

Unveiling the functional specialisation of human circuits with naturalistic stimuli

Marcela Ovando-Tellez

marcela.ovandot@gmail.com

Brain Connectivity and Behaviour Laboratory, Paris, France; Groupe d'imagerie fonctionnelle (GIN), Institut des maladies Neurodegeneratives (IMN) – UMR 5293, CNRS, Bordeaux, France
<https://orcid.org/0000-0002-9521-4620>

Chris Foulon

UCL Queen Square Institute of Neurology, University College London <https://orcid.org/0000-0002-7822-2653>

Victor Nozais

Groupe d'imagerie fonctionnelle (GIN), Institut des maladies Neurodegeneratives (IMN) – UMR 5293, CNRS, Bordeaux, France

Valentina Pacella

University School for Advanced Studies IUSS Pavia <https://orcid.org/0000-0001-6168-9954>

Michel Thiebaut de Schotten

CNRS, University of Bordeaux <https://orcid.org/0000-0002-0329-1814>

Article

Keywords:

Posted Date: July 28th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-7038603/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Abstract

Functional MRI (fMRI) has traditionally focused on grey matter activations, overlooking the contribution of white matter pathways to brain organisation. Yet evidence suggests that white matter constraints, modulates, and integrates signals across distant regions. Here, we introduce a framework to map the functional specialisation of white matter circuits using fMRI scans collected while participants watched naturalistic videos. By leveraging ecologically valid stimuli, this method captures cognition as it naturally unfolds. Integrating fMRI with brain connections, we derived a functionally grounded parcellation of the connectome, challenging conventional cortical-based cognitive taxonomy. Distinct cognitive profiles emerged for association and commissural fibres, affirming the functionally heterogeneous nature of white matter systems. A novel artificial intelligence framework applied to independent datasets confirmed the robustness and functional relevance of the identified parcels. This open-access framework also enables the interactive exploration of white matter functional organisation, offering a new lens on the neural bases of cognition.

Introduction

The intricate relationship between brain and behaviour has led to a progressively refined delineation of the cortex into functional fields over the past three decades. Advances in brain imaging techniques, notably functional magnetic resonance imaging (fMRI), have provided crucial insights into how different brain regions support various cognitive processes and behaviours ¹, as reflected in their different activity patterns and connectivity with the rest of the brain ²⁻⁷. However, task-based fMRI is constrained by predefined cognitive paradigms, in which observed brain activity is tightly linked to specific task demands, limiting the interpretation of function-related activity.

Despite the growing importance of large-scale integrated systems that support the functioning of the brain ⁸, its wiring – the connectome – has been less thoroughly investigated. White matter has typically been divided based on the shape (uncinate, arcuate), the orientation (longitudinal, vertical), and the projections (fronto-occipital) of the tracts ⁹. The functional roles of white matter connections have been inferred indirectly by relating functional impairments to specific white matter regions, associating disconnections of structural pathways caused by lesions with cognitive deficits ¹⁰. However, such approaches are biased by the redundant distribution of lesion locations ^{11,12} and our approximate taxonomy of neuropsychological functions ¹³. Additionally, previous attempts to parcellate white matter based on functional tasks have been hampered by low signal-to-noise ratios existing in the white matter and the absence of a suitable model to fit the data ¹⁴. Hence, a functionally relevant taxonomy of the white matter is still needed.

Recent advances in functional MRI analyses allow us to explore the white matter's contribution to functional connectivity through the Functionnectome method ¹⁵. This method projects fMRI-derived activation from the cortex and subcortical areas onto the white matter using the anatomical priors of the

white matter connectivity. Emerging evidence from projecting resting-state fMRI data on white matter structures using the Functionnectome suggests that this approach can reveal novel links between cortical networks that were traditionally considered independent¹⁶. However, the functional characterisation of resting-state data remains elusive¹⁷, and task-related fMRI is still the gold standard for assessing the relationship between cognitive functions and brain structures. Since it is impossible to assess all cognitive tasks comprehensively, naturalistic stimuli such as movies have emerged as an alternative for studying functional specialisation¹⁸. Passive video-watching engages multiple cognitive and perceptual systems dynamically, enabling functionally meaningful brain activations beyond rigid task designs. It provides a more naturalistic window into the functional brain organisation than conventional task-based fMRI. For example, videos depicting hand movements or requiring face recognition and object discrimination activate respective neural networks for these functions¹⁹. However, how white matter networks could be triggered and segregated by the content of specific video segments remains unknown.

In the present study, we aimed to parcellate white matter circuits according to their functional specificity using a data-driven approach. To this end, we applied the Functionnectome to data from 110 participants acquired with fMRI during video watching and identified functionally specialised white matter circuits independent of preexisting cognitive taxonomies. We then validated these parcellations in three independent external datasets, demonstrating their generalizability across samples and stimuli. To further characterise the functional roles of these parcellations, we developed a novel artificial intelligence (AI) framework grounded in a previously established cognitive morphospace¹³, which we extended to white matter representations. Doing so, we associated dynamic patterns of white matter activity with cognitive terms over time. This Neuro-AI approach leveraged external neuroimaging datasets, including fMRI data acquired during movie watching (Naturalistic Neuroimaging Database - NNDb)²⁰ and naturalistic scene viewing (Natural Scenes Dataset - NSD)²¹.

The Cognitive Parcellation of Association and Commissural Tracts (Cognipact) provides a paradigm shift in functional investigation, offering a novel framework to explore white matter contributions to brain function. We provide an online tool (<http://cognipact.bcblab.com>) accompanying this study that offers a user-friendly platform for researchers to investigate further the relationship between white matter connectivity and cognitive function. Additionally, the parcellations are available for download, visualisation and reuse in NeuroVault (<https://neurovault.org/collections/19566/>).

Results

Robust and reproducible functional topography of the white matter circuits

Each individual's fMRI signals during video watching were projected onto the white matter using the Functionnectome method. This step was performed separately for associative, commissural, and projection fibres on both the discovery and replication datasets. Group-level variation profiles of voxel activity were captured via one-sample t-test applied to the Functionnectome time series at each time

point. These temporal variations were embedded into a two-dimensional space using Uniform Manifold Approximation and Projection algorithm (UMAP) ²², revealing the intrinsic functional organisation of voxels based on their activation patterns. The distribution of these voxels was similar for the discovery and replication groups (Fig. 1), for the associative ($r = .98$), projection ($r = .97$), and commissural ($r = .72$) fibres, supporting the robustness and reproducibility of the Functionnectome analysis. We compared these distributions to a null model, represented by a randomly generated UMAP, and observed a dissimilar distribution between the discovery and replication UMAPs and the null model ($r < .10$).

Given the high similarity between the discovery and replication groups, we merged the two into the brain parcellation group to increase our sample size (Fig. 2a-b). We selected the best Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) ^{23,24} parameters to identify the clusters grouping the voxels with similar activation patterns. For the association fibres, the best parameters were a minimum cluster size of 400 and a minimum sample of 250, reaching a coverage of 67% of the brain voxels and a Density-Based Clustering Validation (DBCV) ²⁵ score of 0.52 (range from - 1 to 1). The HDBSCAN model using these parameters resulted in 36 clusters (Fig. 2c). For the commissural fibres, the best model used a minimum cluster size of 50 and a minimum sample of 400, reaching a coverage of 90% of the brain voxels and a DBCV score of 0.50. This HDBSCAN model resulted in 40 clusters (Fig. 2d). The voxels that the model did not cover were classified as noise. We did not identify any model that converged to cluster the projection fibres based on their functional activations. The clusters identified for the association and commissural fibres were mapped back onto the brain, leading to a white matter parcellation for association (Fig. 2e) and commissural (Fig. 2f) fibres based on functional activations. This parcellation was visualised in relation to key white matter tracts, providing anatomical context for the functionally derived clusters (Fig. 2g-h and **Supplementary Fig. 1**).

Out-of-sample validation of functional specificity

To evaluate the generalisability and functional relevance of our association and commissural parcellations, we first assessed whether the temporal activation patterns observed in the brain parcellation group are observed in an independent out-of-sample group (i.e., the validation group), exposed to the same naturalistic movie-watching paradigm. We projected the parcellations onto the validation group and extracted the group-level temporal variation series from each parcel. We then compared the resulting activation profiles to those from the brain parcellation group by computing correlation matrices across parcels. We observed significant correlations between the brain activations in the validation group and those in the brain parcellation group (Fig. 3a-d and **Supplementary Fig. 4**; all $p < .001$). For the association parcellation, the mean correlation was 0.78 (SD = 0.099; range = 0.54–0.92). For the commissural parcellation, the mean correlation was 0.78 (SD = 0.13; range = 0.48–0.94). This suggests that the identified parcels consistently reflect functionally meaningful activation patterns driven by the naturalistic video stimuli. This result supports the reliability and reproducibility of our white matter parcellation based on functional activations across independent datasets, reinforcing its utility for investigating integrative brain mechanisms in fMRI studies using the Functionnectome.

Next, we assessed the functional coherence of each parcellation by computing their overall homogeneity, which quantifies the similarity of the temporal variation series among voxels within each parcel. A homogeneity value closer to one indicates high similarity among voxel time series, suggesting that the parcellation captures a meaningful functional unit. We applied this analysis to an independent group (i.e., the validation group) to ensure that the observed functional coherence was not specific to the original dataset. We applied the parcellation to the validation group and computed the homogeneity metric for both the associative and commissural parcellations. The associative parcellation yielded a homogeneity metric of 0.92, while the commissural parcellation had a homogeneity metric of 0.96, indicating strong within-parcel similarity in both cases. To evaluate whether these values were significantly higher than expected by chance, we compared them to the distribution of homogeneity metrics obtained from 1000 random parcellations. For each parcellation, a Z-score was calculated by subtracting the mean of the null distribution and dividing by its standard deviation, quantifying how many standard deviations the actual parcellation's homogeneity deviated from the null mean. The resulting Z-scores were 8.05 for the association parcellation and 21.03 for the commissural parcellation. These high Z-scores indicate that the functional homogeneity of the actual parcellations lies outside the range expected by chance. Thus, both parcellations demonstrated greater functional coherence than expected under a null model. Together, these results validate the functional specificity and robustness of the parcellations.

Linking white matter parcellation to cognition via a cognitive morphospace

To relate our white matter parcellation to cognitive processes, we developed a cognitive morphospace: a neuro-AI framework embedding meta-analytic fMRI activation maps¹³ into a low-dimensional space built from activation profiles across white matter parcels (Fig. 3e, **see Supplementary Data 2 for an interactive 3D file**). We projected each time point into the morphospace using the group-level temporal variation series from the brain parcellation dataset. We identified the closest cognitive term based on minimum Euclidean distance. This approach allowed us to determine the most likely cognitive processes associated with each time point based on its projections into the morphospace, linking dynamic brain activation patterns to cognitive terms.

To validate this approach, we projected three external datasets into the morphospace: the movie-watching paradigm (HCP - Validation Dataset), the Naturalistic Neuroimaging Database (NNDb – 500 Days of Summer)²⁰, and the Natural Scenes Dataset (NSD)²¹. Each time point was positioned within the morphospace according to its parcel-wise white matter activation profile. We identified the closest cognitive term in the morphospace for each time point in these datasets. While not all terms were observed in all datasets, we highlighted representative cognitive terms consistently activated across tasks and samples (Fig. 3f). These included functions such as emotion, actions, hand movements, and facial recognition, supporting the generalisability of the morphospace across different experimental contexts.

Finally, ChatGPT 4o was employed to produce experimenter-unbiased descriptions of the video frames and relate them to the cognitive terms activated by each parcel (available as **Supplementary Data 1**).

Association and commissural fibres have different cognitive profiles.

The association parcellation revealed a functionally organised white matter structure with distinct patterns of hemispheric specialisation. Using the cognitive morphospace of white matter, we identified the most frequently associated cognitive terms linked to the activation of association parcels. These included *priming*, *reading*, *semantics*, *auditory stimulus*, and *actions*. In contrast, the commissural parcellation revealed the functional involvement of fibres supporting a wide range of motor, cognitive, and social-emotional functions across both hemispheres (Fig. 4). The cognitive morphospace highlighted terms most frequently associated with commissural parcels, including *endogenous*, *associative*, *spatial-temporal*, and *hearing*.

To further characterise how white matter organisation supports distinct cognitive domains, we compared the frequency of cognitive term activations during video-watching between association and commissural parcels in the original brain parcellation. For each term, we computed the normalised difference in activation frequency between the two sets of parcels, accounting for the overall term frequency across the movie (Fig. 4). Cognitive terms more frequently associated with commissural parcels included *arousal*, *face*, *selective attention*, and *salience*, among others, while those more regularly related to association parcels included *word form*, *suppression*, and *visuomotor*, among others. These findings confirm that specific white matter tracts differentially support distinct cognitive domains and that the morphospace captures these patterns in a biologically meaningful way.

The activation patterns and the concatenated videos for significant activation time points for each parcel for associative and commissural parcellations are provided on the website <http://cognipact.bcblab.com> and in Supplementary Material (**Supplementary Figs. 2 and 3**)

Discussion

This study redefines the functional landscape of the human brain by revealing that white matter circuits, long considered mere anatomical conduits, exhibit a rich and distinct functional organisation, mapped here for the first time comprehensively using naturalistic fMRI and the Functionnectome approach. By projecting fMRI signals onto white matter pathways, we produced the first comprehensive, cognition-driven parcellation of association and commissural fibres. This finding challenges the prevailing view of white matter as a passive relay system and instead positions it as an active substrate underpinning diverse cognitive functions. Critically, the robustness and reproducibility of this functional architecture were validated in independent datasets underscoring its potential for generalisability and translational neuroscience.

Task-based fMRI is usually the gold standard for linking brain regions to function. However, while task-based fMRI is valuable in isolating specific cognitive processes, it is inherently limited by the predefined

nature of the tasks. It may fail to capture the richness of real-world cognition. The naturalistic stimuli employed in our study allow for the emergence of functionally relevant networks, which contrast with artificially segmented cognitive tasks. Additionally, white matter research has focused on lesion-based (e.g. ^{10,13,26,27}), electrical stimulation (e.g. ²⁸) or anatomical classifications (e.g. ^{29,30}) to infer functions indirectly. While these methods have provided valuable insight, they often fail to capture comprehensively the complexity of white matter's contribution to brain function (^{12,13}). Our approach bridges these gaps by integrating functional and structural data in a naturalistic setting, revealing a detailed map of functionally specialised white matter pathways. Unlike previous studies that primarily explored cortical functional divisions ²⁻⁷, our study highlights the dynamic role of white matter in creating a networked brain architecture, enabling the emergence of cognitive functions ⁸.

The Cognipact parcellation reveals distinct patterns of functional specialisation within association and commissural fibres that can be directly explored in our online resource (<http://cognipact.bcblab.com>). This openly available parcellation is designed to facilitate the integration white matter analysis into functional brain mapping, complementing existing cortical atlases ²⁻⁷. For instance, association fibres were predominantly linked to language processing, being more involved in the activation of cognitive terms such as semantics, priming, and reading. Commissural fibres demonstrated bilateral activations associated with motor control, spatial processing, and internal, self-generated cognitive functions, and endogenous processes, as summarised by the ChatGPT generative language model. These findings suggest that white matter actively supports complex, behaviourally relevant functions rather than being a mere relay system.

Surprisingly, we did not observe a functional specialisation within the projection fibres, possibly due to their simpler contribution in relaying information between cortical and subcortical regions. Another explanation is that passive, naturalistic viewing of video does not involve the subcortical networks like active tasks do ³¹. This lack of differentiation could also reflect methodological limitations, such as the resolution of fMRI data, the functional connectome, or the suitability of naturalistic stimuli for detecting such granularity in functional distinctions. Future studies could address this gap using task-based fMRI centred on the functional subdivision of subcortical areas (e.g. ³²) to investigate projection fibres in greater detail.

The parcellations of association and commissural fibres provide a powerful resource for studying brain connectivity and functions. This framework allows future researchers to build more holistic and nuanced models of cognition, by integrating white matter support to complex functional networks. This opens up new avenues of fMRI research, as future studies take into account the contribution of broad dynamics to specific task-based activations. These maps can serve as a reference for investigating disease-specific disruptions in white matter pathways together with their functional association, offering new insights into the fabric of our cognition as well as the neural bases of neurological and psychiatric disorders.

The development of the cognitive morphospace provides a robust framework for interpreting white matter dynamics in terms of cognitive functions. By embedding cognitive terms derived from meta-

analytic fMRI maps into a low-dimensional white matter space, we could link moment-to-moment white matter activity to specific cognitive functions. Notably, the successful projection of independent datasets, collected under different paradigms and conditions, into this morphospace highlights its robustness and generalizability. The consistent identification of functions such as emotion, action, and face perception across these datasets suggests that the morphospace captures fundamental organisational principles of white matter cognition. Furthermore, our analysis revealed distinct cognitive profiles linked with association and commissural fibres, demonstrating the functional heterogeneity of white matter systems. When compared directly, commissural tracts were more frequently involved with cognitive processes involving attention, perception, and inner-state cognition. In contrast, association tracts supported language, memory, visuomotor integration, and executive functions. These findings highlight the utility of the morphospace in disentangling how different white matter systems contribute to support distinct cognitive domains. This approach opens new avenues for comparing cognitive dynamics across tasks, populations, and modalities and exploring how white matter supports flexible cognitive processing in naturalistic settings.

While our study provides strong evidence for the functional organisation of the white matter, the naturalistic stimuli used in our analysis may not fully capture all the functional diversity of the human brain. Future investigation may reveal additional parcels for white matter pathways not characterised in the present study. Additionally, our group analysis does not account for inter-individual variability. However, previous research has demonstrated a high degree of synchrony in neural responses across individuals during naturalistic stimuli, such as movie-watching.¹⁹ Finally, interpreting cognitive functions from naturalistic stimuli remains a challenge, particularly given the data-driven nature of our approach, which reduces dependence on predefined cognitive taxonomies. However, this lack of reliance on often biased existing frameworks¹² enables a more comprehensive and unbiased exploration of brain functions that have been underrepresented in the classical paradigms. To help overcome interpretation difficulties, we provide a tool that derives each parcel's activation to the associated movie frames, and maps systematically parcels and cognitive functions across time, facilitating a more intuitive and context-rich interpretation.

In conclusion, this study provides a robust, data-driven parcellation for understanding the functional organisation of white matter openly accessible at <http://cognipact.bcblab.com>. These results lay the foundation for exploring the intricate interplay between white matter and cognition and provide novel insight and applications in neuroscience.

Methods

Participant data

We analysed data from 110 participants (55 women) available in the HCP 7T release (<https://www.humanconnectome.org/hcp-protocols-ya-7t-imaging>). Each participant's data consists of 3655 volumes corresponding to 1 hour of scanning, including alternate fMRI sessions of resting and

video-watching. The data has been preprocessed and registered to the MNI152 reference space as described in ³³. The acquisition parameters included an EPI gradient-echo sequence, with TR = 1000 ms, TE = 22.2 ms, flip angle = 45 deg, FOV = 208 x 208 mm (RO x PE), slice thickness = 1.6 mm; 85 slices; 1.6 mm isotropic voxels, multiband factor = 5, Image Acceleration factor = (iPAT), and echo spacing = 0.64 ms. The dataset was split into two groups of 55 subjects; the first group served as the discovery group, and the second group served as the replication group.

Video stimuli in fMRI sessions

The fMRI sessions included four short videos featuring two types of content. The first type consisted of two independent short clips sourced from freely available movies under Creative Commons licensing. The second type included two excerpts from Hollywood films previously published by ³³. Each four stimuli included a repeated clip intended for validation across scans. A 20-second rest period was presented before each clip within every movie and after the entire movie. During the rest period, participants were presented with a black screen with the word "REST" in white text.

Applying the Functionnectome method to naturalistic videos

The fMRI data were projected onto the Functionnectome to estimate the contribution of white matter circuits to the cortical variation during the video watching (i.e., Functionnectome time series). The Functionnectome is generated by projecting fMRI signals from grey matter voxels into the white matter, weighing them by the probability of normative structural connectivity between each grey matter voxel and the rest of the brain ¹⁵. Anatomical priors of the brain circuits give this probabilistic structural relationship. Three different anatomical priors were used for this analysis: association, commissural, and projection fibres. Subsequently, at each time point in the Functionnectome time series, we extracted the corresponding volume across all participants and concatenated these volumes to create a group-level dataset for that specific moment. We then computed a voxel-wise one-sample t-test across participants for each time point using the randomise tool available in FSL ³⁴. This step resulted in a 3D group-level map of functional variation in the brain at each time point. These group-level maps were computed separately for the discovery and replication groups. Finally, by examining the sequence of group-level maps, we can look at the variation in each voxel across the video-watching period to explore brain activity variation in relation to the content of the video. From here on, these will be referred to as group-level temporal variation series.

Parcellation of white matter based on functional organisation

We applied dimensionality reduction to the group-level temporal variation series obtained from the Functionnectome projections to examine the large-scale organisation of white matter voxel activity during naturalistic stimulation. Specifically, we aimed to identify whether voxels could be meaningfully organised according to their functional activation patterns across time, potentially revealing an intrinsic structure in white matter functional signals. For this, we used the UMAP ²², a non-linear dimensionality

reduction technique that preserves both local and global data structure. The analysis reduces high-dimensional data into a two-dimensional space where voxels with similar covariation patterns are placed near each other, while dissimilar ones are further apart. We used the implementation available at <https://umap-learn.readthedocs.io/en/latest> with default parameters ($n_neighbors = 15$, $min_dist = 0.1$), which have shown to be the optimal trade-off between maintaining the global structure without spurious neighbouring connections¹³. This allowed us to visualise and compare the spatial-functional organisation of white matter activation patterns. The analysis was conducted separately in the discovery and replication datasets to assess the reproducibility of the observed structure.

To assess this reproducibility, we randomly sampled 1000 points from each UMAP embedding (i.e., UMAP of both discovery and replication groups) using a fixed random seed to ensure consistency across datasets. Pairwise Euclidean distance matrices were computed for each group, capturing the spatial configuration of the points within the two-dimensional embedding space. Pearson correlations between these distance matrices were then calculated to the similarity in spatial organization between the discovery and replication groups. To ensure the robustness of our findings and verify that they are not merely due to chance, we created a null model by generating a third UMAP embedding using shuffled voxel-level data. Specifically, for each subject in the discovery group, the voxel order was randomly permuted before transposing the time series into subject-by-voxel matrices, thereby destroying voxel wise spatial correspondence across subjects. The resulting randomized data were used to generate a new UMAP embedding, trained using the same parameters as the original. We then computed the Euclidean distance matrix of this null embedding and correlated it with those from the actual discovery and replication UMAPs. Lower correlations with the null model would suggest that the observed structure in the actual UMAP embeddings reflects consistent functional organization across individuals.

Given the similar distribution of the discovery and replication groups (see Fig. 1), we combined data from both groups for the clustering data analysis. Time course variation profiles for each brain voxel across all 110 participants (including discovery and replication groups) were entered into a UMAP. Clusters within the UMAP space were identified using the HDBSCAN model^{23,24}. HDBSCAN is a hierarchical clustering algorithm that extracts a flat clustering based on the stability of clusters while accommodating varying densities and identifying outliers as noise points. We used the package available at <https://hdbscan.readthedocs.io/en/latest/index.html#>.

To determine the optimal HDBSCAN model, we explored different parameters of *min_cluster_size* and *min_samples*, ranging from 50 to 500, with an interval of 50. To evaluate the performance of each parameter combination, we employed a fast approximation of the DBCV score²⁵ as a relative measure ranging from -1 to 1, where a higher DBCV score indicated better clustering performance. Finally, the clusters identified using the HDBSCAN clustering algorithm were projected back onto the brain. This step provided the first parcellation of the white matter based on functional activations.

Out-of-sample validation group

We analysed data from a third group comprising 20 participants (16 women) available in the HCP 7T release. These participants underwent the same acquisition protocol described for the previous 110 participants. However, this data was acquired before updates were made to the movies used during the HCP 7T release scanning sessions. Therefore, in this data, the start and end points of the movies are not synchronised with the TR of the fMRI scans, with deviations ranging from 0 to approximately 200 ms for three of the four movie files.

We performed the same analysis described before: the fMRI data were projected onto the Functionnectome, and each time point of the time series was entered into a one-sample t-test.

To validate our association and commissural parcellations, we assessed the consistency of voxel-wise temporal variations within each parcel across the brain parcellation and validation groups. Specifically, we computed Pearson correlations between the temporal variation series of each parcel in the two groups.

For further validation, we evaluated the functional homogeneity of the parcellations within the validation group by comparing them to a null model. A parcellation homogeneity metric was computed to quantify the similarity of the group-temporal variation series within each parcel, reflecting the functional coherence of the brain regions defined by the parcellation. We extracted the time series data of all voxels for each parcel from the group-level temporal variation series. We then computed a correlation matrix for each parcel using a pairwise Pearson correlation coefficient between all voxel pairs. These coefficients were Fisher-Z transformed, and the homogeneity value for each parcel was defined as the average of these Fisher Z-transformed correlations. To account for parcels with a large number of voxels, a block-wise approach was applied when the voxel account exceeded a threshold of 40000 voxels (it only occurs for one parcel in the commissural parcellation). This method involved dividing the parcel into smaller blocks of 5000 voxels, computing correlations within and between blocks, and then averaging these correlations to obtain the final homogeneity value. To compute the overall homogeneity of the parcellation, each parcel's homogeneity value was weighted by its number of voxels, and the weighted sum was divided by the total number of voxels across all parcels (as in ⁶), as follow:

$$\text{Overall homogeneity} = \sum_{i=1}^N h_i * v_i / \sum_{i=1}^N v_i$$

where h_i is the homogeneity of parcel i , and v_i is the number of voxels in that parcel. A value closer to 1 indicates that voxels within each parcel exhibit highly similar activity patterns, reflecting a more homogeneous parcellation.

To assess whether the observed homogeneity of the actual parcellations exceeded chance levels, we generated a null distribution using 1000 random parcellations with the same number and size as in the original parcels. For each parcellation, associative and commissural, the random parcellations were created using a region-growing algorithm that started from a randomly selected seed voxel within the brain. Each parcel was grown iteratively by adding neighbouring voxels until it reached the size of the corresponding original parcel, ensuring contiguity within parcels and no overlap between parcels. We

extracted the group-level temporal variation series for each brain voxel for each random parcellation and computed the overall homogeneity, as described above. These random parcellations provided a null distribution of homogeneity values, which we used to compare against the homogeneity of the actual parcellation to evaluate its statistical significance. We calculated a Z-score to quantify how far the actual overall homogeneity deviated from the null distribution. It was calculated as the difference between the homogeneity of the actual parcellation and the distribution of random homogeneities:

$$Z = \frac{(\text{homogeneity of the actual parcellation} - \text{mean of the random homogeneities})}{\text{standard deviation of random homogeneities}}$$

High absolute Z-scores indicate a statistical difference between the homogeneity of the actual and that of the random parcellations. A greater Z-score can be interpreted as a better homogeneity of the actual parcellation than the distribution of homogeneity values from the random parcellations. In addition, a permutation based p-value was computed by comparing the observed difference between the real homogeneity and the mean of the null distribution against a distribution of differences obtained from shuffling the real and null values 10000 times.

Mapping white matter–cognition relationships via neuro-AI

Naturalistic Neuroimaging Database (NNDb)

We used publicly available data from the Naturalistic Neuroimaging Database (NNDb v2.0; <https://www.naturalistic-neuroimaging-database.org>)²⁰, which includes full-length fMRI recordings of 86 participants (<https://openneuro.org/datasets/ds002837/versions/2.0.0>). From this dataset, we selected 20 participants (10 women, mean age 27.7 ± 10.1, min = 19 max = 53) who watched the same movie, 500 Days of Summer, a film in the romance genre. All participants were right-handed native English speakers. Functional MRI data were acquired using a 1.5T Siemens MAGNETOM Avanto scanner with a 32-channel head coil (Siemens Healthcare, Erlangen, Germany), employing multiband EPI with a multiband factor of 4 and no in-plane acceleration. Acquisition parameters were as follows: TR = 1 s, TE = 54.8 ms, flip angle = 75°, 40 interleaved slices, and 3.2 mm isotropic resolution. The movie lasted 5470 seconds and was split into two scanning sessions on separate days, with at least one break between them due to a software limitation of the EPI sequence. Breaks were kept to a minimum. For full details on data acquisition and preprocessing procedures, see²⁰. Briefly, preprocessing was conducted using AFNI³⁵ and included slice timing correction, despiking, volume registration to the mean functional image (centre of the time series), alignment to the anatomical scan, and normalisation to the MNI template. The data were then detrended. We used the preprocessed, detrended, and concatenated time series with ICA-based artifact removal applied (sub-_task-_bold_blur_censor_ica.nii). As an additional step, we registered the data to MNI space at 2 mm resolution using FSL's FLIRT. Finally, we performed the same analysis as for the HCP data: the fMRI data were projected onto the Functionconnectome, and for each time point in the resulting time series, a voxel-wise one-sample *t*-test was conducted across participants.

The Natural Scenes Dataset (NSD)

We used data from 8 right-handed participants (6 women; mean age = 26.5 ± 4.2 years; range: 19–32) with no known cognitive impairments or colour blindness, and with normal or corrected-to-normal vision. High-resolution functional MRI data were obtained from a publicly available dataset (see ²¹ for full details). Functional data were collected at 7T using gradient-echo EPI with 1.8 mm isotropic resolution, TR = 1.6 s, TE = 22 ms, and multiband acceleration (factor 3), providing whole-brain coverage including the cerebellum. The full dataset includes 40 scanning sessions per subject, with 12–14 runs per session. Preprocessed time series and experimental design files were downloaded from <http://naturalscenesdataset.org>.

In the NSD experiment, participants performed a long-term continuous recognition task while viewing thousands of natural scene images. Each participant viewed approximately 10,000 distinct images, each repeated three times across the sessions. Among these, a subset of 1,000 images was shared and presented to all participants, while the remaining images were unique to each individual. For our analysis, we identified 515 images presented to all subjects three times. The images were selected from Microsoft's COCO dataset ³⁶. Each trial lasted 4 seconds, with a 3-second image presentation followed by 1 second of fixation, and runs included 75 trials (63 image trials and 12 blank trials). The temporal ordering of trials was identical across subjects. The data were temporally resampled for preprocessing to correct for slice timing differences and spatially resampled for motion, EPI distortion, and gradient nonlinearities. We applied further preprocessing including high-pass filtering and nuisance regression to remove motion-related confounds, with demeaned regression residuals. The cleaned data registered to MNI space at 2 mm resolution using FSL's FLIRT tool.

Next, we identified the indices corresponding to the 515 shared images presented to all participants. For each image, we selected the volume corresponding to each image presentation and averaged the two subsequent volumes (approximately 3.2 to 4.8 seconds post-stimulus) to account for the delayed hemodynamic response, across the three repetitions to obtain a mean activation map per concept per participant. These mean maps were then concatenated across participants to form a group-level dataset, which was used as input for the Functionconnectome analysis. Finally, each time point in the resulting time series was entered into a voxel-wise one-sample t-test, following the same steps as described for the HCP processing.

White matter cognitive morphospace

To characterise the organisation of cognitive processes in the brain's white matter, we created a morphospace following the procedure described in Pacella et al ¹³. We used 506 meta-analytic association maps from Neurosynth (www.neurosynth.org), which are statistical representations of fMRI activation patterns linked to specific cognitive terms ¹. These maps were projected from grey matter to white matter using Functionconnectome, yielding white matter-based meta-analytic representations of cognition. Each projected map was then parcellated using our association and commissural white

matter parcellations. This process reduced each meta-analytic map to a vector of average intensities within each parcel, creating a structured representation of the cognitive terms in anatomical white matter space. The resulting matrix had 506 rows (one per cognitive term) and 76 columns (one per parcel, including association and commissural parcels). For further details on the meta-analytic dataset, see Pacella et al ¹³.

To explore how cognitive functions are represented and related within white matter, we applied UMAP to the parcel-wise intensity matrix to generate a 3-dimensional embedding called the white matter cognitive morphospace (Fig. 3e and **Supplementary Data 2**). Each point corresponds to a cognitive term in this space, and the distance between points reflects the similarity in their spatial distribution across parcel-wise white matter representation. The coordinates of the 506 cognitive terms and the full embedding model were saved to allow the projection of new data.

Projecting white matter functional dynamics into the cognitive morphospace

We projected the group-level temporal variation series from our brain parcellation into the previously built cognitive morphospace to explore the temporal dynamics of cognitive representations in white matter. These maps, containing significant t-values within each of the 76 parcels across 3655 time points, were transformed using the pre-trained UMAP model. This resulted in a set of new 3D coordinates, with each point capturing the cognitive profile of a specific time point within the same space as the original meta-analytic cognitive terms.

We repeated this projection for the validation group and two additional independent datasets to validate our morphospace approach. This confirmed that mapping white matter functional dynamics onto cognitive coordinates was reliable and generalisable across different samples, enabling us to identify the cognitive terms engaged at each time point based on white matter variation.

Decoding cognitive representations from white matter activity

To interpret the cognitive content of each time point, we computed the Euclidean distance between each projected time point and the 506 cognitive term coordinates in the morphospace. This resulted in a distance matrix with dimensions [Number of time points × 506 cognitive terms]. We identified the nearest cognitive term for each time point by selecting the one with the smallest distance value. The corresponding term index and label were extracted for further analysis.

Parcel activations during naturalistic video viewing

To identify activations within each parcel while viewing the naturalistic videos, we computed the mean and standard deviation (SD) for each time point in the group-temporal variation series belonging to each parcel. Time points with a mean exceeding + 2 SD exhibited heightened activation. Subsequently, we matched these activated time points with the corresponding moments in the naturalistic videos, accounting for a delay of 5 seconds between each time point and the video time. All this material is available at <http://cognipact.bcblab.com>. Users can select specific brain parcels and view the

corresponding video segments that activate them, facilitating deeper investigation into the functional roles of distinct white matter tracts.

Differential cognitive profiles of association and commissural parcellations

For each cognitive term, we computed a normalised difference in activation frequency across tract classes, considering how often the term appeared throughout the movie. Specifically, we calculated the frequency of activation in association and commissural parcels relative to the total number of time points in which each term was identified, and then applied the formula:

$$(\text{Association} - \text{Commissural}) / (\text{Association} + \text{Commissural})$$

This yielded a measure of parcel-specific activation preference that corrects for variability in term prevalence across the stimulus.

Declarations

Acknowledgments

This work is supported by HORIZON-INFRA-2022 SERV (Grant No. 101147319) 'EBRAINS 2.0: A Research Infrastructure to Advance Neuroscience and Brain Health', by the European Union's Horizon 2020 research and innovation program under the European Research Council (ERC) Consolidator grant agreement No. 818521 (DISCONNECTOME), the University of Bordeaux's IdEx 'Investments for the Future' program RRI 'IMPACT', the IHU 'Precision & Global Vascular Brain Health Institute – VBHI' funded by the France 2030 initiative (ANR-23-IAHU-0001) and the NextGenerationEU PNRR grant No. SOE_000013 (EFFORT).

Data Availability

All data used in this study are publicly available. Human Connectome Project (HCP): <https://www.humanconnectome.org>, Natural Scenes Dataset (NSD): <https://naturalscenesdataset.org>, Naturalistic Neuroimaging Database (NNDb): <https://openneuro.org/datasets/ds002837/versions/2.0.0>. The association and commissural parcellations from this study, along with all associated video frame annotations for each parcel, are available at <http://cognipact.bcblab.com>. The parcellation NIfTI files can also be accessed via NeuroVault: <https://neurovault.org/collections/19566/>

Code Availability

All scripts used for data processing and analysis are available on GitHub: <https://github.com/MarcelaOvando/CogniPACT>

References

1. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).
2. Smith, S. M. *et al.* Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci.* **106**, 13040–13045 (2009).
3. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
4. Gordon, E. M. *et al.* Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cereb. Cortex N. Y. N 1991* **26**, 288–303 (2016).
5. Shen, X., Tokoglu, F., Papademetris, X. & Constable, R. T. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *NeuroImage* **82**, 403–415 (2013).
6. Schaefer, A. *et al.* Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb. Cortex N. Y. N 1991* **28**, 3095–3114 (2018).
7. Yeo, B. T. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
8. Thiebaut de Schotten, M. & Forkel, S. J. The emergent properties of the connected brain. *Science* **378**, 505–510 (2022).
9. Catani, M. & Schotten, M. T. de. *Atlas of Human Brain Connections*. (Oxford University Press, Oxford, 2012).
10. Talozzi, L. *et al.* Latent disconnectome prediction of long-term cognitive-behavioural symptoms in stroke. *Brain J. Neurol.* **146**, 1963–1978 (2023).
11. Mah, Y.-H., Husain, M., Rees, G. & Nachev, P. Human brain lesion-deficit inference remapped. *Brain* **137**, 2522–2531 (2014).
12. Thiebaut de Schotten, M., Foulon, C. & Nachev, P. Brain disconnections link structural connectivity with function and behaviour. *Nat. Commun.* **11**, 5094 (2020).
13. Pacella, V. *et al.* The morphospace of the brain-cognition organisation. *Nat. Commun.* **15**, 8452 (2024).
14. Marussich, L., Lu, K.-H., Wen, H. & Liu, Z. Mapping white-matter functional organization at rest and during naturalistic visual perception. *NeuroImage* **146**, 1128–1141 (2017).
15. Nozais, V., Forkel, S. J., Foulon, C., Petit, L. & Thiebaut de Schotten, M. Functionnectome as a framework to analyse the contribution of brain circuits to fMRI. *Commun. Biol.* **4**, 1–12 (2021).
16. Nozais, V. *et al.* Atlasing white matter and grey matter joint contributions to resting-state networks in the human brain. *Commun. Biol.* **6**, 726 (2023).
17. Gillig, A. *et al.* GINNA, a 33 resting-state networks atlas with meta-analytic decoding-based cognitive characterization. Preprint at <https://doi.org/10.21203/rs.3.rs-4803512/v1> (2024).
18. Finn, E. S. & Bandettini, P. A. Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *NeuroImage* **235**, 117963 (2021).

19. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).
20. Aliko, S., Huang, J., Gheorghiu, F. et al. A naturalistic neuroimaging database for understanding the brain using ecological stimuli. *Sci Data* **7**, 347 (2020). <https://doi.org/10.1038/s41597-020-00680-2>
21. Allen, E.J., St-Yves, G., Wu, Y. et al. A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nat Neurosci* **25**, 116–126 (2022). <https://doi.org/10.1038/s41593-021-00962-x>
22. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. Preprint at <https://doi.org/10.48550/arXiv.1802.03426> (2020).
23. McInnes, L., Healy, J. & Astels, S. hdbscan: Hierarchical density based clustering. *J. Open Source Softw.* **2**, 205 (2017).
24. Campello, R. J. G. B., Moulavi, D. & Sander, J. Density-Based Clustering Based on Hierarchical Density Estimates. in *Advances in Knowledge Discovery and Data Mining* (eds. Pei, J., Tseng, V. S., Cao, L., Motoda, H. & Xu, G.) vol. 7819 160–172 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013).
25. Moulavi, D., Jaskowiak, P. A., Campello, R. J. G. B., Zimek, A. & Sander, J. Density-Based Clustering Validation. in *Proceedings of the 2014 SIAM International Conference on Data Mining* 839–847 (Society for Industrial and Applied Mathematics, 2014). doi:10.1137/1.9781611973440.96.
26. Foulon, C. et al. Advanced lesion symptom mapping analyses and implementation as BCBtoolkit. *GigaScience* **7**, (2018).
27. Matsulevits, A. et al. Deep learning disconnectomes to accelerate and improve long-term predictions for post-stroke symptoms. *Brain Commun.* **6**, fcae338 (2024).
28. Herbet, G. & Duffau, H. Revisiting the Functional Anatomy of the Human Brain: Toward a Meta-Networking Theory of Cerebral Functions. *Physiol. Rev.* **100**, 1181–1228 (2020).
29. Catani, M., Jones, D. K. & Ffytche, D. H. Perisylvian language networks of the human brain. *Ann. Neurol.* **57**, 8–16 (2005).
30. Thiebaut De Schotten, M., Dell’Acqua, F., Valabregue, R. & Catani, M. Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex* **48**, 82–96 (2012).
31. Shulman, G. L. et al. Common Blood Flow Changes across Visual Tasks: II. Decreases in Cerebral Cortex. *J. Cogn. Neurosci.* **9**, 648–663 (1997).
32. Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & Frith, C. D. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042–1045 (2006).
33. Cutting, J. E., Brunick, K. L. & Candan, A. Perceiving event dynamics and parsing Hollywood films. *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 1476–1490 (2012).
34. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *NeuroImage* **62**, 782–790 (2012).

35. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical research **29**(3), 162-73 (1996).
36. Lin, T.-Y. et al. Microsoft COCO: Common Objects in Context. In European Conference on Computer Vision 740–755 (Springer, 2014).

Figures

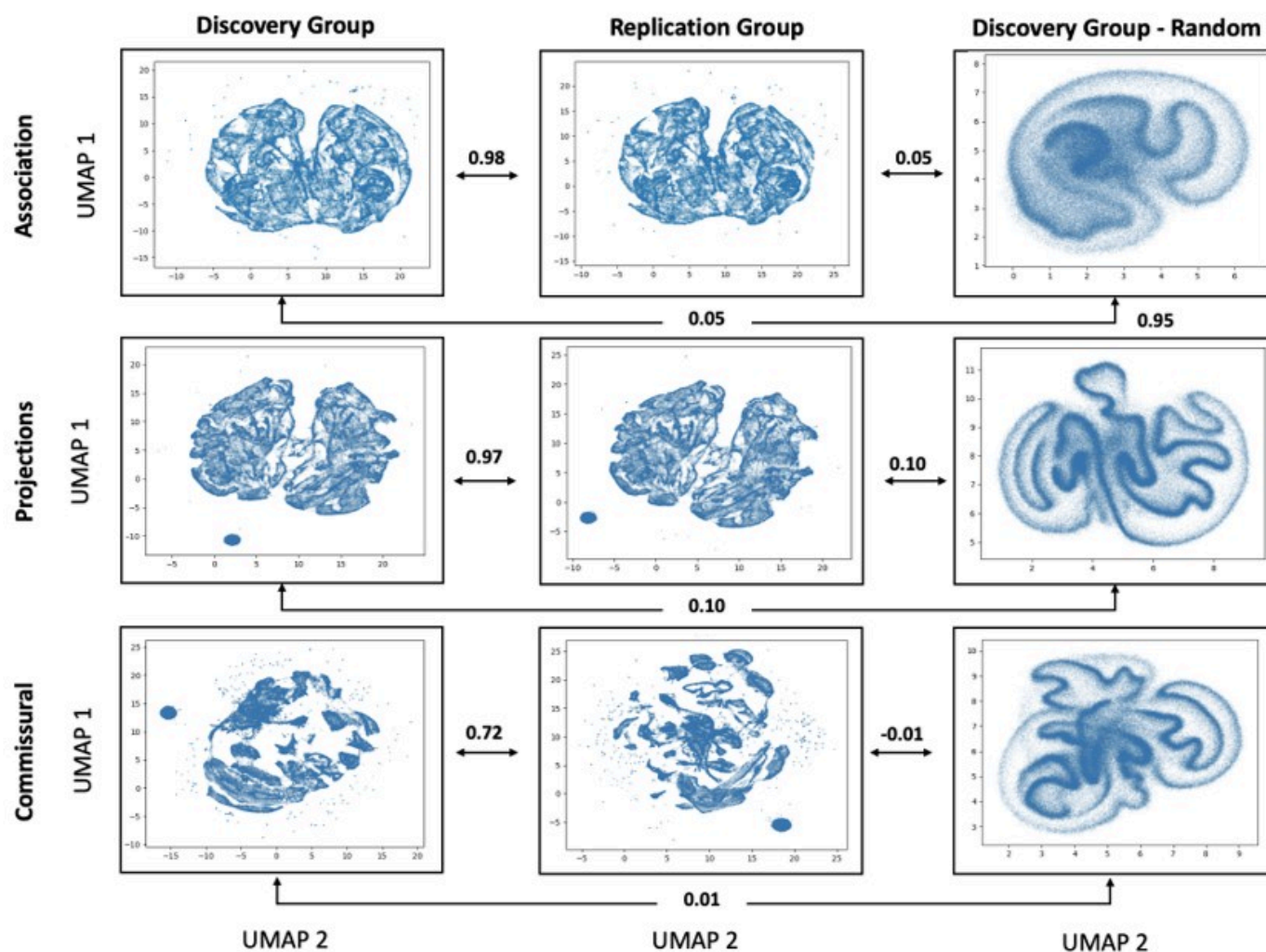


Figure 1

Visualisation of data into the two-dimensional UMAP embedding. UMAP plots illustrating the distribution of data points for the discovery group (left), replication group (middle), and randomised data of the discovery group (right). Each point represents a voxel, and the UMAP algorithm is used for dimensionality reduction to visualise the underlying patterns in the data during naturalistic video-watching.

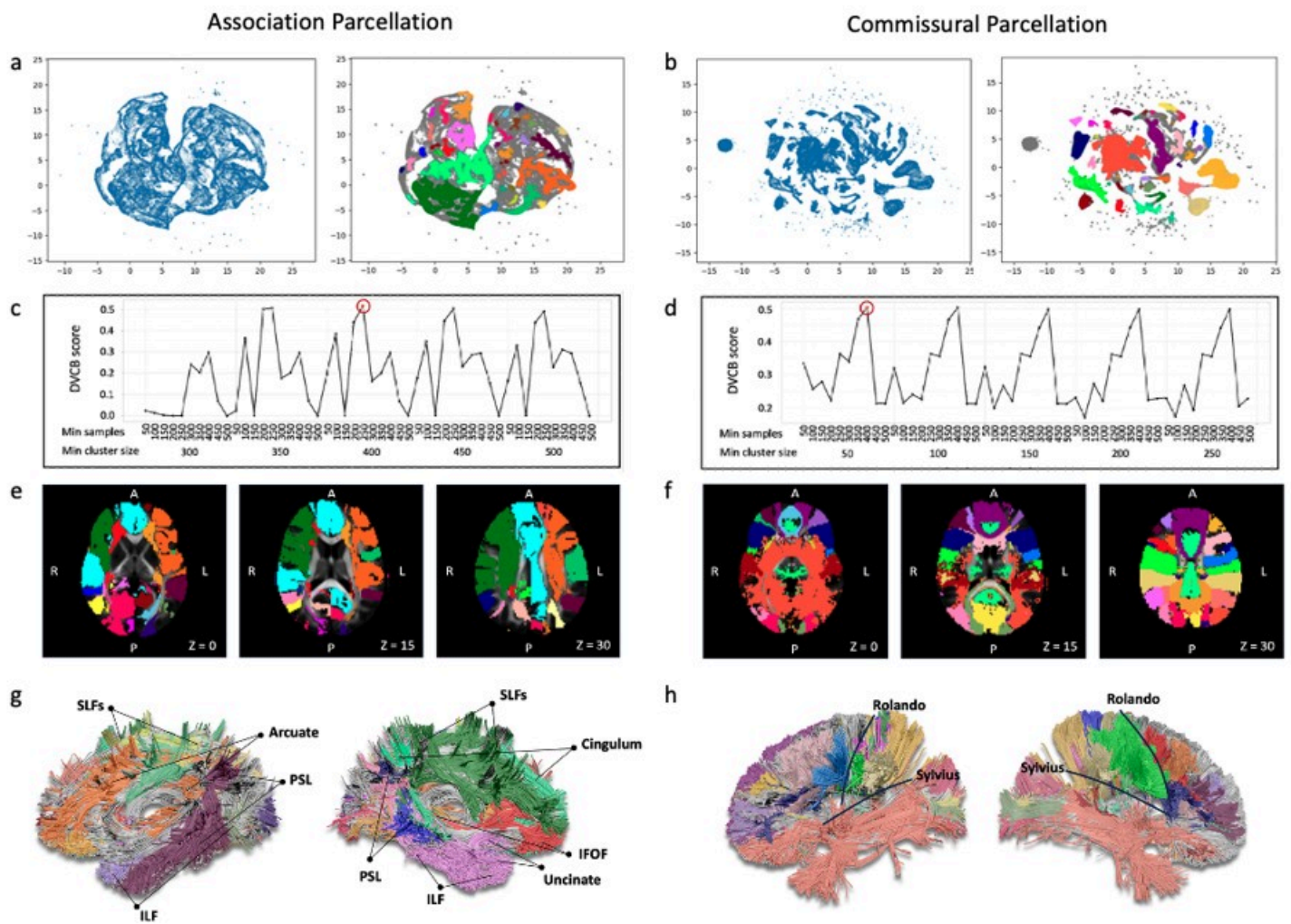


Figure 2

White matter parcellation and tract visualisation for association (left) and commissural (right) fibres. (a-b) UMAP plots show the data points distribution for the combined brain parcellation group (discovery + replication groups). Each point represents a voxel. The UMAP plot in the right panel is colour-coded to indicate different clusters identified by the HDBSCAN model. (c-d) Grid of the clustering models tested for different values of minimum cluster size and minimum number of samples. The model with the highest DBCV score is circled in red, indicating optimal parameters. (e-f) Brain parcellation visualisation in axial slices. A: anterior; P: posterior; R: right; L: left (g-h) Visualisation of the white matter tracts corresponding to parcels shown in (c). For association fibres, principal bundles include the SLFs (Superior Longitudinal Fasciculi), ILF (Inferior Longitudinal Fasciculus), Arcuate, PSL (Posterior Segment of the arcuate fasciculus), Cingulum, and IFOF (Inferior Front-Occipital Fasciculus). Given their large spread and modest anatomical labellisation,⁹ major anatomical landmarks are shown for commissural fibres, including the Rolando fissure (central sulcus) and Sylvian fissure (lateral sulcus).

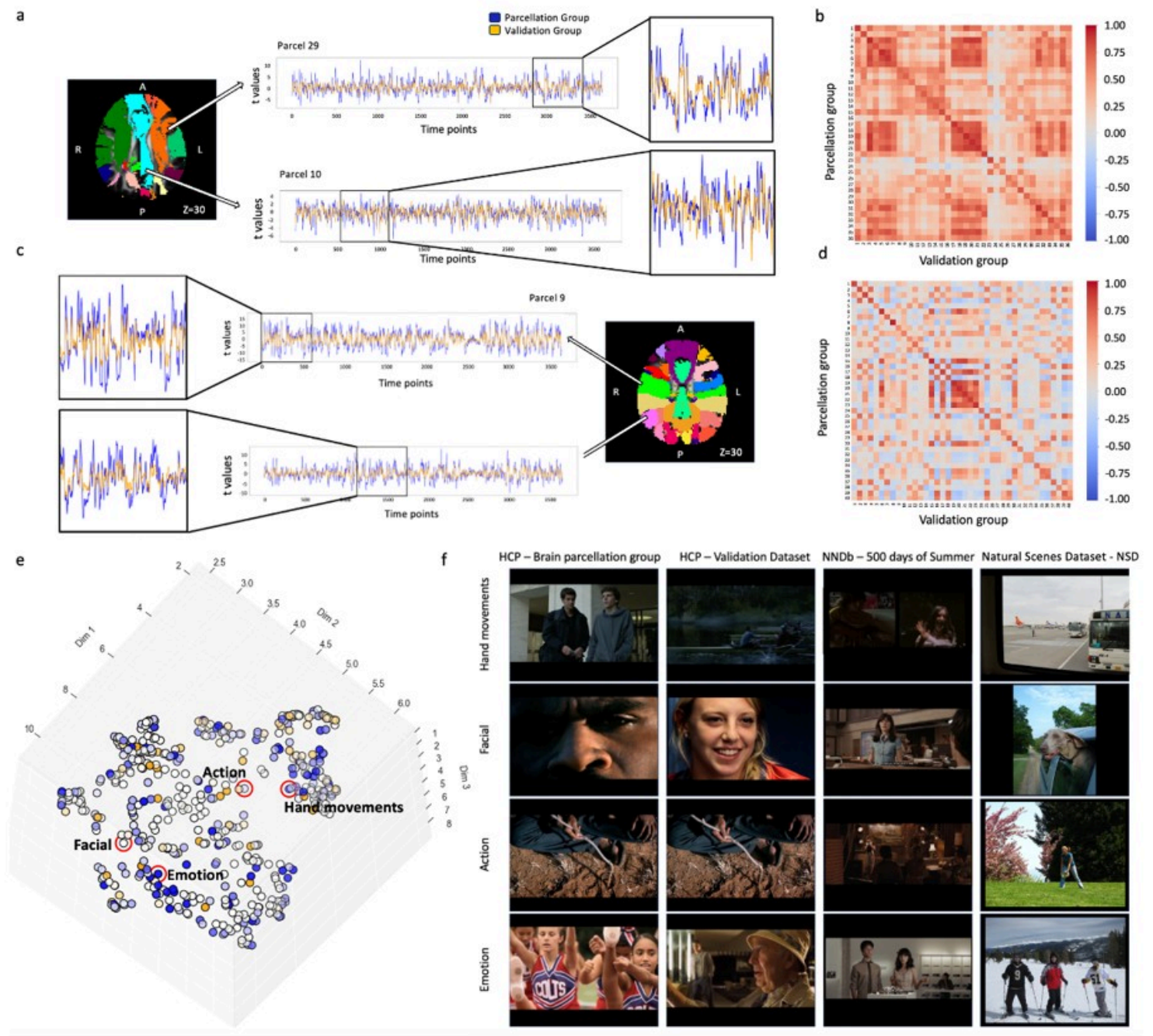


Figure 3

Brain activations in the parcellation and validation groups, and their cognitive embedding in a white matter morphospace. (a, c) Representative time series activation for associative (a) and commissural (c) parcels, showing the overlapping brain activations per time frame for the parcellation (blue) and validation (orange) groups. Zoomed-in segments highlight the similarity in activation patterns across groups for corresponding parcels. (b, d) Correlation matrices for associative (b) and commissural (d) parcellations display parcel-wise correlations between the parcellation and validation groups. Each cell represents the correlation coefficient between matching parcels, demonstrating consistent dynamics

across groups. **(e)** A cognitive morphospace was built by embedding meta-analytic fMRI activation maps of cognitive terms into a low-dimensional space defined by white matter organisation. Each point represents a cognitive term positioned based on its white matter parcel activation profile. Points are coloured based on the normalized difference in activation frequency between association and commissural parcels during video watching. Terms with stronger activation in association parcels appear more orange, and those more engaged in commissural parcels appear more blue (see **Supplementary Data 2** for an interactive 3D visualization). Highlighted in red are terms corresponding to hand movements, facial, action, and emotion. **(f)** To validate the morphospace, three independent datasets were projected into it: a validation sample from the same movie-watching paradigm (HCP - Validation Dataset), the Naturalistic Neuroimaging Database (NNDb - 500 Days of Summer), and the Natural Scenes Dataset (NSD). Each time point was matched to its closest cognitive term based on its white matter activation profile. The figure highlights time frames across datasets that mapped onto the same cognitive terms emphasized in (e). These recurring patterns across distinct experimental contexts support the robustness and functional relevance of the morphospace framework.

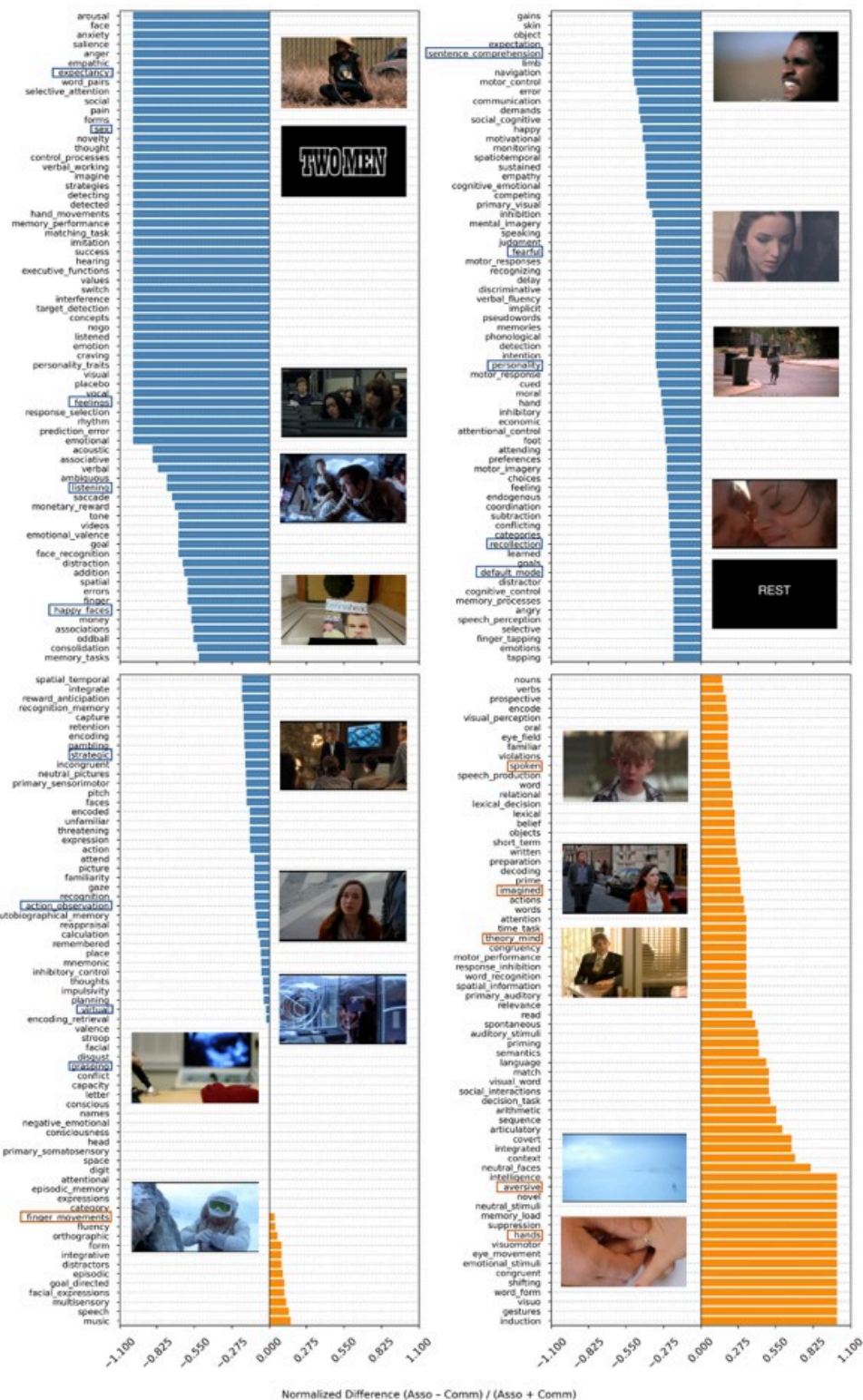


Figure 4

Cognitive term activation in association and commissural parcels. Normalised difference of the activation frequency between association and commissural parcels during the video watching, calculated by (association – commissural)/ (association + commissural). Longer blue barplots (i.e., more negative values) represent the cognitive terms that are more activated in commissural parcels. In contrast, longer orange barplots (i.e., more positive values) represent the cognitive terms that are more

activated in association parcels. To provide context for the cognitive processes involved, we included selected movie frames corresponding to some of the terms shown.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [interactive3Dplot.html.zip](#)
- [20250703SM.docx](#)