not require xDF because it is based on static feature profiles, with temporal autocorrelation deliberately retained to preserve biologically meaningful properties.

To investigate whether functionally connected regions display similar features, we calculated Pearson correlations coefficients for FS within and between networks defined by resting state FC. Then we tested whether the correlations for FS within networks were greater than the correlations between networks by conducting paired *t*-test.

Additionally, we estimated the similarity between FC and FS by calculating their correlations at rest and for each task. To control for multiple comparison, we FWE-corrected the p values using permutation-based maximum r values. Specifically, we created a null distribution using permutation for each task and chose the maximum values among all the tasks. We then compared the observed correlation value with the null distribution to examine whether the correlation between FC and FS was significantly greater than that expected from null distribution.

Dimension reduction analysis of FC and FS. To examine whether FC and FS captured similar principal components, we performed dimension reduction analysis on both resting state FC matrix and FS matrix derived from the HCP dataset. For the FC matrix, we first calculated resting state FC for each run of each participant, as detailed in Method section "Constructing fMRI FC matrices". Subsequently, these individual connectivity matrices were then averaged to calculate a group-level connectivity matrix. We extracted ten group-level gradients from the group-level connectivity matrix (dimension reduction technique = diffusion embedding, kernel = None, sparsity = 0.9), in line with previous studies[37,51] using the Brainspace Toolbox[52]. This analysis resulted in ten group-level gradients explaining maximal whole-brain connectivity variance in descending order. A parallel analysis was performed for the FS matrix, employing identical procedures except for the input, which consisted of the FS matrix as delineated in Method Section "Constructing fMRI FS matrices". This analysis resulted in ten group-level gradients explaining maximal whole-brain FS variance in descending order. We retained the components explaining the most variance by looking at the eigenvalues of each component in the scree plots shown in Fig. 3.

Finally, we examined the similarities between the first three components captured by FC and FS, respectively via calculating the Pearson correlations between corresponding components. Due to the spatial autocorrelation present in each principal component, we created a null distribution using spin permutation implemented in BrainSMASH[53]. This approach simulates brain maps, constrained by empirical data, that preserve the spatial autocorrelation of cortical parcellated brain maps. We then compared the observed correlation value with the null distribution for left hemisphere to examine whether the correlation between two components of parcels in the left hemisphere was significantly greater than that expected from spatial autocorrelation alone. Similarly, we examined the correlation between two components of parcels in the right hemisphere using the same methods. This analysis was performed for the two hemispheres separately

because the geodesic distance between parcels was used to generate the spatial-autocorrelation-preserving surrogate maps when creating the null distribution. It was only possible to measure within-hemisphere geodesic distance between parcels because the left and right hemisphere surface maps were not on the same mesh.

Timescale analysis. The first component captured by FS matrix was similar to the timescale gradient map reported before[7], as determined by visual comparison. To further understand this component, we calculated its correlation with the intrinsic timescale gradient map, which represents the temporal duration of ongoing inputs that the brain can process. The intrinsic timescale for each parcel was characterized by the decay of the temporal autocorrelation function, as the time taken for the autocorrelation function to reach a threshold of $r = 0.5$ (i.e., half of the full width at half maximum), consistent with prior studies[7,54]. A higher value of the intrinsic timescale indicates longer ongoing inputs that the parcel can process. Given the spatial autocorrelation present in this component and the timescale gradient map, we created a null distribution using spin permutation implemented in BrainSMASH[53]. We subsequently compared the observed correlation value with the null distribution for each hemisphere to determine whether the real correlations were significantly greater than that expected by spatial autocorrelation alone.

Calculating the correlations across tasks in FC and in FS. To test whether FS was more sensitive to task modulation than FC, we calculated correlations across tasks for both measures and compared their sensitivities by assessing if FS showed weaker correlations than FC. We used two methods: (i) we first calculated Pearson correlations (r1) for FC across tasks for each possible pair of tasks at the group level (for instance, between the spatial working memory and math tasks) using the task mean FC matrices. Then, we repeated the process for FS (r2), using the task mean FS matrices. Finally, we then compared the resulting correlation values (r1 versus r2). (ii) We calculated the Pearson correlation coefficients between the FC matrix of the spatial working memory task and the FC matrix of math task at the individual level for each participant given that each participant completed both spatial working memory task and math task. Then we repeated the process for FS, using the FS matrix of each participant. Finally, we converted the r values to z values using Fisher transformation and compared the z values of the FC with the ones of the FS by conducting paired t-tests. The same procedures were applied to the two long-term memory tasks (semantic feature matching versus association).

Classification analysis—decoding task labels using FS matrix. To determine whether feature similarity captures task information, we conducted classification analyses to predict task labels (spatial working memory versus math) using FS matrices of the working memory tasks of all the participants, respectively. We employed scikit-learn's[55] linear support vector machine classifier (SVC) with 5-fold cross-validation to avoid overfitting and ensure reliable performance. To assess statistical significance, we performed permutation tests by shuffling task labels 1000 times, creating a null distribution for comparison[56]. The same approach was used for long-term memory tasks (semantic feature matching versus semantic association) using feature similarity matrices of long-term memory tasks.

Comparing FC difference and FS difference between networks across tasks. To investigate whether FS captured varying interaction patterns across tasks that could not be captured by FC, we first calculated the average FC between DAN-A and Visual network and between DAN-A and DMN, across all runs per participant per task. Subsequently, we calculated the relative FC difference by subtracting the FC between DAN-A and DMN from the FC between DAN-A and Visual network. Finally, we conducted paired-*t* tests for each task to examine the significance of the FC difference between network pairs. We further investigated whether the task influenced the FC difference between network pairs using the maximum/minimum permutation test. We calculated the mean FC difference between DAN-A and Visual network versus DAN-A and DMN for each task and calculated the mean FC difference between each task pair. To access statistical significance, we permutated the task label 10000 times and then calculated the mean FC difference between these two tasks to build a null distribution for each task pair. To control the family-wise error (FWE) rate ($p = 0.05$, FWE-corrected) given the inclusion of multiple task pairs, we utilized the permutation-based maximum mean FC difference and minimum mean FC difference values in the null distribution for each task pair. To evaluate significance, if the observed mean difference value was positive, we counted the percentage of times that mean difference values in the maximum null distribution were greater than the observed 'true' mean difference values. Conversely, if the observed mean difference value was negative, we counted the percentage of times of mean difference values in the minimum null distribution were less than the observed 'true' mean difference values.

We then examined whether task influences the FS differences between targeted network pairs (DAN-A-Visual versus DAN-A-DMN) using the same procedures but applied to FS. Our results showed that FS is more sensitive to task modulation than FC in the DAN-A-Visual and DAN-A-DMN comparison. To assess whether this holds for other networks, we repeated the analysis, comparing FPCN-A-Visual versus FPCN-A-DMN.

Classification analysis—decoding network labels of parcels. Given heterogeneity of DAN, we used a data-driven approach to identify the subnetwork most likely to vary in interaction patterns during tasks using the extracted features. Specifically, we conducted a classification analysis on a normalized feature matrix (400 parcels by about 7000 features by 4 runs) to decode parcel network labels. After confirming accurate classification, we examined the confusion matrix to identify classification errors (Detailed information is provided in Supplementary Materials and Methods, Section 1.2.3). This allowed us to explore network similarity, as functionally similar networks are more likely to be misclassified.

Constructing fMRI statistics of pairwise interaction matrices. Many techniques have been developed to measure pairwise interactions in complex systems, such as Pearson correlation in fMRI and mutual information in signal processing. These methods, which range from contemporaneous correlation coefficients to causal inference approaches, are based on distinct quantitative theories and define interactions differently. As a result, each method captures different information and has varying sensitivity. However, since most methods focus on a single aspect, their sensitivity may still be weaker than FS, which considers multiple dimensions of information. To test this, we calculated many SPIs using the Python Toolkit for Statistics for Pairwise Interactions (pyspi; v0.4.1[1];. The parcel-based task fMRI time series were z-scored along the time dimension before the calculation.

Different systems involve distinct interactions, causing SPIs to perform variably across datasets (e.g., smartwatch activity, EEG, fMRI), with some methods excelling only in specific datasets[1]. Therefore, it's essential to select SPIs that capture the relevant interactions for a given dataset. From the original list of SPIs[1], we initially selected all 67 SPIs that significantly classified states in the fMRI film dataset of comparable duration (770 s versus. 612 s for our semantic tasks) and then refined this set to 49 SPIs with reasonable computational requirements (<5 h per run per task per subject). This approach ensured the selected SPIs were both relevant and practical for our analysis.

The 49 SPIs used in this study were derived from 20 interaction measures across six categories: (1) Basic Statistics: This category includes measures such as covariance, which quantifies the linear relationship between two variables, and precision, which captures pairwise associations while controlling for the effects of other variables in the dataset.

(2) Distance-Based Similarity: These measures quantify statistical similarity or independence using pairwise distances between bivariate observations. Examples include Euclidean distance and correlation distance, which evaluate how similar or dissimilar two variables are in their distributions. (3) Information-Theoretic Measures: Metrics in this category include mutual information, which quantifies the total dependency between two variables, and joint entropy, which measures the combined uncertainty or information content of two variables. (4) Causal Inference: Methods in this category are designed to infer directed relationships. For example, regression error-based causal inference determines the causal direction by analysing which variable's residuals are more independent of the predictor, indicating potential causal influence. (5) Spectral Measures: These measures are computed in the frequency or time-frequency domain. Examples include spectral coherence magnitude, which quantifies the alignment of frequency components in phase and amplitude, and phase locking value, which measures the consistency of phase differences between two signals over time. (6) Miscellaneous Methods: This category includes techniques such as linear model fits, which estimate relationships through regression, and other statistical tools that quantify pairwise interactions but do not fit into the above categories. See Supplementary Material 1.3.3 for a full list of SPIs used in this study. The detailed information of these SPIs can be found in Supplementary Table 1.

We calculated the pairwise statistics for the selected SPIs across all subjects and compared the interaction matrices across tasks, focusing on key networks (DAN-A, Visual, and DMN). We conducted paired t-tests for each task to assess differences between network pairs and conducted FDR correction to control for multiple comparisons. The same methods were used when comparing FC and FS, with the only difference being the input matrices derived from SPIs.

Statistics and reproducibility

We computed FC between and within brain networks using Pearson correlation. We derived FS from multidimensional feature-based correlations across regional time-series features. We applied diffusion embedding to FC and FS matrices to extract the principal connectivity components. We used paired *t*-tests to examine task-related differences between FC and FS and conducted classification analyses to identify subnetworks within the DAN. We applied family-wise error (FWE) correction and spin permutation tests to control for multiple comparisons. We also used spin permutation to account for spatial autocorrelation in cortical data, ensuring that anatomical proximity did not bias statistical significance.

All statistical analyses used biologically independent participants, with sample sizes reported in the figure legends. We replicated previous findings showing that regions within the same intrinsic functional network exhibit higher FC and FS than regions across networks and that the first three principal connectivity components of FC replicate those described by Margulies et al. We further showed that FC and FS share two principal components and differ in one. Extending these results, we demonstrated that FS exhibits greater task-dependent variation than FC and captures network interaction patterns across tasks that FC and most other SPIs fail to detect.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The HCP dataset is publicly available at https://www.humanconnectome.org/. Raw data collected at the University of York cannot be shared at this time due to consent limitations under UK GDPR. Researchers interested in accessing this data should contact the Chair of the Research Ethics Committee at the York Neuroimaging Centre. Preprocessed data from the University of York are available on the Open Science Framework[57], and results are accessible at https://github.com/Xiuyi-Wang/Feature_similarity_Project.

Code availability

The analysis code is accessible at https://github.com/Xiuyi-Wang/Feature_similarity_Project.

Received: 9 July 2025; Accepted: 30 October 2025;
Published online: 16 December 2025

References

1. Cliff, O. M., Bryant, A. G., Lizier, J. T., Tsuchiya, N. & Fulcher, B. D. Unifying pairwise interactions in complex dynamics. *Nat. Comput Sci.* **3**, 883–893 (2023).
2. Simony, E. et al. Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat. Commun.* **7**, 12141 (2016).
3. O'Reilly, J. X., Woolrich, M. W., Behrens, T. E. J., Smith, S. M. & Johansen-Berg, H. Tools of the trade: psychophysiological interactions and functional connectivity. *Soc. Cogn. Affect Neurosci.* **7**, 604–609 (2012).
4. Friston, K. Dynamic causal modeling and Granger causality Comments on: the identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. *Neuroimage* **58**, 303–305 (2011).
5. Liu, Z.-Q. et al. Benchmarking methods for mapping functional connectivity in the brain. *Nat Methods*. **22**, 1593–1602 (2025).
6. Wang, X. et al. The brain's topographical organization shapes dynamic interaction patterns that support flexible behaviour based on rules and long term knowledge. *J. Neurosci.* **44**, 22 (2024).
7. Raut, R. V., Snyder, A. Z. & Raichle, M. E. Hierarchical dynamics as a macroscopic organizing principle of the human brain. *Proc. Natl. Acad. Sci. USA* **117**, 20890–20897 (2020).
8. Hansen, J. Y. et al. Integrating multimodal and multiscale connectivity blueprints of the human cerebral cortex in health and disease. *PLoS Biol.* **21**, e3002314 (2023).
9. Zhu, X. et al. Fusing functional connectivity with network nodal information for sparse network pattern learning of functional brain networks. *Inf. Fusion* **75**, 131–139 (2021).
10. Mohanty, R., Sethares, W. A., Nair, V. A. & Prabhakaran, V. Rethinking measures of functional connectivity via feature extraction. *Sci. Rep.* **10**, 1–17 (2020).
11. Prado, P. et al. Harmonized multi-metric and multi-centric assessment of EEG source space connectivity for dementia characterization. *Alzheimer's Dement. Diagn. Assess. Dis. Monit.* **15**, e12455 (2023).
12. Shafiei, G. et al. Topographic gradients of intrinsic dynamics across neocortex. *Elife* **9**, 1–24 (2020).
13. Blank, I. A. & Fedorenko, E. No evidence for differences among language regions in their temporal receptive windows. *Neuroimage* **219**, https://doi.org/10.1016/j.neuroimage.2020.116925 (2020).
14. Honey, C. J. et al. Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* **76**, 423–434 (2012).
15. Ito, T., Hearne, L. J. & Cole, M. W. A cortical hierarchy of localized and distributed processes revealed via dissociation of task activations, connectivity changes, and intrinsic timescales. *Neuroimage* **221**, 117141 (2020).
16. González-García, C., Flounders, M. W., Chang, R., Baria, A. T. & He, B. J. Content-specific activity in frontoparietal and default-mode networks during prior-guided visual perception. *Elife* **7**, https://doi.org/10.7554/eLife.36068 (2018).
17. Wang, X., Gao, Z. & Smallwood, J. Both default and multiple-demand regions represent semantic goal information. *J. Neurosci.* **41**, 3679–3691 (2021).
18. Wolff, A. et al. Intrinsic neural timescales: temporal integration and segregation. *Trends Cogn. Sci.* **26**, 159–173 (2022).
19. Golesorkhi, M. et al. The brain and its time: intrinsic neural timescales are key for input processing. *Commun. Biol*. **4**, 1 (2021).

20. Buckner, R. L. & DiNicola, L. M. The brain's default network: updated anatomy, physiology and evolving insights. *Nat. Rev. Neurosci.* **20**, 593–608 (2019).
21. Dixon, M. L. et al. Interactions between the default network and dorsal attention network vary across default subsystems, time, and cognitive states. *Neuroimage* **147**, 632–649 (2017).
22. Fulcher, B. D., Little, M. A. & Jones, N. S. Highly comparative time-series analysis: the empirical structure of time series and their methods. *J. R. Soc. Interface* **10**, 20130048 (2013).
23. Fulcher, B. D. & Jones, N. S. hctsa: a computational framework for automated time-series phenotyping using massive feature extraction. *Cell Syst.* **5**, 527–531.e3 (2017).
24. Kong, R. et al. Individual-specific areal-level parcellations improve functional connectivity prediction of behavior. *Cerebral Cortex* **31**, 4477–4500 (2021).
25. Fulcher, B. D. Feature-based time-series analysis. in *Feature Engineering for Machine Learning and Data Analytics* 87–116 (CRC Press, 2018).
26. Margulies, D. S. et al. Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl. Acad. Sci. USA* **113**, 12574–12579 (2016).
27. Wang, X., Margulies, D. S., Smallwood, J. & Jefferies, E. A gradient from long-term memory to novel cognition: transitions through default mode and executive cortex. *Neuroimage* **220**, https://doi.org/10.1016/j.neuroimage.2020.117074 (2020).
28. Wang, X. et al. Physical distance to sensory-motor landmarks predicts language function. *Cerebral Cortex* https://doi.org/10.1093/cercor/bhac344 (2023).
29. Wang, X. et al. Macroscale brain states support the control of semantic cognition. *Commun. Biol.* **7**, https://doi.org/10.1038/s42003-024-06630-7 (2024).
30. Glasser, M. F. et al. The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* **80**, 105–124 (2013).
31. Wang, H. T. et al. Neurocognitive patterns dissociating semantic processing from executive control are linked to more detailed off-task mental time travel. *Sci. Rep.* **10**, https://doi.org/10.1038/s41598-020-67605-2 (2020).
32. Blank, I., Kanwisher, N. & Fedorenko, E. A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *J. Neurophysiol.* **112**, 1105–1118 (2014).
33. Fedorenko, E., Behr, M. K. & Kanwisher, N. Functional specificity for high-level linguistic processing in the human brain. *Proc. Natl. Acad. Sci. USA* **108**, 16428–16433 (2011).
34. Fedorenko, E., Duncan, J. & Kanwisher, N. Broad domain generality in focal regions of frontal and parietal cortex. *Proc. Natl. Acad. Sci. USA* **110**, 16616–16621 (2013).
35. Dale, A. M. Optimal experimental design for event-related fMRI. *Hum. Brain Mapp. 8. 2-* **3**, 109–114 (1999).
36. Barch, D. M. et al. Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage* **80**, 169–189 (2013).
37. Wang, X., Margulies, D. S., Smallwood, J. & Jefferies, E. A gradient from long-term memory to novel cognition: transitions through default mode and executive cortex. *Neuroimage* **220**, 117074 (2020).
38. Fischl, B., Sereno, M. I., and Dale, A. M. Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. *Neuroimage* https://doi.org/10.1006/nimg.1998.0396 (1999).
39. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. *Neuroimage* **9**, 179–194 (1999).
40. Salimi-Khorshidi, G. et al. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* **90**, 449–468 (2014).
41. Glasser, M. F. et al. Using temporal ICA to selectively remove global noise while preserving global signal in functional MRI data. *Neuroimage* **181**, 692–717 (2018).
42. Robinson, E. C. et al. MSM: a new flexible framework for Multimodal Surface Matching. *Neuroimage* **100**, 414–426 (2014).
43. Esteban, O. et al. {fMRIPrep}: a robust preprocessing pipeline for functional {MRI}. *Nat Methods* https://doi.org/10.1038/s41592-018-0235-4 (2018).
44. Ciric, R. et al. Mitigating head motion artifact in functional connectivity MRI. *Nat. Protoc.* **13**, 2801–2826 (2018).
45. Cole, M. W. et al. Task activations produce spurious but systematic inflation of task functional connectivity estimates. *Neuroimage* **189**, 1 (2019).
46. McCormick, E. M., Arnemann, K. L., Ito, T., Hanson, S. J. & Cole, M. W. Latent functional connectivity underlying multiple brain states. *Netwk Neurosci*. **6**, 1–42 (2022).
47. Norman-Haignere, S. V., McCarthy, G., Chun, M. M. & Turk-Browne, N. B. Category-selective background connectivity in ventral visual cortex. *Cereb. Cortex* **22**, 391–402 (2012).
48. Cocuzza, C. V., Ito, T., Schultz, D., Bassett, D. S. & Cole, M. W. Flexible coordinator and switcher hubs for adaptive task control. *J. Neurosci.* **40**, 6949–6968 (2020).
49. Allefeld, C., Graben, P. B., and Kurths, J. Advanced methods of electrophysiological signal analysis and symbol grounding. *Nova Publishers* 276–296 (2008).
50. Afyouni, S., Smith, S. M. & Nichols, T. E. Effective degrees of freedom of the Pearson's correlation coefficient under autocorrelation. *Neuroimage* **199**, 609–625 (2019).
51. Mckeown, B. et al. The relationship between individual variation in macroscale functional gradients and distinct aspects of ongoing thought. *Neuroimage* **220**, 117072 (2020).
52. Vos de Wael, R. et al. BrainSpace: a toolbox for the analysis of macroscale gradients in neuroimaging and connectomics datasets. *Commun. Biol.* **3**, 1–10 (2020).
53. Burt, J. B., Helmer, M., Shinn, M., Anticevic, A. & Murray, J. D. Generative modeling of brain maps with spatial autocorrelation. *Neuroimage* **220**, 117038 (2020).
54. Murray, J. D. et al. A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014).
55. Pedregosa FABIANPEDREGOSA, F Scikit-learn: machine learning in Python Gaël Varoquaux bertrand thirion vincent dubourg alexandre passos PEDREGOSA, VAROQUAUX, GRAMFORT ET AL. Matthieu Perrot. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
56. Ojala, M. & Garriga, G. C. Permutation tests for studying classifier performance. Journal of machine learning research **11**, 6 (2010).
57. Wang, X. et al. Feature similarity, a sensitive method to capture the functional interaction of brain regions and networks to support flexible behavior. [Data set]. *Open Sci. Framework*. https://doi.org/10.17605/OSF.IO/T2P7J (2025).

Acknowledgements

We are grateful to Pradeepa Ruwan and Antonia De Freitas for piloting the experiment. We would like to thank Ben D Fulcher for providing codes and a guide to help us extract the features using HCTSA. X.W. discloses support for the research of this work from National Natural Science Foundation of China (Grant Number. 32300881). Y. D. discloses support for the publication of this work from the STI 2030—Major Projects (Grant Number. 2021ZD0201500), and Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences (Grant Number. E2CX3625CX). E.J. discloses support for the research of this work from a European Research Council Consolidator grant (Project ID: 771863 - FLEXSEM). N.L. discloses support for the research of this work from National Natural Science Foundation of China (Grant Number. 32471111).

Author contributions

X.W. and E.J. designed the research. K.K.R. and X.W. collected the data. X.W. and B.L. analysed the data. X.W. wrote the original manuscript. All authors, including N.S., G.S., N.L., J.S., and D.Y. edited the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-025-09165-7.

Correspondence and requests for materials should be addressed to Xiuyi Wang, Elizabeth Jefferies or Yi Du.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: Jasmine Pan.