Final Paper

Shu Hu

Department of Learning Science, Georgia State University

PSYCH 8630: Topics in Neuroimaging Analysis

Dr. Vince Calhoun

April 27, 2026

Introduction

Classical studies of narrative comprehension mainly utilize text-based materials (e.g., Gernsbacher, 1990; Zwaan et al., 1995a). However, advancements in research methods, particularly fMRI, have enabled the use of naturalistic, multimodal stimuli such as films to reveal the cognitive processes involved in narrative and event comprehension (Magliano et al., 2001; Hasson et al., 2004; Zacks et al., 2001). Narrative comprehension involves multiple cognitive processes, including language processing, semantic knowledge retrieval, mental model building and updating, and participatory responses, etc. For instance, Event Indexing Model suggests that readers and viewers continuously monitor situational dimensions such as time, space, causality, and intentionality to build and update mental models (Zwaan et al., 1995b). When one or multiple situational dimensions change dramatically, people are likely to store the previous event as an integrated model and build a new mental model of event. Similarly, Event Segmentation Theory (EST) suggests that people segment continuous input into discrete event models based on prediction errors (Zacks et al., 2011). Since all these processes can happen simultaneously, traditional introspection- and extrospection-based hypothesis generation processes may limit our ability to capture the underlying cognitive mechanisms. As such, data-driven approaches are becoming more and more important for expanding the observational scope and generating new theories or hypotheses regarding narrative comprehension and event cognition.

Group Independent Component Analysis (ICA) is an important data-driven approach for extracting meaningful spatial and temporal features from fMRI and other data (Calhoun et al., 2001). The two main types of it are spatial ICA (sICA), which identifies components based on spatial independence, and temporal ICA (tICA), which focuses on temporal independence. sICA is more popular in fMRI research as standard fMRI scans generally have far fewer data points on the time dimension compared to the spatial dimension. This would lead to relatively unreliable or unusable tICA components. However, event cognition studies using naturalistic stimuli mainly have longer scanning as they involve tasks such as film or video watching. This could make tICA feasible. More importantly, because participants viewing naturalistic stimuli share a continuous, unfolding narrative dynamic, they are highly likely to exhibit synchronized temporal patterns of cognitive and brain activity (Hasson et al., 2004; Loschky et al., 2015).

While sICA provides valuable insights about how different brain regions cooperate to comprehend, tICA may reveal additional insights about the shared temporal patterns across subjects during narrative comprehension. Understanding the relationship between these sICA and tICA is important for selecting the most appropriate analysis tools in future related research. The present project serves as an exploratory study that compares the spatial and temporal properties of sICA and tICA. Specifically, this project aims to address two questions. First, if tICA focuses more on temporal independence, do tICA components capture temporal patterns better than sICA components? Second, if tICA focuses less on spatial independence, do tICA components still reflect anatomic structure? Or do they involve multiple spatially independent networks? As such, this project aims to explore the spatial and temporal relationship between sICA components and tICA components.

Method

Data

The present project uses open-source functional data from Chen et al. (2017). The original study recruited 22 participants, and data of 17 right-handed native English speakers were kept for analysis. One subject was excluded due to excessive head motion or falling asleep. The participants watched a 50-minute clip of the first episode of the BBC TV series *Sherlock*, which they had never seen before. The viewing process was separated into two functional runs (run 1 and run 2), each lasting around 25 minutes. Participants were asked to pay attention to the movies as normal, knowing that they would be asked to recall about it later. Scanning was performed on a 3T scanner with a TR of 1.5 s, TE of 28 ms, and a flip angle of 64 degrees. The original study also provided researchers labeled scene boundary, which is based on the shift of space. While this is different from event boundary, it can be used as an approximation.

Preprocessing

The preprocessing pipeline fMRIPrep was used for normalization, slice-timing correction, and motion correction (Esteban et al., 2019; 2020). In addition, the preprocessed data of each subject was cropped based on the instruction provided by the original paper. More specifically, the onset and offset timestamp of stimuli for each subject was provided in the dataset, we removed the first and last few TRs of each subject's scan to align their data.

gICA

First we used GIFT to run blind sICA and tICA. The number of independent components was set to 20. As such, we obtained 20 sICA components and 20 tICA components for both run 1 and run 2 data. The algorithm used was Infomax with a default mask. The back reconstruction used GICA. For sICA, the group ICA type was set to spatial, and for tICA, the group ICA type was set to temporal. Both results were scaled using Z-scores, and the number of data reduction (PCA) steps was set to 2. The number of principal components for step 1 was 30, and the number of principal components for step 2 was 20.

Labeling

For the run 1 and run 2 data, the Neuromark 1.0 templates were applied to sICA and tICA components using a maximum correlation approach (Du et al., 2020). Specifically, we calculated the spatial Pearson correlation between the unthresholded spatial map of each component and all templates. Components were labeled by the template with the highest correlation that satisfies $r > 0.3$. sICA components without a template satisfying this requirement were not included in further analysis. For tICA, because there were fewer labeled components, all components without labels were labeled as "Others." Python packages nilearn and nibabel were used for this and later analysis.

Spatial Comparison

To compare the spatial pattern captured by sICA and tICA, we conducted a one-to-many comparison procedure. Given that only one tICA component was identified in the labeling process, and tICA components are not spatially independent, we compared the single tICA

component against all subsets of sICA components (including singletons). For each subset of sICA components, the spatial map of components was aggregated by taking the maximum voxel-wise value across the components. Using maximum voxel-wise value allows us to obtain a union-like representation that preserves the spatial features of all components within the subset.

We used two methods to assess the similarity between the spatial map of the tICA component and the sICA component groups. First, we used the Dice coefficient to evaluate the overlap between the tICA and the aggregated sICA maps. The continuous spatial maps were thresholded at $Z > 2.0$ for binarization. Then, we used spatial Pearson correlation to evaluate the relationship between the voxel weights of the unthresholded maps. We also conducted a permutation test with 1,000 simulations to determine the statistical significance of the spatial correlation.

Temporal Comparison

To compare the temporal patterns captured by sICA and tICA, we conducted a one-to-many comparison procedure. As in spatial comparison, we compared the temporal dynamics of the single tICA component against all subsets of sICA components using a sliding-window approach. Two window sizes were used, 10 TR (15 s) and 30 TR (45 s), and the step size was 1 TR.

To explore the dynamic relationships from different perspectives, we used three different combinations of metrics and a lagged cross-correlation approach (maximum lag was set to 10 TRs). First, to investigate the raw shape similarity, the windowed mean time course of the tICA component was compared with the windowed mean time course of the sICA component groups. Second, to investigate the relation between the magnitude of tICA fluctuations and the co-fluctuation of the sICA component groups, the windowed variance of the tICA time course was compared against the windowed covariance of the sICA component groups' time courses. Third, to investigate the relation between the magnitude of tICA fluctuations and the connectivity within sICA component groups, the windowed variance of the tICA time course was compared with the internal coherence of the sICA component groups' time courses. The internal coherence is calculated based on the dynamic Functional Network Connectivity (dFNC) matrices by averaging the pairwise correlations in Fisher z-transformed space and subsequently transforming the average back to r-space. The statistical significance was evaluated using a group-level permutation test with 1,000 simulations.

dFNC States

While sICA captures independent structures and tICA captures independent time courses, it is valuable to investigate spatial and temporal patterns as a whole. To address this, dFNC methods allow us to explore the temporal dynamics of network connectivity through state-based clustering.

We constructed dFNC correlation matrices using a sliding window with a window size of 30 TRs with 1TR sliding step. Standard Pearson correlation was used to estimate dynamic connectivity. A group-level K-means clustering analysis was then conducted on the aggregated dFNC features across all subjects. For sICA, all labeled components were included in the analysis. For tICA, all components were used as only one is labeled by the template.

To determine the optimal number of connectivity states, we evaluated a cluster range from 2 to 10 using the Davies-Bouldin Index and Silhouette scores. This range was determined to limit the number of states and reduce the complexity of analysis. The K-means clustering outcomes for sICA and tICA were visualized for direct comparison.

Temporal Functional Mode

Another way to investigate spatial and temporal patterns as a whole is Temporal Functional Mode (TFM; Smith et al., 2012), which combines sICA and tICA. For each subject, the extracted sICA time courses were z-score normalized and concatenated along the temporal dimension to create a group-level time course matrix. Utilizing sICA time course provides cleaner input for tICA, and the concatenated matrix provides more data for tICA. Notice that this method assumes there exist shared meta-states among the subjects. To approximate the optimal number of tICA components, a PCA was performed on the concatenated matrix to evaluate the cumulative explained variance. The optimal number of components or meta-states was selected based on the threshold of 85% cumulative explained variance. Then a tICA was conducted on the concatenated matrix with the estimated component number. The weights of the sICA components within each meta-state were visualized, and their group-level temporal dynamics were plotted for direct comparison.

Results

Labeling

In run 1, 1 tICA and 7 sICA components were matched to the templates. In run 2, 1 tICA and 5 sICA components were matched. These results suggest that sICA matched with the templates better than tICA overall. Moreover, for components matched to the same template, sICA shows stronger spatial correlations than tICA. As shown in Figure 1 (auditory template AU_6), for run 1, the spatial correlation between the sICA component and the template is 0.664, yet for tICA it is 0.332. A similar pattern was found in run 2 (see Figure 2). This is not surprising given that the Neuromark templates were obtained based on group sICA and were designed to maximize spatial independence (Du et al., 2020).

Figure 1

Comparison of sICA and tICA components matched to the AU_6 template using run 1 data

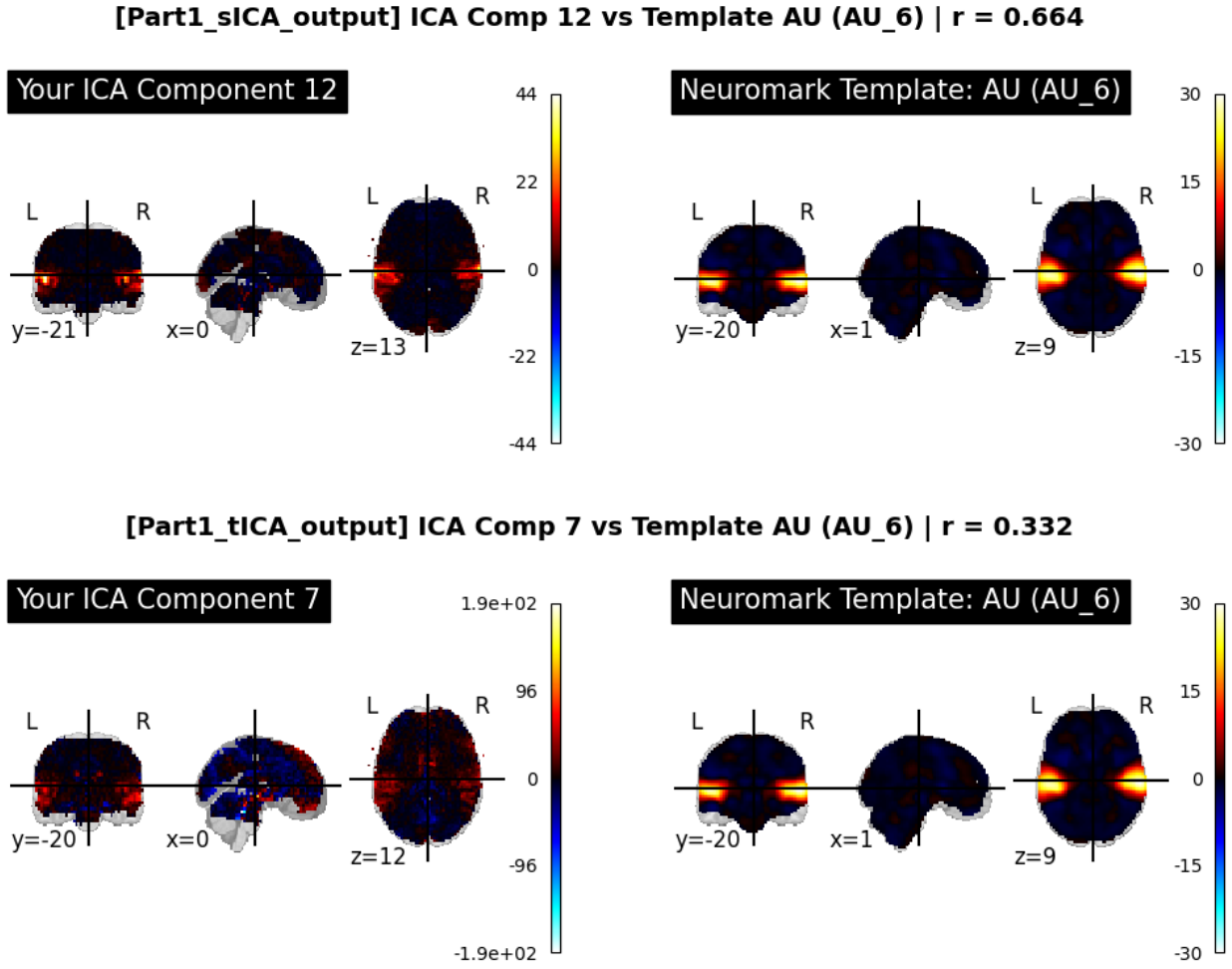
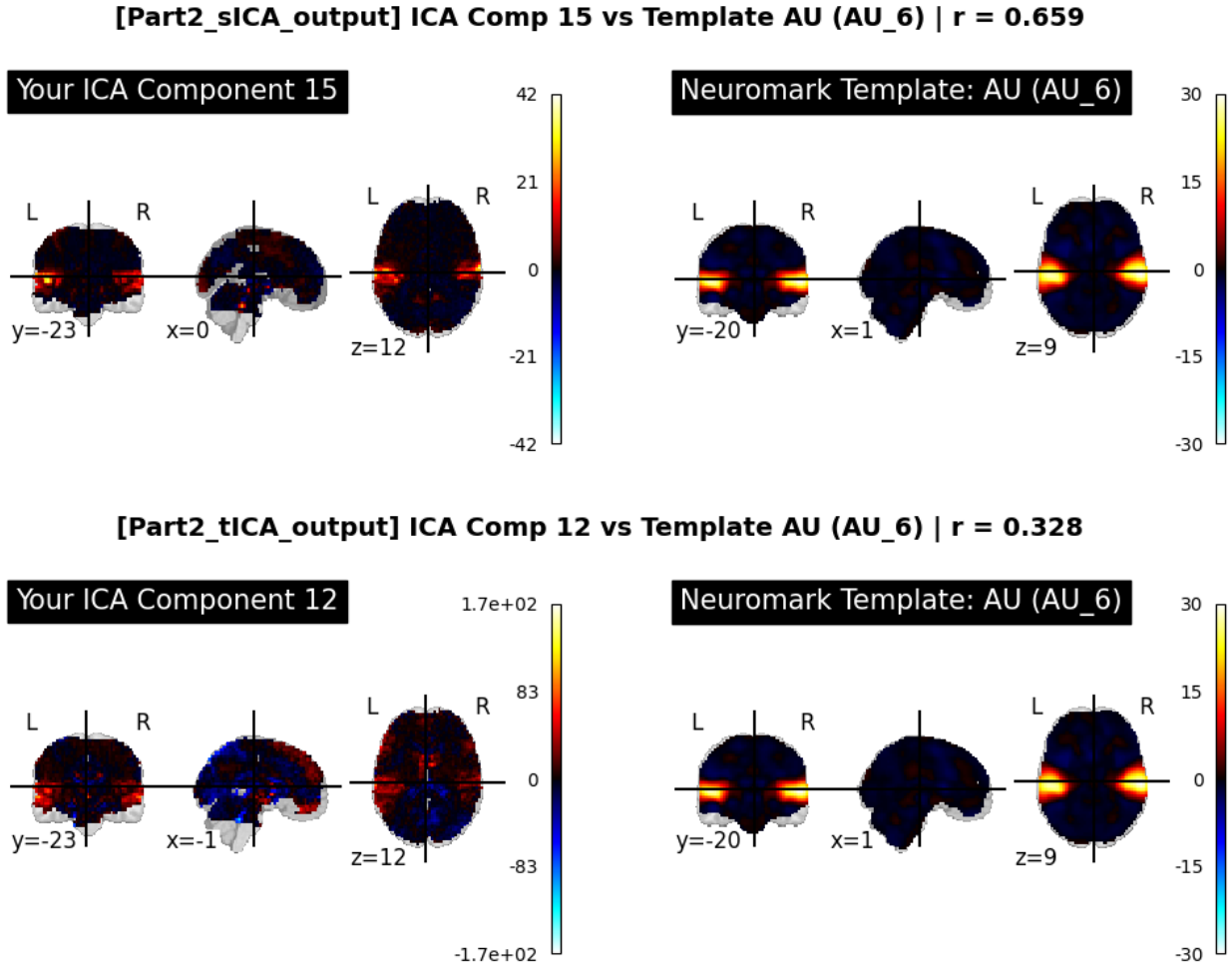


Figure 2

Comparison of sICA and tICA components matched to the AU_6 template using run 2 data



Spatial Comparison

Using the Dice coefficient, the best-matching spatial pairs between the single tICA component and the aggregated sICA component subsets for run 1 and run 2 are listed in Table 1. For run 1, the highest matching pair suggests that the tICA component c7 (labeled by template AU_6) shares the greatest spatial overlap with the aggregated map of all labeled sICA components (dice coefficient = 0.59, $p < .001$). Similarly, for run 2, the best-matching pair indicates that tICA component c12 (template AU_6) shares the most spatial overlap with the aggregated map of all labeled sICA components (dice coefficient = 0.53, $p < .001$).

Using spatial correlation, the best-matching spatial pairs between the single tICA component and the aggregated sICA component subsets for run 1 and run 2 are listed in Table 2. For run 1, the highest matching pair suggests that the tICA component c7 (labeled by template AU_6) is negatively correlated with the aggregated map of sICA component c12 (AU_6) and c19 (DM_43) ($r = -0.34$, $p < .001$). For run 2, the best-matching pair indicates that tICA

component c12 (template AU_6) is negatively correlated with the sICA component c9 (DM_43) ($r = -0.21$, $p < .001$).

These results demonstrate a robust spatial overlap between the single tICA component and the aggregation of the sICA component group, despite the lack of strong linear correlation. This aligns with the property of ICA, where tICA enforces temporal independence rather than spatial independence. Thus, a single tICA component can capture a broader, spatially distributed network that encompasses several distinct, spatially independent sICA components. However, spatial overlap doesn't mean they have shared weights, because sICA emphasizes spatial independence, and the voxels in shared anatomical regions can have different weights across components.

Table 1

Top spatial matching pairs for run 1 and run 2 using dice

Run	tICA component	sICA components	Spatial correlation
1	AU_6 (c7)	AU_6 (c12), CC_27 (c9), DM_48 (c1), DM_45(c13), DM_43 (c19), VI_24(c5),VI_17 (c16)	0.59
2	AU_6 (c12)	AU_6 (c15), DM_43 (c9), DM_45 (c12), SM_15 (c3), VI_17 (c5)	0.53

Table 2

Top spatial matching pairs for run 1 and run 2 using spatial correlation

Run	tICA component	sICA components	Spatial correlation
1	AU_6 (c7)	AU_6 (c12), DM_43 (c19)	-0.34
2	AU_6 (c12)	DM_43 (c9)	-0.21

Temporal Comparison

For run 1, the best-matching combinations for the three comparison metrics are listed in Table 3. The timecourse of tICA component c7 (AU_6) was strongly negatively correlated with the windowed mean time course of the sICA component group (DM_48, DM_43, and VI_24) at both 10TR ($r = -0.79$, $p < .001$) and 30TR ($r = -0.78$, $p < .001$) window sizes. As shown in Figure 3, this strong temporal association is driven by the time courses of multiple sICA components.

Importantly, ICA uses arbitrary signs for component weights and time courses. Therefore, a negative temporal correlation implies that if sICA and tICA capture the same brain activity, their spatial maps should have opposite signs as well. A spatial correlation confirmed this, showing a correlation of $r = -0.23$ between the temporal matching pair in run 1 (see Figure 5).

The correlation between the windowed variance of the tICA time course and the windowed covariance of the sICA components' time courses showed a moderate positive relationship (e.g., $r = 0.49$ for the 30TR window). Similarly, the relation between the tICA variance and the internal coherence of the sICA component group was relatively small but statistically significant ($r = 0.36$ for the 30TR window). These findings suggest that the magnitude of tICA fluctuations reflects the connectivity across sICA components at some level.

Similar relationships between tICA and sICA were observed in run 2 (Table 4 and Figure 4), where the time course of tICA AU_6 and the mean time course of sICA components (DM_43 and VI_17) showed a negative temporal correlation at both 10TR ($r = -0.48$, $p < .001$) and 30TR ($r = -0.42$, $p < .001$) window sizes. This also corresponded to a negative spatial correlation ($r = -0.13$; see Figure 6). Dynamic metrics in run 2, such as variance-to-covariance and variance-to-internal coherence, demonstrated significant but small correlations compared to run 1.

Overall, the temporal matching between sICA and tICA was stronger in run 1 than in run 2. This decrement of similarity might be a consequence of the unfolding of the narrative plots. As the narrative unfolds, participants likely engage in more divergent and endogenous cognitive processes, such as updating complicated mental models or generating different participatory responses. These factors could introduce higher inter-subject differences and reduce the group-level temporal patterns captured by the components.

Table 3

Top time course matching pairs for run 1

Comparison	Window Size (TR)	tICA Component	sICA Components	Correlation
Raw vs. mean	10	AU_6 (c7)	DM_48 (c1), DM_43 (c19), VI_24 (c5)	-0.79
	30	AU_6 (c7)	DM_48 (c1), DM_43 (c19), VI_24 (c5)	-0.78
Variance vs. Covariance	10	AU_6 (c7)	DM_43 (c19), VI_24 (c5)	0.42
	30	AU_6 (c7)	CC_27 (c9), DM_48 (c1), DM_43 (c19), VI_24 (c5)	0.49
Variance vs. Internal Coherence	10	AU_6 (c7)	CC_27 (c9), DM_48 (c1), DM_43 (c19), VI_24 (c5)	0.29
	30	AU_6 (c7)	CC_27 (c9), DM_48 (c1), DM_43 (c19), VI_24 (c5)	0.36

Table 4

Top time course matching pairs for run 2

Comparison	Window Size (TR)	tICA Component	sICA Components	Correlation
Raw vs. mean	10	AU_6 (c12)	DM_43 (c9), VI_17 (c5)	-0.48
	30	AU_6 (c12)	DM_43 (c9), VI_17 (c5)	-0.42
Variance vs. Covariance	10	AU_6 (c12)	DM_43 (c9), VI_17 (c5)	0.25
	30	AU_6 (c12)	AU_6 (c15), DM_43 (c9), SM_15 (c3), VI_17 (c5)	0.29
Variance vs. Internal Coherence	10	AU_6 (c12)	AU_6 (c15), DM_43 (c9), DM_45 (c12), SM_15 (c3)	0.16
	30	AU_6 (c12)	DM_45 (c12), SM_15 (c3)	0.21

Figure 3

Time course for the mean-to-mean best-matching pair of run 1 with a 10TR window

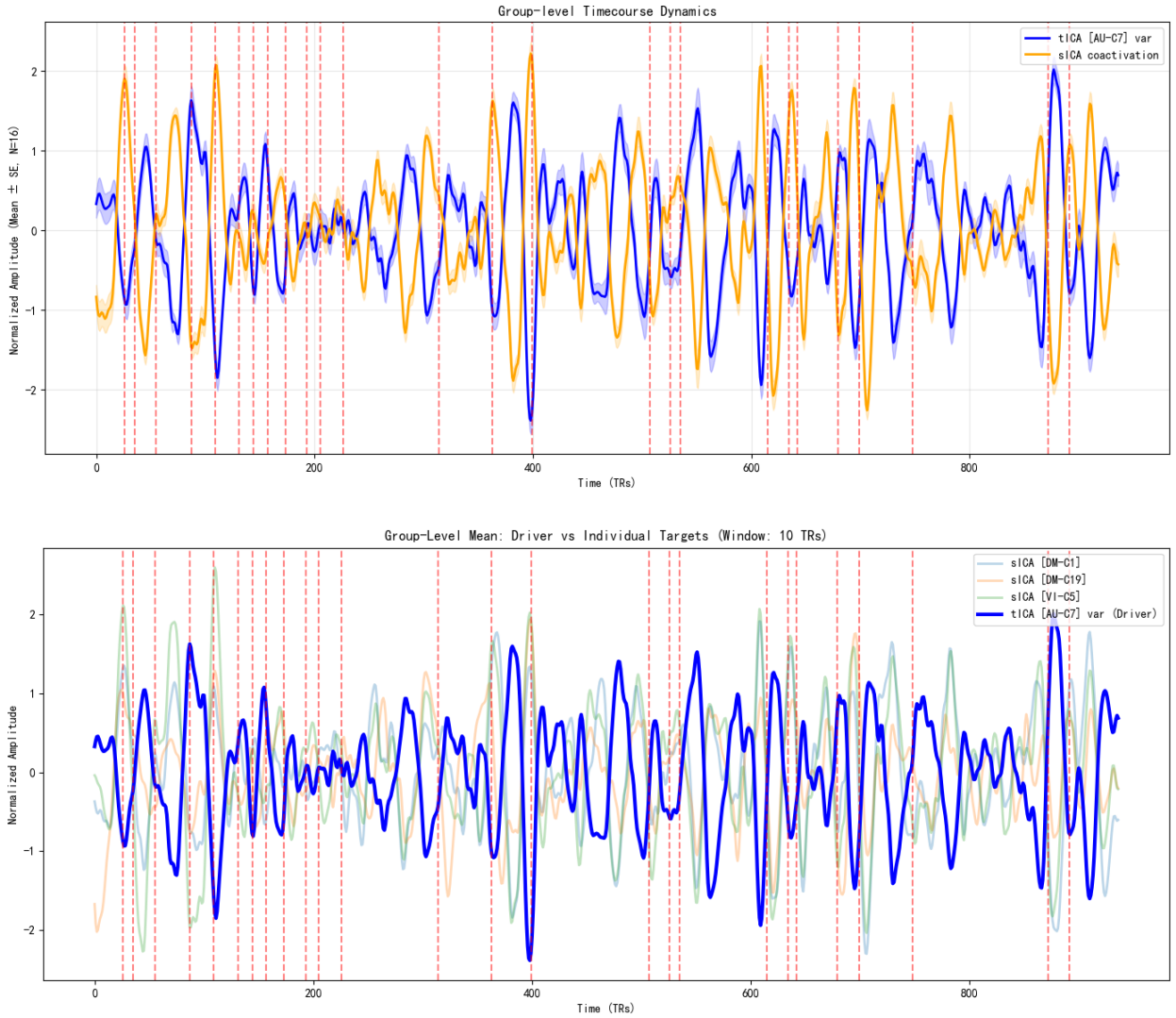


Figure 4

Time course for the mean-to-mean best-matching pair of run 2 with a 10TR window

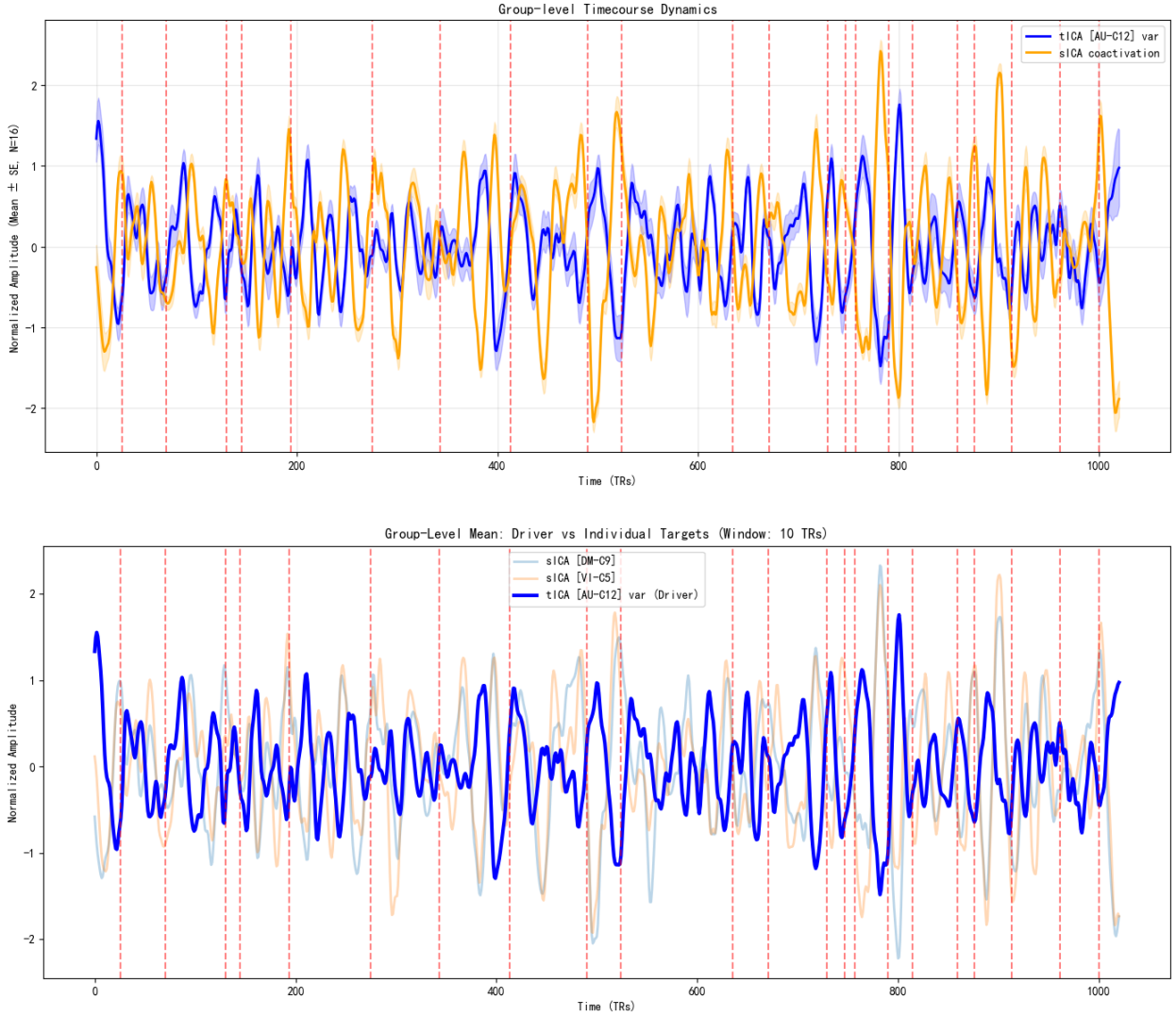


Figure 5

Spatial map of tICA component and sICA component group for run 1

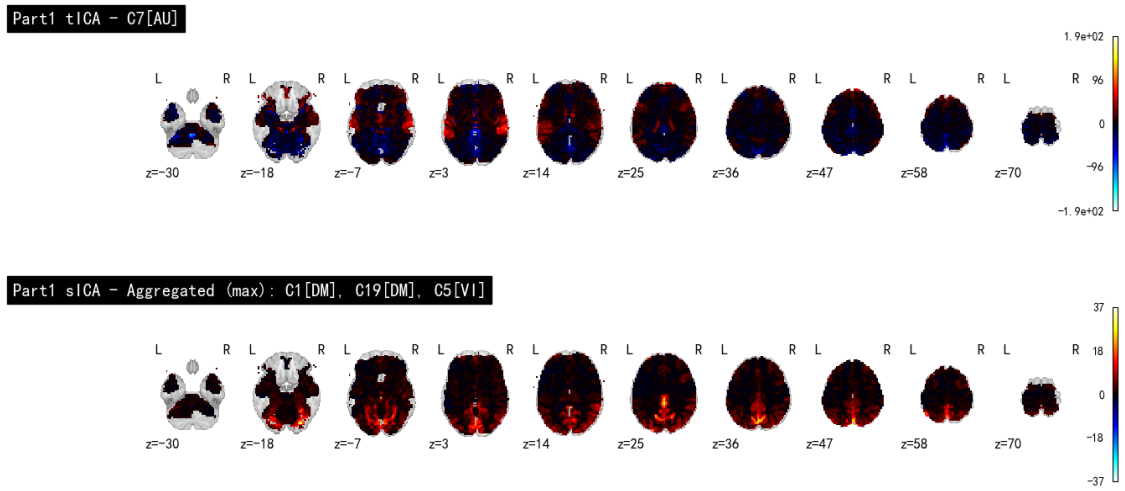
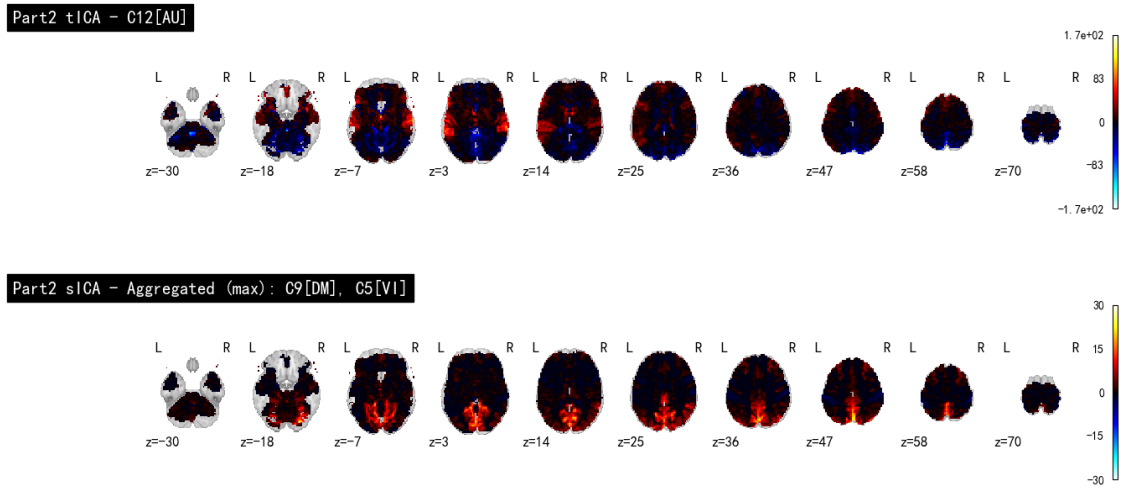


Figure 6

Spatial map of tICA component and sICA component group for run 2



dFNC States

The optimal number of clusters (k) was determined by evaluating the DBI and Silhouette scores, balancing the scores and simplicity to ensure the network states are interpretable. For run 1, the selected k values for sICA and tICA were 6 and 3, respectively (see Figure 7). For run 2, a k of 5 was selected for both sICA and tICA (see Figure 8).

Figure 7

DBI and Silhouette scores for tICA and sICA in run 1

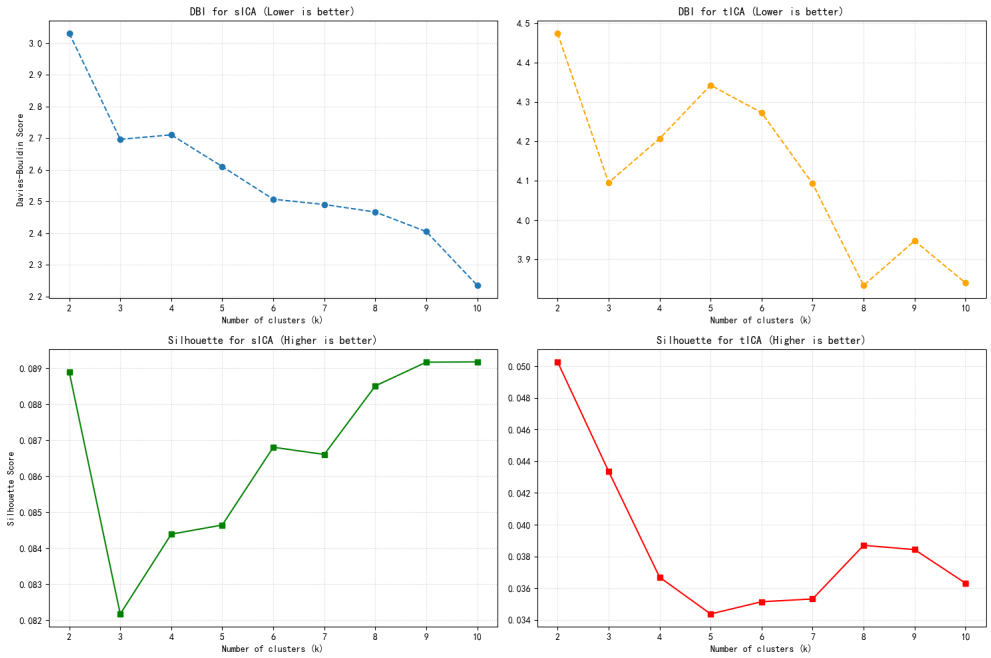
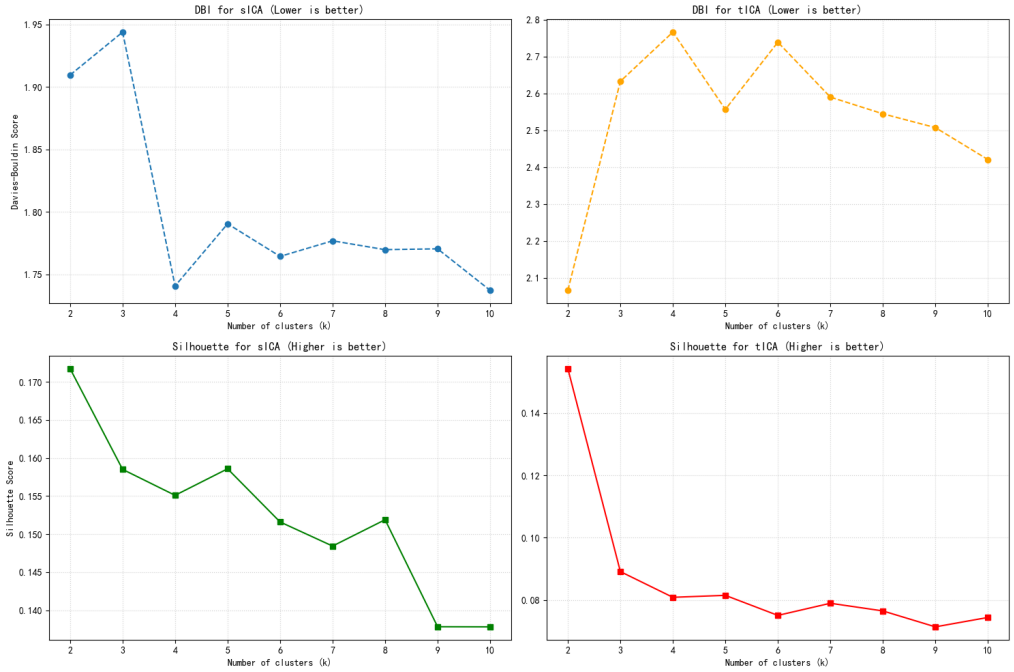


Figure 8

DBI and Silhouette scores for tICA and sICA in run 2



The spatial connectivity patterns for each state are shown in the top panels of Figure 9-12. The corresponding group-level probabilities over time are plotted in the bottom panels of the

figures. The red vertical lines represent researcher-labeled scene boundaries. Visually, the probability dynamics of sICA state appear to align closer with the scene boundaries. In contrast, the probability dynamics of tICA states show less variability and tend to be dominated by one state. Moreover, interpreting the tICA states is more challenging, given that the majority of the components were not classified by the template and may contain signal artifacts.

Figure 9

Spatial connectivity and group-level state probabilities over time for run 1 sICA

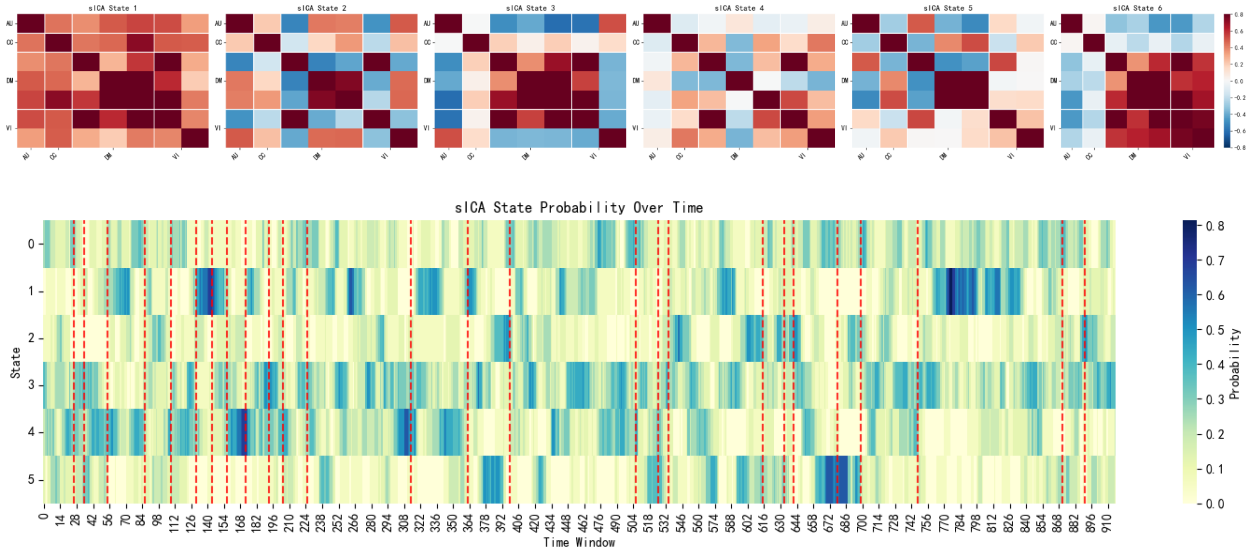


Figure 10

Spatial connectivity and group-level state probabilities over time for run 1 tICA

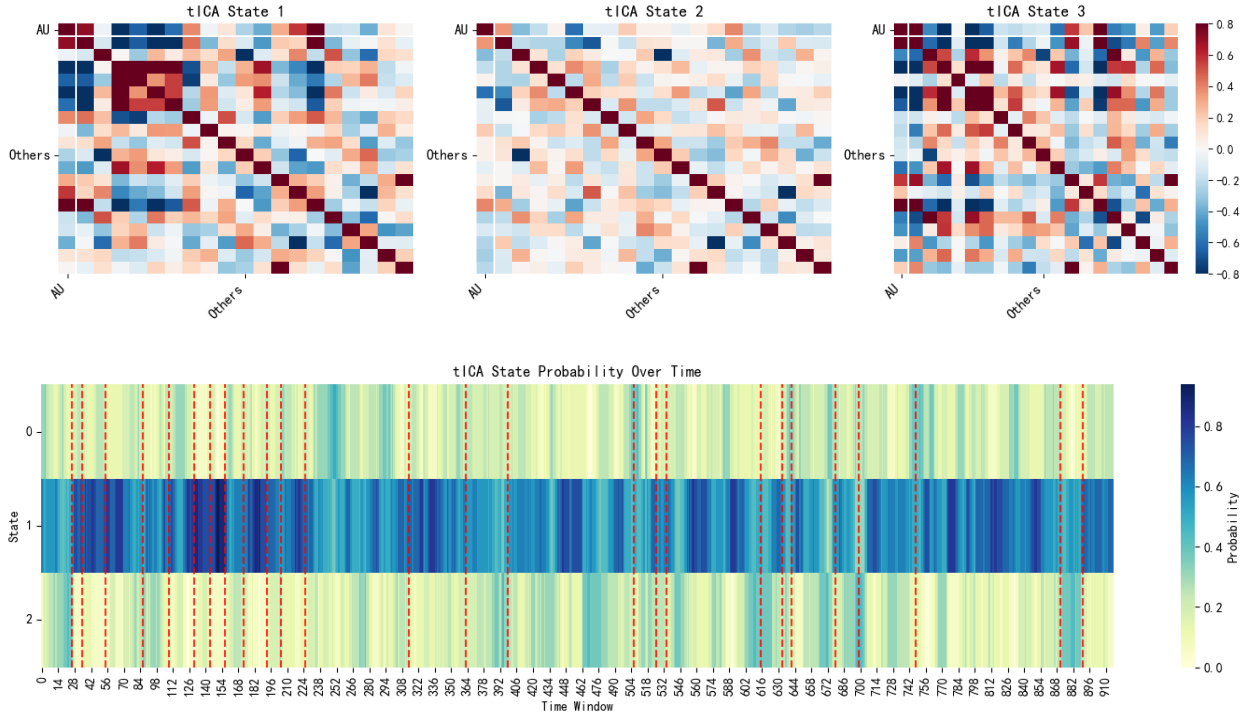


Figure 11

Spatial connectivity and group-level state probabilities over time for run 2 sICA

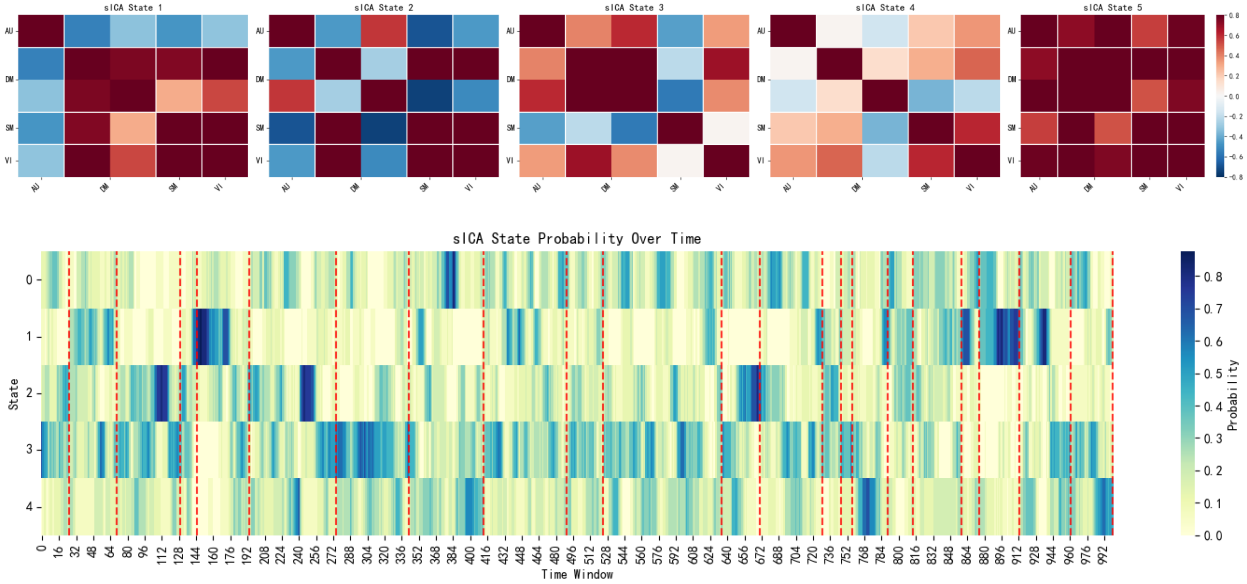
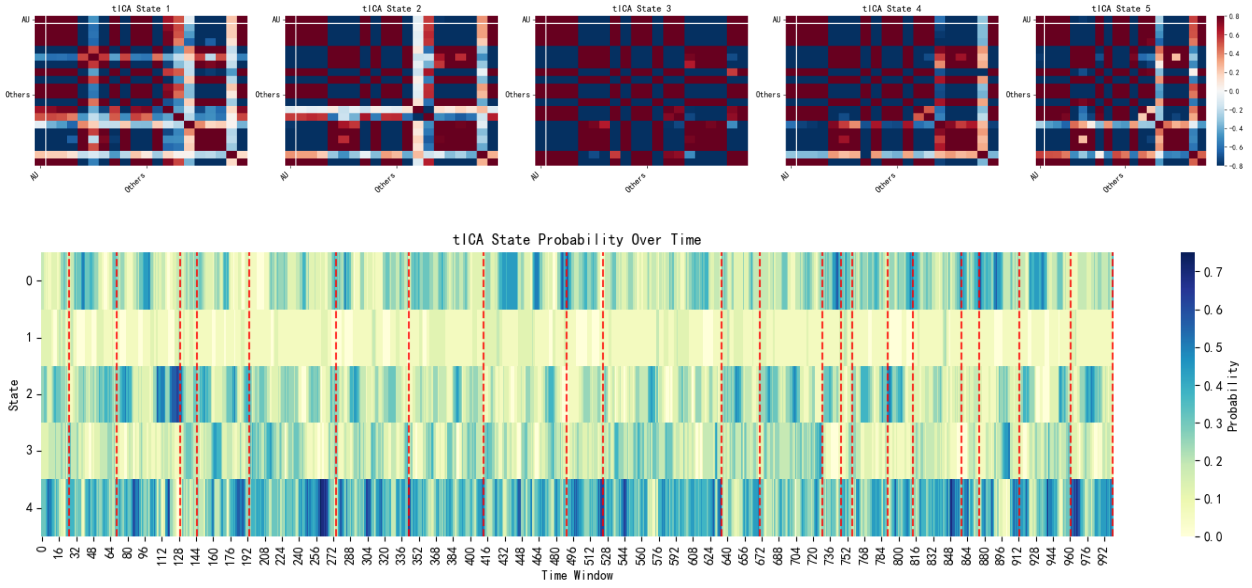


Figure 12

Spatial connectivity and group-level state probabilities over time for run 1 tICA



Temporal Functional Mode

Based on the cumulative explained variance of the PCA, the optimal number of tICA components (meta-states) was estimated to be 5 and 3 for run 1 and run 2, respectively.

Figure 13 shows the weights of sICA components contributing to each of the 5 meta-states in run 1. These meta-states represent unique functionally connected networks. For example, meta-state 1 represents a strong positive contribution from AU_6 (c12) and VI_17 (c16), with negative weights of DM_45 (c13) and DM_43 (c19). In contrast, meta-state 5 likely represents a globally shared temporal pattern except VI_17 (c16) as the sign is arbitrary. Similarly, Figure 15 demonstrates the 3 meta-states identified in run 2. Notably, meta-state 3 in run 2 exhibits a globally shared temporal pattern across AU_6 (c15), DM_43 (c9), DM_45 (c12), SM_15 (c3), and VI_17 (c5). It's worth noticing that the dFNC states of sICA also revealed 2 globally shared temporal patterns.

The group-level mean temporal values of these meta-states are plotted in Figure 14 (run 1) and Figure 16 (run 2) with researcher-labeled scene boundaries (red vertical lines). The meta-states show rapid and high-frequency fluctuations throughout the scans, showing rapid switching across meta-states. With a preliminary observation, it seems like scene boundaries are likely to be associated with sudden meta-state shifts. However, given the high frequency of meta-state shifting, this can be a consequence of coincidence and would require formal investigation. Overall, the meta-states found through TFM are less stable and have a shorter temporal scale compared to the states identified via sICA dFNC.

Figure 13

Weights of sICA components in tICA components of run 1

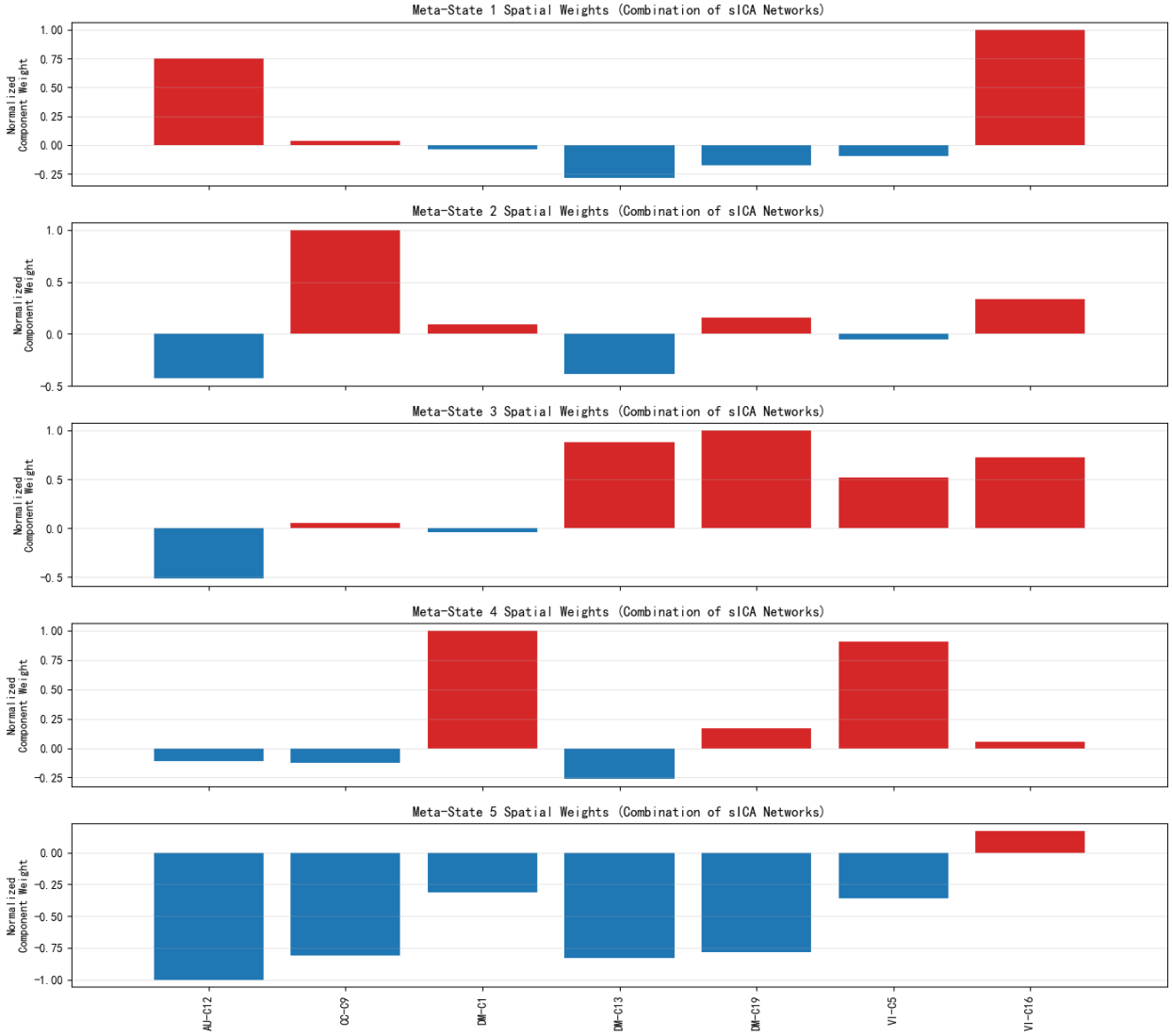


Figure 14

Temporal dynamics of meta-states for run 1

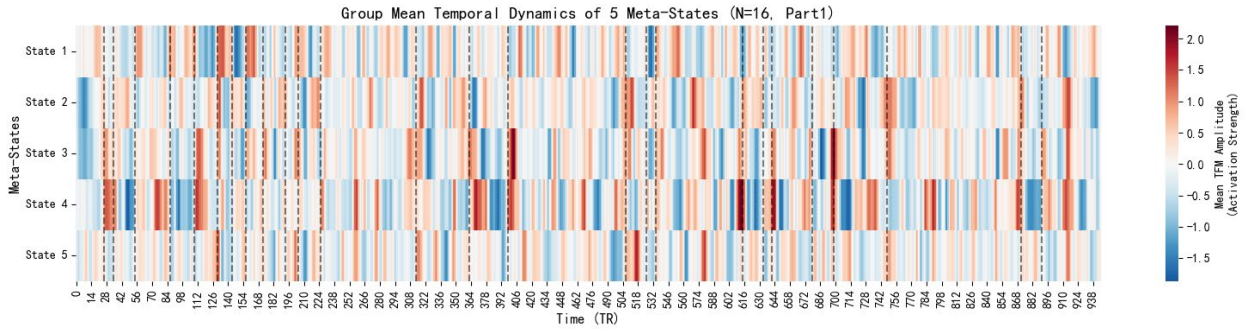


Figure 15

Weights of sICA components in tICA components of run 2

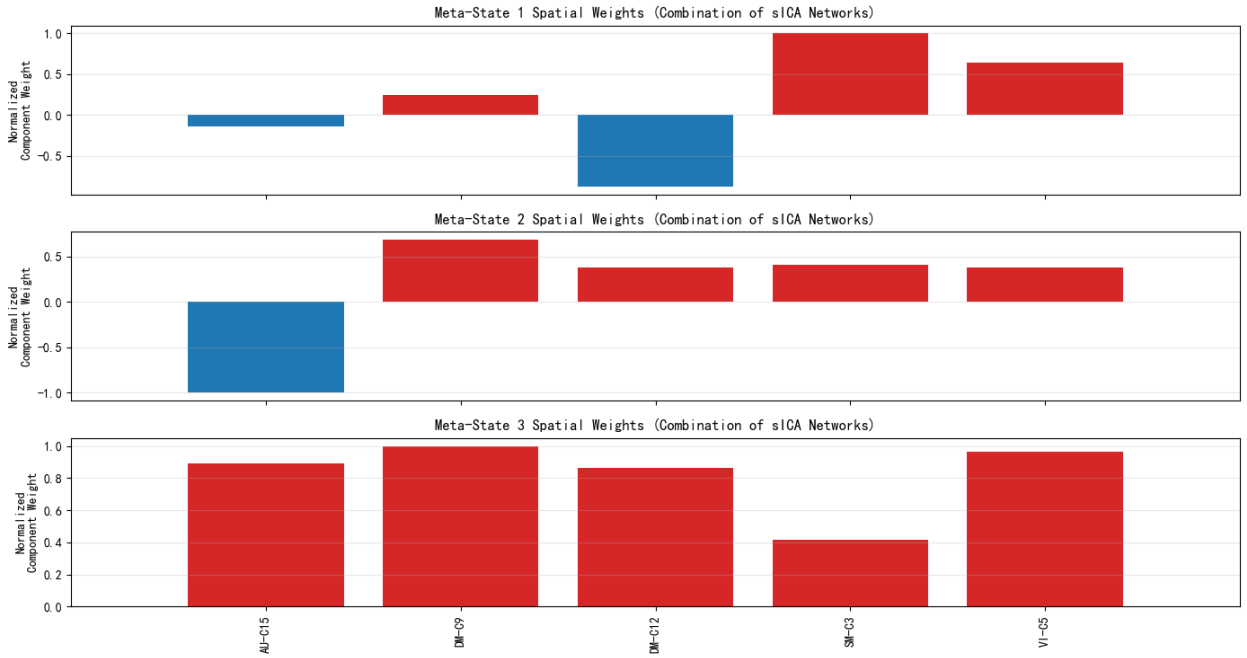
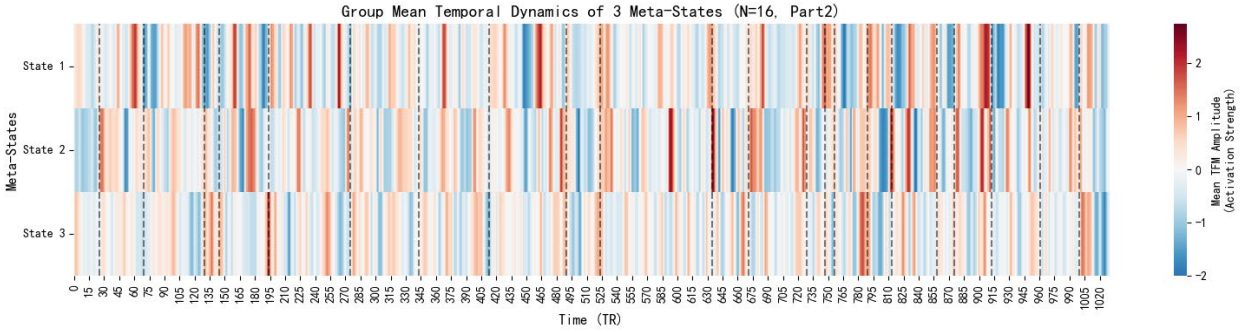


Figure 16

Temporal dynamics of meta-states for run 2



Discussion

This exploratory study aimed to evaluate the spatial and temporal relationship between sICA and tICA in the context of film watching or naturalistic narrative comprehension. Based on the results, the continuous and highly synchronized nature of film viewing allowed both methods to capture meaningful, overlapping spatial and temporal patterns at some level. Overall, the results indicate that while tICA and sICA share some spatial and temporal patterns, sICA provides higher interpretability and more stable network dynamics.

The difference in interpretability is likely a consequence of both data features and template biases. Although using longer naturalistic stimuli, the available amount of data on the time dimension is still less compared to the amount of data on the spatial dimension. With a 1.5s TR and a scan around 25 minutes long, there are only around 1000 data points on temporal dimension. In contrast, the amount of voxel is around 150k. As such, the available data point on spatial dimension is about 150 times the temporal dimension. Which could lead to relatively low quality components for tICA compared to sICA. Furthermore, sICA matched the Neuromark templates significantly better than tICA. This is expected, given that these templates were derived from group-level sICA and naturally optimize for spatial independence. Interestingly, despite these constraints, our spatial comparisons demonstrated that a single tICA component often shares substantial overlap with an aggregation of several sICA components. This suggests that because tICA requires temporal independence more than spatial independence, it can capture combined networks that involve multiple spatially independent but temporally dependent regions.

When evaluating network connectivity state transitions, a comparison between sICA dFNC states and TFM meta-states showed some differences. The probability dynamics of sICA dFNC states appeared to be more stable and showed a visual alignment with researcher-labeled scene boundaries. In contrast, the meta-states identified by TFM showed rapid, high-frequency fluctuations and shorter durations. These differences in temporal scale suggest that TFM (which concatenates sICA time courses for a tICA) might be more sensitive to lower-level processes. On the other hand, sICA dFNC states captured states with relatively larger temporal scales, which might be associated with high-level event comprehension. However, this would require

further formal analysis. In addition, there exist other ICA-based methods that can capture both spatial and temporal patterns simultaneously. For instance, the Tensor Independent Component Analysis (Tensor CA; Beckmann & Smith, 2005). This method directly aggregates group-level data into three-dimensional tensors (time, space, subjects) as the input of ICA and balances spatial and temporal independence. Although the present project does not include Tensor ICA due to time limits, it should be explored in the future.

The present project has multiple limitations that are worth consideration for future research. First, the observed alignment between sICA state transitions and scene boundaries remains subjective. A formal and strict future analysis should utilize Generalized Linear Mixed Models (GLMMs), treating the scene boundary as the dependent variable and the state transition distance as the independent variable, to formally quantify the relation between state shifting and scene boundary. Second, scene boundaries are merely labeled based on shifting of spatial information. In event cognition and narrative comprehension, event boundaries are generated based on shifts of situational information beyond just spatial information, and the process is tied to subjective and endogenous cognitive processes rather than simple visual cuts. As mentioned previously, the Event Indexing Model suggests event boundaries are associated with shifts of multiple situational dimensions (Zwaan et al., 1995b). Therefore, event boundary data would be important to better understand the relationship between connectivity state shifting and cognitive processes. Additionally, the network resolution in this study was constrained by the use of a relatively coarse template (Neuromark 1.0) and a limited component number (k=20). Implementing ICA with more components and updated templates such as Neuromark 3.0 could capture more nuanced states and dynamics of the brain activity during film watching.

Finally, while exploring the temporal structure of brain states and its relationship to event boundary is important, narrative comprehension and event cognition is fundamentally driven by content, such as the retrieval of event schemas, emotional valence, and semantic knowledge. For instance, Event Segmentation Theory suggests that detection of event boundaries is directly related to prediction error, which would require a generative model to predict, which is equivalent to mental model (Zacks et al., 2011). Understanding the content of the mental models would be useful for better understanding how narrative comprehension and event cognition are achieved. Complex cognitive processes like utilization of event schema require temporal directionality that simple association matrices cannot fully capture. In the future, it is worth considering integrating structural data through multimodal methods with time-respecting models, such as Hidden Markov Models (HMM) or Greedy State Boundary Search (GSBS) (Baldassano et al., 2017; Geerligs et al., 2021). Moreover, to transition from structural boundary detection to true content decoding, incorporating neural networks or large language models (LLMs) offers a possible path. These models could be used as semantic dictionaries to extract hierarchical event schemas or semantic concepts from narrative stimuli, which can then be mapped onto the spatiotemporal networks identified by ICA through methods like representational similarity analysis. This could help bridge the gap between data-driven neural patterns and cognitive theories of event comprehension.

References

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, *95*(3), 709–721.e5. https://doi.org/10.1016/j.neuron.2017.06.041

Beckmann, C. F., & Smith, S. M. (2005). Tensorial extensions of independent component analysis for multisubject FMRI analysis. *NeuroImage*, *25*(1), 294–311. https://doi.org/10.1016/j.neuroimage.2004.10.043

Calhoun, V. D., Adali, T., Pearlson, G., and Pekar, J. (2001). "Group ICA of functional MRI data: separability, stationarity, and inference," in Proceeedings of the International Conference on ICA and BSS (San Diego, CA), 155.

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature neuroscience*, *20*(1), 115–125. https://doi.org/10.1038/nn.4450

Du, Y., Fu, Z., Sui, J., Gao, S., Xing, Y., Lin, D., Salman, M., Abrol, A., Rahaman, M. A., Chen, J., Hong, L. E., Kochunov, P., Osuch, E. A., Calhoun, V. D., & Alzheimer's Disease Neuroimaging Initiative (2020). NeuroMark: An automated and adaptive ICA based pipeline to identify reproducible fMRI markers of brain disorders. *NeuroImage. Clinical*, *28*, 102375. https://doi.org/10.1016/j.nicl.2020.102375

Geerligs, L., van Gerven, M., & Güçlü, U. (2021). Detecting neural state transitions underlying event segmentation. *NeuroImage*, *236*, 118085. https://doi.org/10.1016/j.neuroimage.2021.118085

Gernsbacher, M.A. (1990). Language Comprehension As Structure Building (1st ed.). Psychology Press. https://doi.org/10.4324/9780203772157

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science (New York, N.Y.)*, *303*(5664), 1634–1640. https://doi.org/10.1126/science.1089506

Loschky, L. C., Larson, A. M., Magliano, J. P., & Smith, T. J. (2015). What Would Jaws Do? The Tyranny of Film and the Relationship between Gaze and Higher-Level Narrative Film Comprehension. *PloS one*, *10*(11), e0142474. https://doi.org/10.1371/journal.pone.0142474

Magliano, J. P., Miller, J., & Zwaan, R. A. (2001). Indexing space and time in film understanding. *Applied Cognitive Psychology, 15*(5), 533–545. https://doi.org/10.1002/acp.724

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L., & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature neuroscience*, *4*(6), 651–655. https://doi.org/10.1038/88486

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction error associated with the perceptual segmentation of naturalistic events. *Journal of cognitive neuroscience*, *23*(12), 4057–4066. https://doi.org/10.1162/jocn_a_00078

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995b). The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. *Psychological Science*, *6*(5), 292–297. http://www.jstor.org/stable/40063035

Zwaan, R.A., Magliano, J.P., & Graesser, A.C. (1995a). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 386-397.

Appendix

Code Repository: https://github.com/angushushu/PSYC8690