

A nested cortical hierarchy of neural states underlies event segmentation in the human brain

Linda Geerligs¹, Marcel van Gerven¹, Karen L. Campbell² and Umut Güçlü¹

1. Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands

2. Department of Psychology, Brock University, St. Catharines, Canada

Correspondence to:

Linda Geerligs

Donders Institute for Brain, Cognition and Behaviour

Thomas van Aquinostraat 4, 6525 GD Nijmegen, The Netherlands

+31 243616091

lindageerligs@gmail.com

Keywords: Event segmentation, naturalistic viewing, fMRI, greedy state boundary search, timescales, functional networks, hierarchy

Abstract

A fundamental aspect of human experience is that it is segmented into discrete events. This may be underpinned by transitions between distinct neural states. Using an innovative data-driven state segmentation method, we investigate how neural states are organized across the cortical hierarchy and where in cortex neural state and perceived event boundaries overlap. Our results show that neural state boundaries are organized in a temporal cortical hierarchy, with short states in primary sensory regions and long states in anterior temporal pole and lateral and medial prefrontal cortex. Neural state boundaries overlap with event boundaries across large parts of this hierarchy. State boundaries are shared within and between groups of brain regions that resemble well known functional networks, such as the default mode network that fractionates into two subnetworks – one fast, one slow. Together these findings suggest that a nested cortical hierarchy of neural states forms the basis of event segmentation.

Introduction

Segmentation of information into meaningful units is a fundamental feature of our conscious experience in real life contexts. Spatial information processing is characterized by chunking spatial regions into objects (e.g. DiCarlo and Cox, 2007). In a similar way, temporal information processing is characterized by segmenting our ongoing experience into separate events (Kurby and Zacks, 2008; Newton et al., 1977). Segmentation improves our understanding of ongoing perceptual input (Zacks et al., 2001a) and allows us to recall distinct event from our past (Flores et al., 2017; Sargent et al., 2013; Zacks et al., 2006). Recent work has shown that the end of an event triggers an evoked response in the hippocampus (Baldassano et al., 2017; Ben-Yakov and Henson, 2018), suggesting that events form the basis of long-term memory representations. Events that are identified in movie or auditory narratives are often very similar across individuals and can be segmented hierarchically on different timescales (Newton and Rindner, 1979; Zacks et al., 2001a). According to event segmentation theory (EST), perceptual systems spontaneously segment activity into meaningful events as a side-effect of predicting future information (Zacks et al., 2007). That is, event boundaries are perceived when predictions become less accurate, which can be due to a change in motion or features of the situation such as characters, causes, goals and spatial location (Zacks et al., 2009).

While much is known about temporal event segmentation at a behavioral level, less is known about its neural underpinnings. A number of studies have investigated which brain regions show evoked responses around event boundaries. Although the exact regions vary across studies, commonly identified regions include the precuneus and medial visual cortex, as well as area V5 and the intraparietal sulcus (Kurby and Zacks, 2018; Speer et al., 2007; Zacks et al., 2001b; Zacks et al., 2010). Increased brain responses at event boundaries in these regions likely reflect updating processes that occur when shifting to a new event model (Ezzyat and Davachi, 2011). Recently a different approach has been introduced to investigate the neural underpinnings of event segmentation (Baldassano et al., 2017). They applied a data-driven method, based on Hidden Markov Models (HMMs) to functional magnetic resonance imaging (fMRI) data obtained during movie watching to identify time points where brain activity in a particular region transitioned from one temporarily stable activity pattern to a different pattern. Here, we will refer to these periods of relative stability as neural states to distinguish them from subjectively experienced events (Zacks et al., 2007).

These neural states occur on different timescales across the cortical hierarchy, with short lived states in early sensory regions and long lasting states in the precuneus and angular gyrus (Baldassano et al., 2017), in line with previous observations of a temporal hierarchy of information processing in the brain (Hasson et al., 2008; Honey et al., 2012; Lerner et al., 2011; Stephens et al., 2013). Interestingly, for a set of four brain regions, Baldassano and colleagues (2017) showed that neural state boundaries overlap across different regions as well as with subjectively experienced event boundaries. These results suggest that neural state segmentation could be the source of perceived event boundaries and that states may be organized in a nested cortical hierarchy, such that the boundaries of faster states in regions lower in the cortical hierarchy are nested within the boundaries of slower regions higher up the hierarchy. The idea is that each brain region integrates information within discretized neural states that may align with sensory units in the external environment (e.g., phonemes, words, sentences) and provide its output to the brain regions higher in the cortical hierarchy (Nelson et al., 2017), until neural states at the highest level of the hierarchy align with subjectively experienced events. This way information traveling up the hierarchy is gradually integrated into complex and long-lasting

multimodal representations (Hasson et al., 2015). Although this is an intriguing hypothesis, the evidence for it is limited, as it has only been investigated in one previous study using four pre-defined regions of interest (Baldassano et al., 2017). In addition, it remains unknown which brain regions show neural state boundaries that align with perceived event boundaries and what the temporal hierarchy of state segmentation looks like across large parts of the cortex that remain unexplored.

The aim of the current study was to investigate whether event segmentation is indeed underpinned by neural state segmentation occurring in a nested cortical hierarchy and to characterize the temporal hierarchy of neural state segmentation across the entire cortex. If the brain segments ongoing input in a nested hierarchical fashion, we would expect to find especially long lasting neural states in the frontal cortex, which can be considered to be at the top of the cortical hierarchy (Fuster, 2001). We would also expect to find overlap between neural state boundaries and event boundaries across all levels of the cortical hierarchy, although this overlap should be strongest for areas at higher levels of the hierarchy where the timescales of neural state segmentation should closely match the experienced timescale of events. Finally, we would expect that state boundaries are most strongly shared between groups of brain regions involved in similar cognitive functions and to a lesser extent between more distinct sets of brain areas as some boundaries are expected to traverse the entire cortical hierarchy.

To test these hypotheses we used a novel data-driven state segmentation method, that has been shown to be a more valid and reliable method for neural state segmentation than the previously used HMM-based method (Geerligts et al., 2020). By using a unique movie fMRI dataset from the Cam-CAN project (Shafto et al., 2014) that shows reliable stimulus-driven activity (i.e., significant inter-subject correlations) over nearly all cortical brain regions (Geerligts et al., 2018), we were able to study neural state segmentation across the entire cortex for the first time.

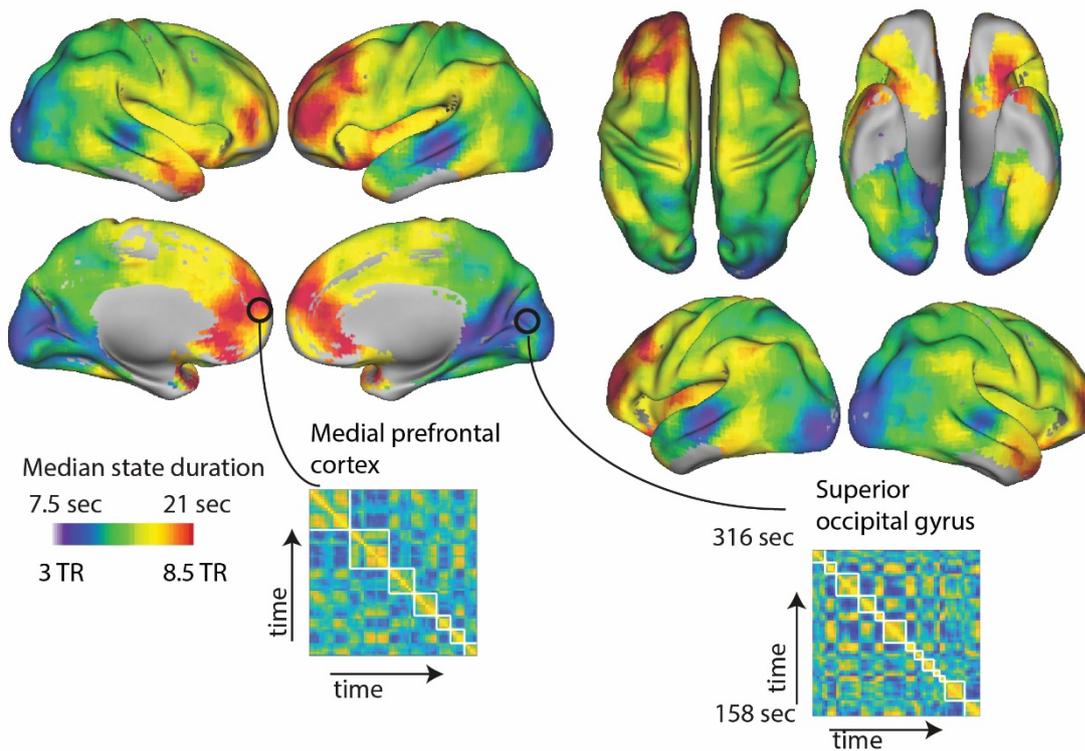
Results

To identify neural state boundaries, we applied greedy state boundary search (GSBS; Geerligs et al., 2020) to a large fMRI dataset in which 265 participants were viewing an 8-minute Alfred Hitchcock movie (Shafto et al., 2014). After hyperaligning the data (Guntupalli et al., 2016), GSBS was applied to multi-voxel brain activity time courses from overlapping spherical searchlights covering the entire cortex. GSBS provides a set of neural state boundaries for each searchlight and for different numbers of states (K). The t-distance metric was used to identify the optimal number of state boundaries in each searchlight (Geerligs et al., 2020). This metric identifies the optimal number of state boundaries such that the correlations of time points within a state are maximized and correlations of time points in consecutive states are minimized.

The median duration of neural states differed greatly between searchlights, ranging from 7 seconds in the regions with fast transitions up to 25 seconds in the slowest regions (see figure 1A). Most searchlights showed median state durations between 15 and 17 seconds per state. However, we also observed that even regions with a slow timescale can show fast state transitions in some cases (see figure 1B). To determine whether regional differences in state duration were reliable, neural state boundaries were identified in 15 independent groups of participants. Across these 15 groups, there was a medium to large correlation between the estimates of median state duration across searchlights (mean $r=0.46$). To further examine the reliability of regional variations in state duration, we tested whether each searchlight's median state duration deviated significantly from the median state duration across all searchlights. The results show that there is reliable variability in the duration of states across the cortex (see figure S1).

Figure 1A shows that there were particularly fast transitions between stable activity patterns in early visual cortex and early auditory cortex. State transitions were slower in areas further up the ventral visual stream, such as the inferior occipital gyrus and the fusiform gyrus, as well as up the dorsal visual stream in the intraparietal sulcus. Particularly slow state transitions were observed in high level-regions such as the medial prefrontal gyrus and the temporal pole and the left insula and the left middle frontal gyrus. For the subset of regions investigated in previous work, we observe a strong correspondence in our findings (Baldassano et al., 2017; Hasson et al., 2008; Lerner et al., 2011).

A. Median state durations across the cortical hierarchy



B. Distribution of state durations for groups of regions with different timescales

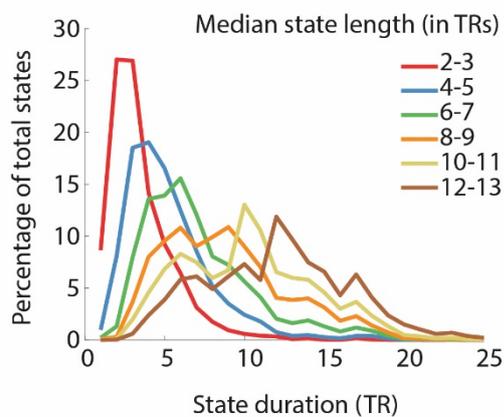


Figure 1. A. The optimal number of states varied greatly across regions, with many shorter states in the primary visual and auditory cortices and few longer states in the association cortex, such as the medial prefrontal gyrus and temporal pole. Parts of the correlation matrices for two selected searchlights are shown in the insets, representing approximately 2.5 minutes of the movie. The white lines in these insets are the neural state boundaries that are detected by GSBS. B. Shows the distribution of state durations for regions grouped according to the median state duration in each region. In panel A, we show the median state duration across all 15 independent participant groups, while panel B shows all observations across the 15 groups.

Neural states and event boundaries

If neural states are organized in a nested cortical hierarchy, with some of the state boundaries at lower levels reoccurring throughout the hierarchy and reaching conscious awareness, we would expect to see that state boundaries align with event boundaries across all of the different levels of the hierarchy. However, the overlap would be more pronounced in higher-level cortical areas, where the number of states should more closely align with the number of perceived events. Event boundaries were determined by asking participants to indicate when they felt one event (meaningful unit) ended and another began (Ben-Yakov and Henson, 2018). To determine the similarity between neural state boundaries and event boundaries, we estimated the correlation between the boundary vectors for different delays to account for the delay in the hemodynamic response. These boundary vectors contained zeros when there was no change in state/event and ones when there was a transition to a different state/event. For each searchlight, the correlation for the optimal delay was contrasted with the expected correlation based on randomized event boundaries, resulting in an adjusted correlation metric that is zero when the correlation is equal to the expected correlation and one when the event and neural state boundary vectors were identical.

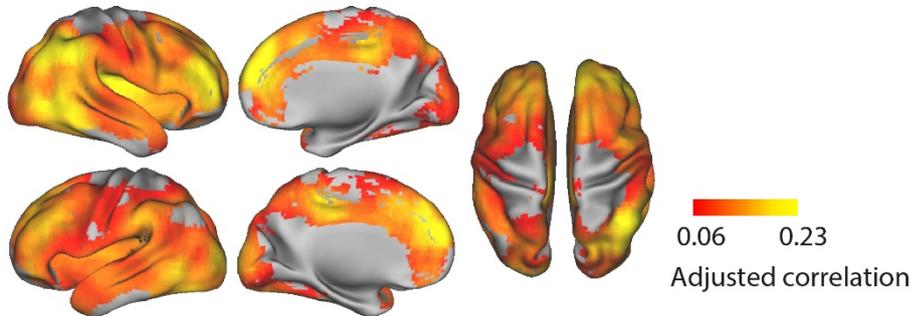
We found that a large number of brain regions, throughout the cortical hierarchy, showed significant alignment between neural state boundaries and event boundaries after false-discovery-rate (FDR) correction for multiple comparisons (see figure 2A; Benjamini and Yekutieli, 2011). This was most prominent in the right middle temporal visual areas (area V5), the right middle frontal gyrus and most areas of the default mode network (DMN), including the medial frontal gyrus, anterior parts of the precuneus and right angular gyrus. To determine whether regional differences in the overlap between event and state boundaries were a side-effect of regional differences in the optimal number of states, we repeated the analyses using a fixed number of states that was the same as the number of subjective events ($k=19$, see figure S2A). The results were highly similar to those in figure 2A, suggesting that regional differences were due to the timing and not just the number of state boundaries in these regions.

While estimating the association between neural state boundaries and event boundaries, we also identified the delay between the neural state boundaries and event boundaries within each searchlight. These delay differences across searchlights were similar to BOLD signal delays described previously (Spearman $r=0.43$; see figure S2B; Amemiya et al., 2020), further supporting our interpretation that the neural state boundaries across all these brain regions are associated with the experience of an event boundary.

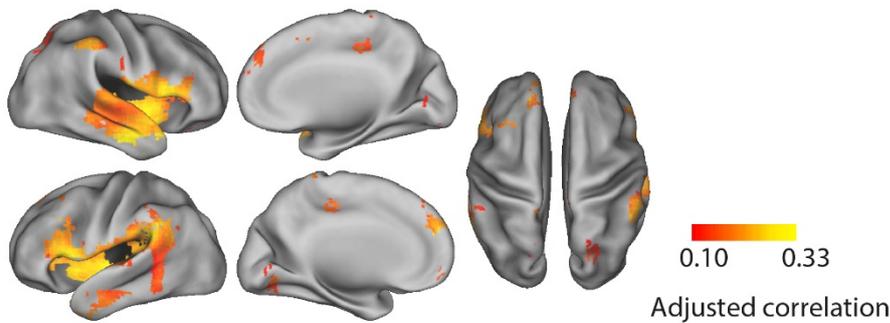
To determine what dissociates neural state boundaries that are not associated with an event boundary from those that are, we investigated two boundary properties; boundary strength and boundary co-occurrence. Boundary strength is defined as the correlation distance between the neural activity patterns of consecutive neural states. Boundary co-occurrence is defined as the proportion of searchlights that show a boundary at the same time. We found that stronger boundaries, so boundaries that coincided with a larger shift in brain activity patterns, were more often associated with an experienced event boundary, particularly in the anterior parts of the temporal gyrus bilaterally and the left inferior frontal gyrus (see figure 2B). Boundary co-occurrence was also an important determinant of experienced event boundaries. The results in figure 2C suggest that at some points in time, state boundaries were shared between large numbers of brain regions across the temporal hierarchy. Those state boundaries that co-occured across many brain regions were much more likely

to be associated with a perceived event boundary than state boundaries that only occurred in a few brain regions (average adjusted correlation = 0.27, average unadjusted correlation = 0.34, $p < 0.001$).

A. Association between neural state boundaries and event boundaries



B. Association between neural state boundary strength and event boundaries



C. Relation between state boundary co-occurrence and subjective event boundaries

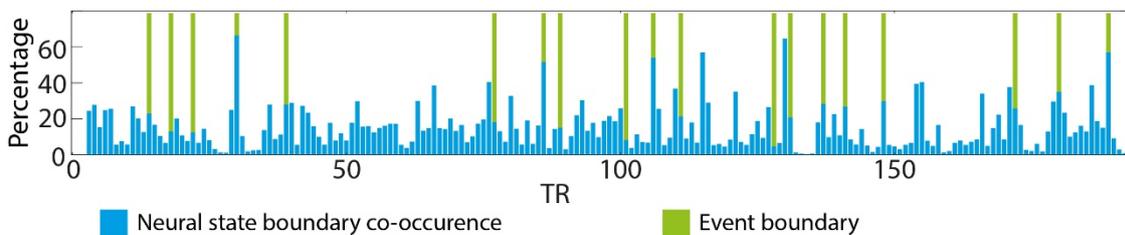


Figure 2. A. Regions across large parts of the hierarchy show a significant association between the neural state boundaries and the experienced event boundaries identified by a different group of participants outside the scanner. B. Association between the strength of the neural state boundaries and event boundaries. Regions with strong neural state boundaries (i.e., a large change between successive states) were more likely to overlap with perceived event boundaries than weak boundaries. In A and B, all of the colored regions showed a significant association after FDR correction for multiple comparisons. C. State boundaries that co-occurred across many brain regions were much more likely to be associated with a perceived event boundary than state boundaries that only occurred in a few brain regions. Neural state boundary co-occurrence and event boundaries are shown over the entire time course of the movie. Event boundaries were shifted to account for a six second HRF delay. The data shown in panels A-C were averaged across all 15 independent groups of participants.

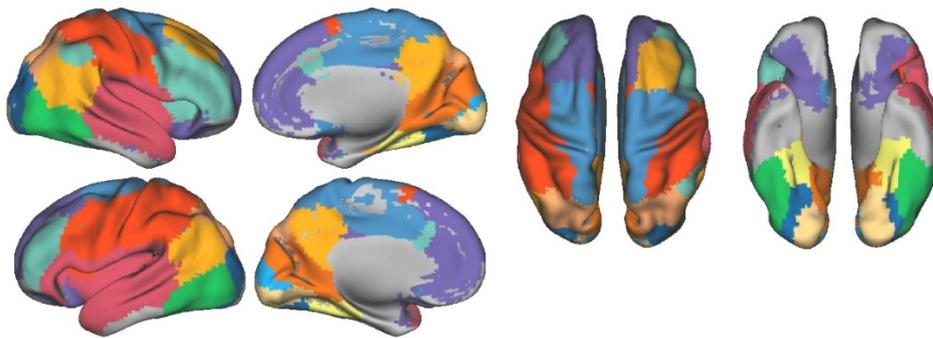
Neural state networks

Because neural state boundary co-occurrence seems to be an important factor in driving the overlap between neural states and event boundaries, we investigated how neural state boundaries are shared across regions. We identified networks of brain regions that shared state boundaries, by computing the Pearson correlation between the boundary vectors of each searchlight and using consensus partitioning based on Louvain modularity maximization to identify networks (Blondel et al., 2008; Lancichinetti and Fortunato, 2012). We found that state boundaries were shared within long-range networks. Some of these networks resembled the functional networks that are typically identified based on (resting state) timeseries correlations (see figure 3A). To quantify this, we computed the proportion of searchlights overlapping with each of the networks defined by Power et al. (2011, see table S1). We identified an auditory network, that extended into regions involved in language processing in the left inferior frontal gyrus, a fronto-parietal task control network (FPCN) and a default mode network (DMN), which was fractionated into an anterior and a posterior component. The somatomotor network (SMN) we identified was split into a lateral and medial component. We also identified a network overlapping with the dorsal attention network (DAN), although the network we identified covered the posterior parts of the DAN but not the frontal eye fields. While the visual network is typically identified as a single network in functional connectivity studies, we observed a set of six fine-grained networks, roughly corresponding to different levels of the visual hierarchy, one of which overlapped with the fusiform gyrus network also present in the data from Power et al. (2011). Figure 3 visualizes for each voxel, which functional network label occurs most frequently for the searchlights overlapping that voxel. In contrast, the full extent of each of the functional networks can be seen in figure 4.

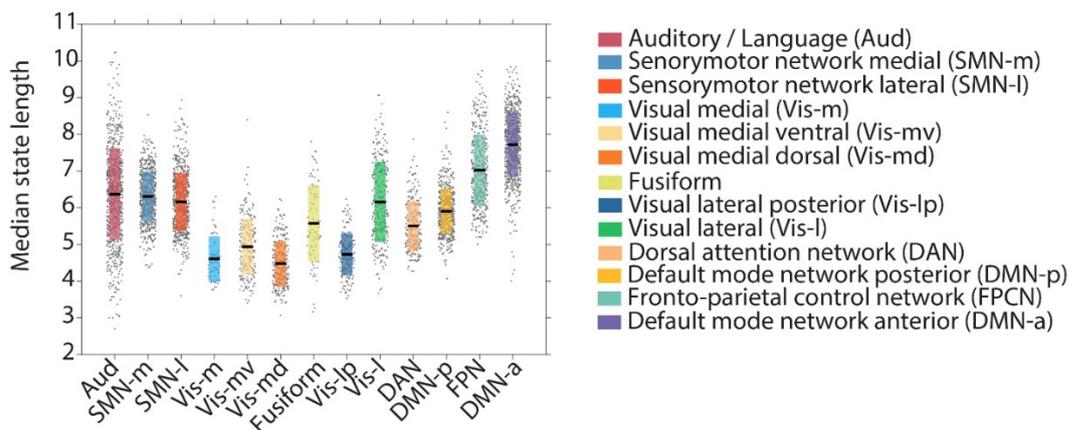
Figure 3B shows the average timescales within each of these function networks. The networks with the longest durations were the anterior DMN and the FPCN, while some of the lower-level visual networks had particularly short durations. Although regions within functional networks tended to operate on a similar temporal scale, we also observed a lot of variability in state duration within networks, particularly in the auditory network (see figure 3C). This suggests that the correlations between state boundary vectors were not simply driven by a similarity in the number of states, but rather by a similarity in the state boundary timings. This is also supported by the results in figure S3A and C, showing that the correlation structure of boundary vectors was quite distinct from the pairwise differences in median state duration ($r=-0.29$).

Although the networks we identified show overlap with previously identified functional networks based on correlations between brain activity timecourses across voxels or regions, they clearly diverged for various networks (e.g. the visual network). This divergence was supported by the limited similarity with the previously identified networks by Power et al. (2011; adjusted mutual information = 0.41), as well as the differences between the correlation matrix that was computed based on the mean activity time courses in each searchlight (figure S3B) and the correlation matrix that was computed based on the boundary vectors in each searchlight (figure S3A; $r=0.43$). Regions with strongly negatively correlated mean activity time courses typically showed weak or positive correlations between boundary vectors.

A. Functional network parcellation



B. Timescales of searchlights within each network



C. Graph visualization of partial correlations between networks

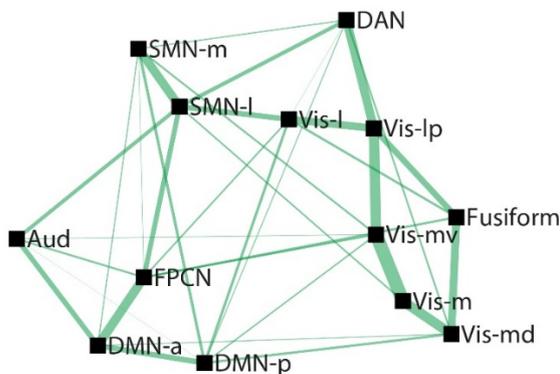


Figure 3. Neural state boundaries are shared within and across distinct functional networks that span the cortical hierarchy A. Visualization of the detected functional networks. The network label at each voxel is determined by the functional network that occurs most often in all the searchlights that overlap with that voxel. B. Visualization of the neural state durations within each network. Each searchlight is shown as a dot. The bars show the mean and standard deviation around the mean for each network. C. A graph visualization of the partial correlation between the averaged boundary vectors in the different functional networks is shown (controlling for the boundary vectors in all other networks). The organization of the graph shows a clear correspondence to the known hierarchical organization of the cortex. The thickness of the line corresponds to the strength of the partial correlation. All lines are significant after FDR-correction for multiple comparisons. The data shown in panels B-C were averaged across all 15 independent groups of participants.

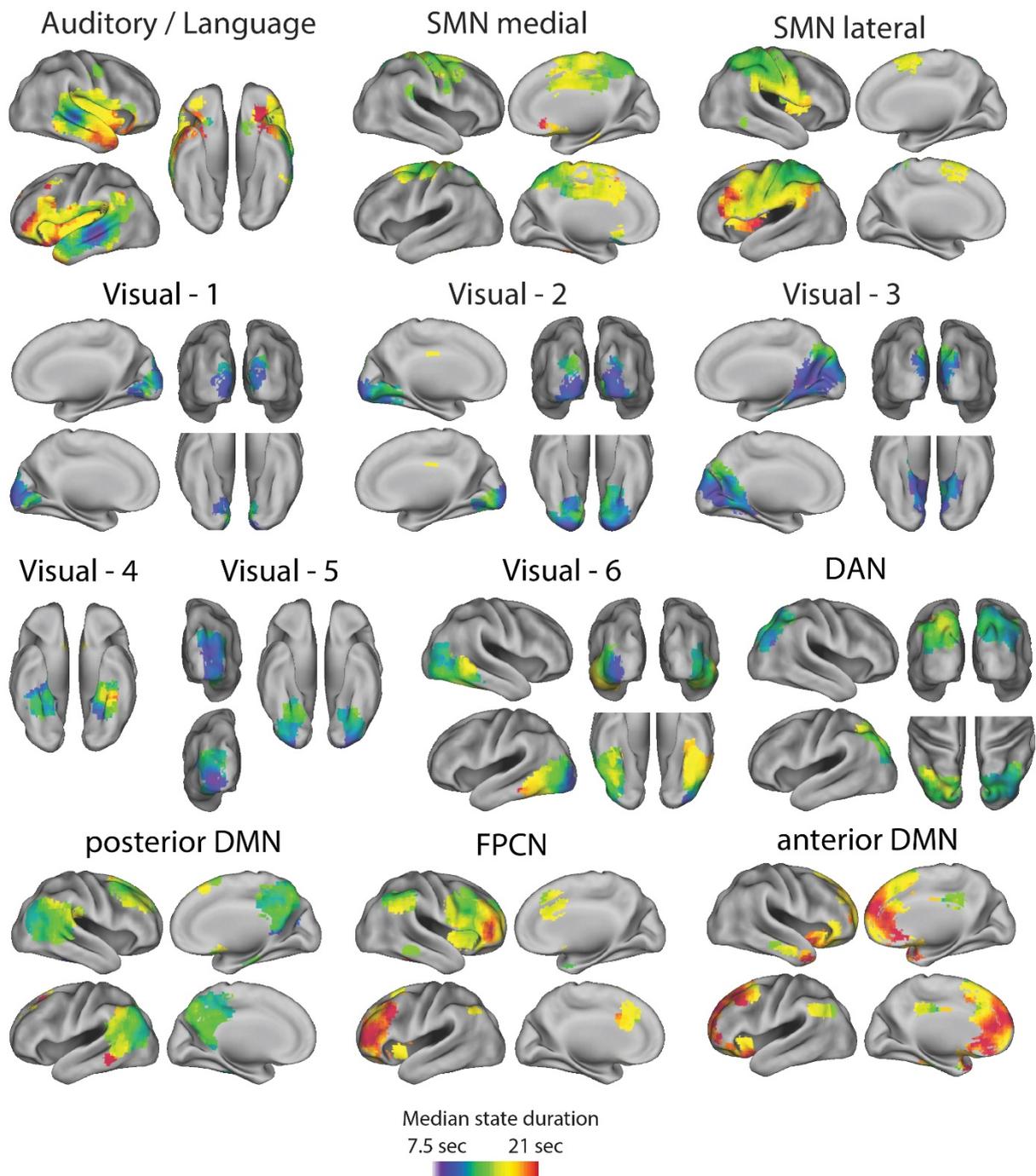


Figure 4: Separate visualizations for each of the identified functional networks, showing the optimal number of states for each of the searchlights within the functional network. SMN=somato-motor network, DMN=default mode network, FPCN=fronto-parietal control network, DAN=dorsal attention network. The figure shows the median state duration estimates across all 15 groups of participants.

If state boundaries form a nested cortical hierarchy, we expect that they are shared not only within brain regions involved in similar functions but also across different functional networks. To investigate this, we computed partial correlations between the average boundary vectors in each network, controlling for the boundary vectors in all other networks. Using partial, rather than regular Pearson

correlations allowed us to focus on direct connections between networks (Smith et al., 2011). We found evidence that neural state boundaries are indeed shared across different networks, in a way that closely aligns with the hierarchical organization of these different networks (see figure 3C). For example, the different visual networks and the dorsal attention network tended to share boundaries, as well as the higher order networks (anterior and posterior DMN, FPCN) and the sensorimotor networks. There were also many significant partial correlations between each of these networks. The lower-level visual networks shared boundaries with the higher-level visual networks, which in turn connected to the FPCN and the DAN, as well as the posterior DMN. The FPCN and posterior DMN also received input from the auditory network and were connected to each other, as well as to the anterior DMN.

Discussion

While event segmentation is a critical aspect of our ongoing experience, the neural mechanisms that underlie this ability are not yet clear. The aim of this paper was to investigate the cortical organization of neural states that may underlie our experience of distinct events. By combining an innovative data-driven state segmentation method, with a movie dataset of many participants, we were able to identify neural states across the entire cortical hierarchy for the first time. We observed particularly fast states in primary sensory regions and long periods of information integration in the left middle frontal gyrus and medial prefrontal cortex. Across the entire cortical hierarchy, we observed associations between neural state and event boundaries and our findings demonstrate that neural state boundaries are shared within long-range functional networks as well as across the temporal hierarchy between distinct functional networks.

A nested cortical hierarchy of neural states

Previous findings have suggested that neural states may be organized in a nested cortical hierarchy (Baldassano et al., 2017). In line with this hypothesis, we observed that neural state boundaries throughout the entire hierarchy overlap with experienced event boundaries, but this overlap is greater for transmodal regions such as the medial prefrontal cortex and angular gyrus that may be involved in constructing ‘situation models’ of an ongoing narrative (Ranganath and Ritchey, 2012; Zwaan and Radvansky, 1998). This finding suggests that some of the neural state boundaries that can be identified in early sensory regions are represented throughout the cortical hierarchy until they are eventually reflected in conscious experience. It also suggests that event segmentation may not be underpinned by a central mechanism that registers the prediction error of the current input in relation to a single event model (Zacks et al., 2007; Zacks et al., 2011), but rather by a local segmentation mechanism that occurs at all levels of the cortical hierarchy.

So what do these neural states represent? Recent work by Chien and Honey (2020) has shown that neural activity around an artificially introduced event boundary can be effectively modeled by ongoing information integration, which is reset by a gating mechanism, very much in line with mechanism that have been proposed to underlie event segmentation (Kurby and Zacks, 2008). Similarly, neural states may represent information integration about a particular stable feature of the environment, which is reset when that feature undergoes a substantial change. This suggests that neural states in early visual cortex may represent short-lived visual features of the external environment, while states in anterior temporal cortex may contain high-level semantic representations related to the ongoing narrative (Clarke and Tyler, 2015). For transmodal regions such as the medial prefrontal cortex, or middle frontal gyrus, that have been associated with many different high level cognitive processes (e.g. Duncan, 2010; van Kesteren et al., 2012; Simony et al., 2016), it is not yet clear what a distinct neural state might represent. Just as event boundaries can be related to changes in one or multiple situational dimensions, such as changes in goals or locations (Clewett et al., 2019; Zacks et al., 2009), neural state boundaries in transmodal cortical areas may not necessarily reflect one particular type of change. State boundaries in these regions are likely also dependent on the goals of the viewer (Wen et al., 2020).

We also investigated the factors that distinguish neural state boundaries that traverse the hierarchy from those that do not. It has previously been shown that changes across multiple aspects of the narrative are more likely to result in an experienced event boundary (Zacks et al., 2010). In line with

this, we observed that boundaries that were represented in more brain regions at the same time were also more likely to be associated with the experience of an event boundary. The strength of the neural state boundary, as measured by the amount of change in neural activity patterns, was also identified as a factor that can to some degree distinguish neural states that travel up the hierarchy and appear in subjective experience from the neural states that do not, particularly in temporal cortex and inferior frontal gyrus. This suggests that a neural state boundary is not an all or none occurrence. Instead, the reset of representations at neural state boundaries (Chien and Honey, 2020) may differ based on what is happening in other brain regions or based on the degree of change in the representations of the environment in that particular brain region.

More evidence for the idea of a nested cortical hierarchy of neural state boundaries comes from our connectivity analyses, which show that neural state boundaries are shared both within and across groups of regions that partly resemble well known functional brain networks. The way neural state boundaries are shared between networks closely aligns with the interconnected functions between those particular networks. For example, different low-level visual networks shared state boundaries with each other but also with higher level networks, such as the DAN and FPCN. This sharing of boundaries across different cortical areas may suggest that neural states in higher level cortical regions represent an overarching representation that corresponds to many distinct states in lower level cortical areas, which all represent different features of that overarching representation (e.g. words spoken, characters on screen or locations within a particular situation). This is in line with previous conceptualizations of events as partonomic hierarchies (Zacks et al., 2001a) and with other models of hierarchical neural representations, such as the hub-and-spokes model for semantic representations, which proposes that semantic knowledge is represented by the interaction between modality-specific brain regions and a transmodal semantic representational hub in the anterior temporal lobe (Lambon Ralph et al., 2010; Rogers et al., 2004). It is also in line with a recently proposed hierarchical representation of episodic memories, in which items that are linked within small-scale events are in turn linked within large-scale episodic narratives (Andermane et al., 2021).

Timescales of information processing across the cortex

While previous studies have been able to show regional differences in the time scale of information processing across part of the cortex (Baldassano et al., 2017; Hasson et al., 2008; Honey et al., 2012; Lerner et al., 2011; Stephens et al., 2013), here we were able to reveal neural state timescales across the entire cortex for the first time. The validity of our results is supported by extensive validations using simulations (Geerligs et al., 2020) and the reliability of our observations across independent groups of participants. It is also supported by the striking similarity between our results and previous findings based on very different approaches, such as experiments with movies and auditory narratives that have been scrambled at different timescales (Hasson et al., 2008; Honey et al., 2012; Lerner et al., 2011), or resting-state fluctuations in electrocorticography (Honey et al., 2012) and functional MRI data (Stephens et al., 2013).

Although we characterized brain areas based on their median state length, we observed that neural states within a region were not of equal duration, suggesting that regional timescales may change dynamically based on the features of the stimulus. This is also in line with the observed correspondence between neural state and event boundaries. Event boundaries have previously been

shown to align with changes in features of the narrative input stimulus, such as characters, causes, goals and spatial locations (Zacks et al., 2009). Therefore, the overlap between state boundaries and event boundaries across the cortex also suggest that characteristics of the sensory input are driving the occurrence of neural state boundaries. Together, these findings show that the timescale of information processing in particular brain regions are not only driven by stable differences in the rate of temporal integration of information, that may be associated with interregional interactions in the neural circuitry (Honey et al., 2012), but also by the properties of the input that is received from the environment. Our results show that some of the areas that were not covered in previous investigations (Baldassano et al., 2017; Hasson et al., 2008; Honey et al., 2012; Lerner et al., 2011; Stephens et al., 2013), such as the medial prefrontal cortex, anterior temporal pole and middle frontal gyrus have the longest timescales of information processing. This suggests these regions at the top of the cortical hierarchy (Clarke and Tyler, 2015; Fuster, 2001) also have the slowest timescales of information processing, in line with expectations based on the hierarchical process memory framework (Hasson et al., 2015).

Functional networks of neural state boundaries

Our results show for the first time that neural state boundaries are shared across brain regions in distinct functional networks. Interestingly, the networks we identify partially resemble the functional networks that are typically found using regular functional connectivity analyses (c.f. Power et al., 2011; Yeo et al., 2011), though there are some differences. For instance, the visual network was segregated into six smaller subnetworks and for other networks, the topographies sometimes deviated somewhat from those observed in prior work.

Our results show that functional networks defined by state boundaries differ in their timescales of information processing. While some networks have a particular temporal mode of information processing, other networks show a broader range of neural state timescales (as indicated by the spread of searchlight timescales in Figure 3B). For the DMN we observed a split into anterior and posterior subnetworks with markedly different timescales, which is similar to previously observed posterior and anterior DMN subnetworks (Andrews-Hanna et al., 2010; Campbell et al., 2013; Lei et al., 2014). The posterior/fast DMN is particularly prominent in the precuneus and angular gyri, which are thought to engage in episodic memory retrieval through connectivity with the hippocampal formation (Andrews-Hanna et al., 2010). The posterior DMN has also been proposed to be involved in forming mental scenes or situation models (Ranganath and Ritchey, 2012). Thus, neural states in this subnetwork may reflect the construction of mental scenes of the movie and/or retrieval of related episodic memories. The anterior/slow DMN is particularly prominent in the medial prefrontal cortex which has been related to self-referential thought, affective processing, and integrating current information with prior knowledge (Benoit et al., 2014; Gilboa and Marlatte, 2017; Van Kesteren et al., 2012; Northoff et al., 2006). The current results suggest that these processes require integration of information over longer timescales.

Real life experience

Although event segmentation is thought to be a pivotal aspect of how information is processed in real life (Zacks et al., 2007), it is often not considered in experimental settings, where events are predetermined by the trial or block structure. This study and previous work (Baldassano et al., 2017) shows that we are now able to investigate brain activity as it unfolds over time without asking participants to perform a task. This allows us to study brain function in a way that is much more similar to our daily life experience than typical cognitive neuroscience experiments (Hamilton and Huth, 2020; Lee et al., 2020; Willems et al., 2020). This opens the door for investigations of neural differences during narrative comprehension between groups of participants, such as participants with autism who may have trouble distinguishing events that require them to infer the state of mind of others (Baron-Cohen, 2000; Hasson et al., 2009a), or participants with Alzheimer's disease, who may have trouble encoding events in memory (Zacks et al., 2006).

Conclusion

Here, we demonstrate that event segmentation is underpinned by neural state boundaries that occur in a nested cortical hierarchy. This work also provides the first cortex-wide mapping of timescales of information processing and shows that the DMN fractionates into fast and slow subnetworks. Together, these findings provide new insights into the neural mechanisms that underlie event segmentation, which in turn is a critical component of real-world perception, working memory and episodic memory formation. What remains to be addressed is how timescales of different brain regions relate to the types of neural representations that are contained within these regions. For example, does the dissociation between the posterior and anterior default mode network reflect relatively fast construction of mental scenes and slow integration with existing knowledge, respectively? Studying brain function from this perspective provides us with a new view on the organizational principles of the human brain.

Acknowledgements

LG was supported by a Veni grant [451-16-013] from the Netherlands Organization for Scientific Research. MVG was supported by a Vidi grant [639.072.513] from the Netherlands Organization for Scientific Research. KC was supported by the Natural Sciences and Engineering Research Council of Canada (Grant RGPIN-2017-03804 to KC) and the Canada Research Chairs program. We thank Aya Ben-Yakov for providing data on the subjective event boundaries in the Cam-CAN movie dataset. We thank Shiori Amemiya for provided the HRF time delay maps. Data collection and sharing for this project was provided by the Cambridge Centre for Ageing and Neuroscience (CamCAN). CamCAN funding was provided by the UK Biotechnology and Biological Sciences Research Council (grant number BB/H008217/1), together with support from the UK Medical Research Council and University of Cambridge, UK.

Materials and Methods

Participants

This study included data from 265 adults (131 female) who were aged 18–50 (mean age 36.3, SD = 8.6) from the healthy, population-derived cohort tested in Stage II of the Cam-CAN project (Shafto et al., 2014; Taylor et al., 2017). Participants were native English-speakers, had normal or corrected-to-normal vision and hearing, and had no neurological disorders (Shafto et al., 2014). Ethical approval for the study was obtained from the Cambridgeshire 2 (now East of England - Cambridge Central) Research Ethics Committee. Participants gave written informed consent.

Movie

Participants watched a black and white television drama by Alfred Hitchcock called ‘Bang! You’re Dead’ while they were scanned with fMRI. The full 25-minute episode was shortened to 8 minutes, preserving the narrative of the episode (Shafto et al., 2014). A longer version of this movie has been shown to elicit robust brain activity, synchronized across younger participants (Hasson et al., 2009b). Participants were instructed to watch, listen, and pay attention to the movie.

fMRI data acquisition

The details of the fMRI data acquisition are described in (Geerligs et al., 2018). In short, 193 volumes of movie data were acquired with a 32-channel head-coil, using a multi-echo, T2*-weighted EPI sequence. Each volume contained 32 axial slices (acquired in descending order), with slice thickness of 3.7 mm and interslice gap of 20% (TR = 2470 ms; five echoes [TE = 9.4 ms, 21.2 ms, 33 ms, 45 ms, 57 ms]; flip angle = 78 degrees; FOV = 192mm x 192 mm; voxel-size = 3 mm x 3 mm x 4.44 mm), the acquisition time was 8 minutes and 13 seconds. High-resolution (1 mm x 1mm x 1 mm) T1 and T2-weighted images were also acquired.

Data pre-processing and hyperalignment

The initial steps of data preprocessing for the movie data were the same as in Geerligs et al. (2018) and are described there in detail. Briefly, the preprocessing steps included deobliquing of each TE, slice time correction and realignment of each TE to the first TE in the run, using AFNI (version AFNI_17.1.01; <https://afni.nimh.nih.gov>; Cox, 1996). To denoise the data for each participant, we used multi-echo independent component analysis (ME-ICA), which is a very promising method for removal of non-BOLD like components from the fMRI data, including effects of head motion (Kundu et al., 2012; Kundu et al., 2013). Co-registration followed by DARTEL intersubject alignment was used to align participants to MNI space using SPM12 software (<http://www.fil.ion.ucl.ac.uk/spm>).

To optimally align voxels across participants in the movie dataset, we subsequently used whole-brain searchlight hyperalignment as implemented in the PyMVPA toolbox (Guntupalli et al., 2016; Hanke et al., 2009). Hyperalignment is an important step in the pipeline because the neural state segmentation method relies on group-averaged voxel-level data. Hyperalignment uses Procrustes transformations to derive the optimal rotation parameters that minimize intersubject distances between responses to

the same timepoints in the movie. The details of the procedure are identical to those in Geerligs et al. (2020). After hyperalignment, the data were highpass-filtered with a cut-off of 0.008 Hz.

Neural state boundaries

To identify neural state boundaries in the fMRI data, we used greedy state boundary search (GSBS) (Geerligs et al., 2020). GSBS performs an iterative search for state boundary locations that optimize the similarity between the average activity patterns in a neural state and the (original) brain activity at each corresponding time point. At each iteration of the algorithm, previous boundary locations are finetuned by shifting them by 1 TR (earlier or later) if this further improves the fit. To determine the optimal number of boundaries in each brain region, we used the t-distance metric. This metric identifies the optimal number of states, such that timepoints within a state have maximally similar brain activity patterns, while timepoints in consecutive states are maximally dissimilar. The validity of these methods has been tested extensively in previous work, with both simulated and empirical data (Geerligs et al., 2020). The input to the GSBS algorithm consists of a set of voxel time courses within a searchlight and a maximum value for the number of states, which we set to 100, roughly corresponding to half the number of TRs in our data (Geerligs et al., 2020).

We applied GSBS in a searchlight to the hyperaligned movie data. Spherical searchlights were scanned within the Harvard-Oxford cortical mask with a step size of two voxels and a radius of three voxels (Desikan et al., 2006). This resulted in searchlights with an average size of 97 voxels (max: 123, IQR: 82-115), this variation in searchlight size was due to the exclusion of out-of-brain voxels. Only searchlights with more than 15 voxels were included in the analysis. Previous analyses have shown that neural state boundaries cannot be identified reliably in single-subject data. Instead data should be averaged across at least ~17 participants, to eliminate sources of noise from the data (Geerligs et al., 2020). To obtain valid boundaries while still being able to determine the consistency of our findings across participants, we therefore fit the state boundaries and optimal number of states within 15 independent datasets. Each dataset consisted of averaged data across 17 or 18 randomly selected participants.

Fast and slow regions

To investigate whether different brain regions showed reliable differences in median state length, we tested whether some regions had significantly shorter or longer neural state durations than the average duration across the brain. To this end, we computed the median state duration within each of the 15 independent datasets for each searchlight and tested whether these were significantly different from the averaged median duration across all of the searchlights using the Wilcoxon signed rank test (Wilcoxon, 1945). P-values were corrected for multiple comparisons using the false discovery (FDR) correction procedure that is accurate for any test dependency structure (Benjamini and Yekutieli, 2001).

Comparison of neural state boundaries to event boundaries

Event boundaries in the Cam-CAN movie dataset were identified by Ben-Yakov and Henson (2018) based on data from sixteen observers. These participants watched the movie outside the scanner and indicated with a keypress when they felt “one event (meaningful unit) ended and another began”. Participants were not able to rewind the movie. To make sure the identified boundaries were reliable, only boundaries identified by at least five observers were included. This resulted in a total of 19 boundaries separated by 6.5–93.7 s.

To compare the neural state boundaries across regions to the event boundaries, we computed Pearson correlations between boundary vectors. These boundary vectors contained zeros for within-state/event time points and ones for time points where the current state/event did not match the previous state (i.e., state boundaries). Because it has been shown that the peak of the hemodynamic response can vary in latency (Handwerker et al., 2004), we used different delays to convert the event boundary onsets (in seconds) to boundary vectors (in TRs). In particular, we added a delay of four up to eight seconds, with steps of 0.5 seconds, to the event boundary onset before down sampling from seconds to TRs.

For each delay, we computed the correlation between the event boundary vector and the state boundary vectors. Next, we identified the delay with the maximal correlation. To account for spuriously high correlations introduced by this approach, we used a randomization procedure. In particular, for each delay the event boundary vectors were randomized such that the durations between events was preserved but the order of events was shuffled. This randomization was repeated 1000 times and in each case, we recomputed the correlation with the neural state boundary vectors and identified the maximal correlation across the nine possible delays. The final adjusted correlation metric was computed by scaling the maximum correlation measure of the real data with respect to the average expected correlation measure in the randomized data; zero indicates that the correlation is the same as the average expected correlation and one indicates perfect correspondence between the two boundary vectors.

In addition to investigating the similarity between the event boundary vectors and the state boundary vectors, we also investigated the effect of boundary strength and boundary co-occurrence, using the same approach. Boundary strength is defined as the Pearson correlation distance between the neural activity patterns of consecutive neural states. Boundary co-occurrence is defined as the proportion of searchlights that show a boundary at the same time. For computing the association between event boundaries and boundary strength, only TRs in which a neural state boundary occurred were taken into account.

Each analysis was done separately across each of the 15 independent samples. A p-value was obtained for each region by testing whether the adjusted correlations across all 15 samples differed significantly from zero using a Wilcoxon signed rank test (Wilcoxon, 1945). P-values were corrected for multiple comparisons using FDR correction (Benjamini and Yekutieli, 2001).

To ensure that differences between regions in the correspondence between neural state boundaries and event boundaries were not driven by timescale differences, we ran an additional analysis where we fixed the number of states in each region to be identical to the number of events (K=19; rather than using each region’s optimal number of states).

Identification of functional networks

In order to identify networks of regions that contained the same neural state boundaries, and not simply the same duration of states, we calculated a matrix of Pearson correlations between the boundary vectors in each of the searchlights for each of the 15 independent groups. Functional networks were detected using a consensus partitioning algorithm (Lancichinetti and Fortunato, 2012), as implemented in the Brain Connectivity Toolbox (Rubinov and Sporns, 2010). An initial partition into functional networks was created using the Louvain modularity algorithm (Blondel et al., 2008), which was refined using a modularity fine-tuning algorithm (Sun et al., 2009) to optimize the modularity. The fit of the partitioning was quantified using an asymmetric measure of modularity that assigns a lower importance to negative weights than positive weights (Rubinov and Sporns, 2011).

The partitioning was done separately for each of the 15 independent participant groups and was repeated 100 times within in each participant group to take the stochasticity of the modularity maximization into account. Subsequently, all repetitions were combined into a consensus matrix. Each element in the consensus matrix indicates the proportion of repetitions and groups in which the corresponding two searchlights were assigned to the same network. The consensus matrix was thresholded such that values less than those expected by chance were set to zero (Bassett et al., 2013). This thresholded consensus matrix was used as the input for a new partitioning, using the same method described above, until the algorithm converged to a single partition (such that the final consensus matrix consisted only of ones and zeroes).

The procedure described above was applied for multiple resolutions (varying gamma between 1 and 3, Reichardt and Bornholdt, 2006), which was always matched across the initial and consensus partitioning. We selected the partition with the highest similarity to a previous whole brain network partition (Power et al., 2011), as measured by adjusted mutual information (aMI; Xuan Vinh et al., 2010). To compare our network labels for each searchlight to the voxelwise Power (2011) networks, we labeled each searchlight according to the Power network label that occurred most frequently in the searchlight voxels. The highest similarity was observed for gamma=2.3 (aMI=0.41). We named each functional network we identified in accordance with the Power (2011) network that it overlapped most with, in addition to a descriptive term about the network location (e.g. ventral, posterior).

We also investigated whether neural states are shared in a hierarchical fashion across different functional networks. To this end, we computed partial correlations between the averaged boundary vectors within each functional network, controlling for the average boundary vectors in all other networks. Using partial correlations allowed us to focus on direct, rather than indirect connections between networks (Smith et al., 2011). To identify pairs of networks with partial correlations that were significantly higher than zero, we applied the Wilcoxon signed rank test to the partial correlation estimates across all 15 independent groups of participants (Wilcoxon, 1945). P-values were corrected for multiple comparisons using FDR correction (Benjamini and Yekutieli, 2001). Connections that would be estimated as positive by full correlation, but were estimated as negative by partial correlation were not taken into account because it is not clear how they should be interpreted (Smith et al., 2011).

Data visualization

In all analyses, p-values from the searchlights were projected to the voxel level and averaged across the searchlights that overlapped each voxel before they were thresholded using the FDR-corrected critical p-value (Benjamini and Yekutieli, 2001). When projecting the results of the analyses to the voxel-level, we excluded voxels for which less than half of the searchlights that covered that voxel were included in the analysis. These excluded searchlights had too few in-brain voxels (see section [Neural state boundaries](#)). Data were projected to the surface for visualization, using the Caret toolbox (Van Essen et al., 2001).

References

- Amemiya S, Takao H, Abe O (2020): Origin of the Time Lag Phenomenon and the Global Signal in Resting-State fMRI. *Front Neurosci* 14:1141.
- Andermane N, Joensen BH, Horner AJ (2021): Forgetting across a hierarchy of episodic representations. *Current Opinion in Neurobiology*. Elsevier Ltd.
- Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL (2010): Functional-anatomic fractionation of the brain's default network. *Neuron* 65:550–62.
- Baldassano C, Chen J, Zadbood A, Pillow JW, Hasson U, Norman KA (2017): Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron* 95:709-721.e5.
- Baron-Cohen S (2000): Theory of mind and autism: A review. *Int Rev Res Ment Retard* 23:169–184.
- Bassett DS, Porter M a., Wymbs NF, Grafton ST, Carlson JM, Mucha PJ (2013): Robust detection of dynamic community structure in networks. *Chaos* 23.
- Ben-Yakov A, Henson RN (2018): The Hippocampal Film Editor: Sensitivity and Specificity to Event Boundaries in Continuous Experience. *J Neurosci* 38:10057–10068.
- Benjamini Y, Yekutieli D (2011): The Control of the False Discovery Rate in Multiple Testing under Dependency. *Ann Stat* 29:1165–1188.
- Benjamini Y, Yekutieli D (2001): The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188.
- Benoit RG, Szpunar KK, Schacter DL (2014): Ventromedial prefrontal cortex supports affective future simulation by integrating distributed knowledge. *Proc Natl Acad Sci U S A* 111:16550–16555.
- Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008): Fast unfolding of communities in large networks. *J Stat Mech Theory Exp* p. P10008.
- Campbell KL, Grigg O, Saverino C, Churchill N, Grady CL (2013): Age differences in the intrinsic functional connectivity of default network subsystems. *Front Aging Neurosci* 5:1–12.
- Chien HYS, Honey CJ (2020): Constructing and Forgetting Temporal Context in the Human Cerebral Cortex. *Neuron* 106:675-686.e11.
- Clarke A, Tyler LK (2015): Understanding What We See: How We Derive Meaning From Vision. *Trends in Cognitive Sciences*. Elsevier Ltd.
- Clewett D, DuBrow S, Davachi L (2019): Transcending time in the brain: How event memories are constructed from experience. *Hippocampus* 29:162–183.

Cox RW (1996): AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29:162–173.

Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ (2006): An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31:968–980.

DiCarlo JJ, Cox DD (2007): Untangling invariant object recognition. *Trends Cogn Sci* 11:333–341.

Duncan J (2010): The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn Sci* 14:172–9.

Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH (2001): An integrated software suite for surface-based analyses of cerebral cortex. *J Am Med Informatics Assoc* 8:443–459.

Ezzyat Y, Davachi L (2011): What constitutes an episode in episodic memory? *Psychol Sci* 22:243–52.

Flores S, Bailey HR, Eisenberg ML, Zacks JM (2017): Event segmentation improves event memory up to one month later. *J Exp Psychol Learn Mem Cogn* 43:1183–1202.

Fuster JM (2001): The prefrontal cortex - An update: Time is of the essence. *Neuron*. Cell Press.

Geerligs L, Cam-CAN, Campbell KL (2018): Age-related differences in information processing during movie watching. *Neurobiol Aging* 72:106:120.

Geerligs L, Van Gerven M, Güçlü U (2020): Detecting neural state transitions underlying event segmentation. *bioRxiv*.

Gilboa A, Marlatte H (2017): Neurobiology of Schemas and Schema-Mediated Memory. *Trends Cogn Sci* 21:618–631.

Guntupalli JS, Hanke M, Halchenko YO, Connolly AC, Ramadge PJ, Haxby J V (2016): A Model of Representational Spaces in Human Cortex. *Cereb Cortex* 26:2919–2934.

Hamilton LS, Huth AG (2020): The revolution will not be controlled: natural stimuli in speech neuroscience. *Lang Cogn Neurosci* 35:573–582.

Handwerker DA, Ollinger JM, D’Esposito M (2004): Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21:1639–1651.

Hanke M, Halchenko YO, Sederberg PB, Hanson SJ, Haxby J V., Pollmann S (2009): PyMVPA: a Python Toolbox for Multivariate Pattern Analysis of fMRI Data. *Neuroinformatics* 7:37–53.

Hasson U, Avidan G, Gelbard H, Vallines I, Harel M, Minshew N, Behrmann M (2009a): Shared and idiosyncratic cortical activation patterns in autism revealed under continuous real-life viewing conditions. *Autism Res* 2:220–231.

Hasson U, Chen J, Honey CJ (2015): Hierarchical process memory: Memory as an integral component of information processing. *Trends Cogn Sci* 19:304–313.

Hasson U, Malach R, Heeger DJ (2009b): Reliability of cortical activity during natural stimulation. *Trends Cogn Sci* 14:40–48.

Hasson U, Yang E, Vallines I, Heeger D, Rubin N (2008): A hierarchy of temporal receptive windows in human cortex. *J Neurosci* 28:2539–2550.

Honey CJ, Thesen T, Donner TH, Silbert LJ, Carlson CE, Devinsky O, Doyle WK, Rubin N, Heeger DJ, Hasson U (2012): Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* 76:423–34.

- van Kesteren MTR, Ruitter DJ, Fernandez G, Henson RN (2012): How schema and novelty augment memory formation. *Trends Neurosci* 35:211–219.
- Van Kesteren MTR, Ruitter DJ, Fernández G, Henson RN (2012): How schema and novelty augment memory formation. *Trends Neurosci* 35:211–219.
- Kundu P, Brenowitz ND, Voon V, Worbe Y, Vértes PE, Inati SJ, Saad ZS, Bandettini P a, Bullmore ET (2013): Integrated strategy for improving functional connectivity mapping using multiecho fMRI. *Proc Natl Acad Sci U S A* 110:16187–92.
- Kundu P, Inati SJ, Evans JW, Luh W-M, Bandettini PA (2012): Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *Neuroimage* 60:1759–70.
- Kurby CA, Zacks JM (2008): Segmentation in the perception and memory of events. *Trends Cogn Sci* 12:72–79.
- Kurby CA, Zacks JM (2018): Preserved neural event segmentation in healthy older adults. *Psychol Aging* 33:232–245.
- Lambon Ralph MA, Sage K, Jones RW, Mayberry EJ (2010): Coherent concepts are computed in the anterior temporal lobes. *Proc Natl Acad Sci U S A* 107:2717–2722.
- Lancichinetti A, Fortunato S (2012): Consensus clustering in complex networks. *Sci Rep* 2:336.
- Lee H, Bellana B, Chen J (2020): What can narratives tell us about the neural bases of human memory? *Curr Opin Behav Sci* 32:111–119.
- Lei X, Wang Y, Yuan H, Mantini D (2014): Neuronal oscillations and functional interactions between resting state networks. *Hum Brain Mapp* 35:3517–28.
- Lerner Y, Honey CJ, Silbert LJ, Hasson U (2011): Topographic Mapping of a Hierarchy of Temporal Receptive Windows Using a Narrated Story. *J Neurosci* 31:2906–2915.
- Nelson MJ, El I, Giber K, Yang X, Cohen L, Koopman H, Cash SS, Naccache L, Hale JT, Pallier C, Dehaene S (2017): Neurophysiological dynamics of phrase-structure building during sentence processing.
- Newton D, Engquist GA, Bois J (1977): The objective basis of behavior units. *J Pers Soc Psychol* 35:847–862.
- Newton D, Rindner RJ (1979): Variation in behavior perception and ability attribution. *J Pers Soc Psychol* 37:1847–1858.
- Northoff G, Heinzel A, de Greck M, Bormpohl F, Dobrowolny H, Panksepp J (2006): Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *Neuroimage* 31:440–457.
- Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church J a, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE (2011): Functional network organization of the human brain. *Neuron* 72:665–78.
- Ranganath C, Ritchey M (2012): Two cortical systems for memory-guided behaviour. *Nat Rev Neurosci* 13:713–726.
- Reichardt J, Bornholdt S (2006): Statistical mechanics of community detection. *Phys Rev E* 74:16110.
- Rogers TT, Lambon Ralph MA, Garrard P, Bozeat S, McClelland JL, Hodges JR, Patterson K (2004): Structure and Deterioration of Semantic Memory: A Neuropsychological and Computational Investigation. *Psychol Rev* 111:205–235.
- Rubinov M, Sporns O (2010): Complex network measures of brain connectivity: Uses and

- interpretations. *Neuroimage* 52:1059–1069.
- Rubinov M, Sporns O (2011): Weight-conserving characterization of complex functional brain networks. *Neuroimage* 56:2068–2079.
- Sargent JQ, Zacks JM, Hambrick DZ, Zacks RT, Kurby CA, Bailey HR, Eisenberg ML, Beck TM (2013): Event segmentation ability uniquely predicts event memory. *Cognition* 129:241–255.
- Shafto MA, Tyler LK, Dixon M, Taylor JR, Rowe JB, Cusack R, Calder AJ, Marslen-Wilson WD, Duncan J, Dalgleish T, Henson RN, Brayne C, Cam-CAN, Matthews FE (2014): The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: A cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC Neurol* 14.
- Simony E, Honey CJ, Chen J, Lositsky O, Yeshurun Y, Wiesel A, Hasson U (2016): Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat Commun* 7:12141.
- Smith SM, Miller KL, Salimi-khorshidi G, Webster M, Beckmann CF, Nichols TE, Ramsey JD, Woolrich MW (2011): NeuroImage Network modelling methods for fMRI. *Neuroimage* 54:875–891.
- Speer NK, Reynolds JR, Zacks JM (2007): Human brain activity time-locked to narrative event boundaries. *Psychol Sci* 18:449–455.
- Stephens GJ, Honey CJ, Hasson U (2013): A place for time: The spatiotemporal structure of neural dynamics during natural audition. *J Neurophysiol* 110:2019–2026.
- Sun Y, Danila B, Josić K, Bassler KE (2009): Improved community structure detection using a modified fine-tuning strategy. *Europhys Lett* 86:28004.
- Taylor JR, Williams N, Cusack R, Auer T, Shafto M a., Dixon M, Tyler LK, Cam-CAN, Henson RN (2017): The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: Structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *Neuroimage* 144:262–269.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002): Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289.
- Wen T, Duncan J, Mitchell DJ (2020): Hierarchical representation of multistep tasks in multiple-demand and default mode networks. *J Neurosci* 40:7724–7738.
- Wilcoxon F (1945): Individual Comparisons by Ranking Methods. *Biometrics Bulletin*. Vol. 1.
- Willems RM, Nastase SA, Milivojevic B (2020): Narratives for Neuroscience. *Trends Neurosci* 43:271–273.
- Xuan Vinh N, Epps J, Bailey J (2010): Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance. *Journal of Machine Learning Research*. Vol. 11.
- Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011): The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106:1125–65.
- Zacks JM, Tversky B, Iyer G (2001a): Perceiving, remembering, and communicating structure in events. *J Exp Psychol Gen* 130:29–58.
- Zacks JM, Braver TS, Sheridan MA, Donaldson DI, Snyder AZ, Ollinger JM, Buckner RL, Raichle ME (2001b): Human brain activity time-locked to perceptual event boundaries. *Nat Neurosci* 4:651–

655.

Zacks JM, Kurby CA, Eisenberg ML, Haroutunian N (2011): Prediction error associated with the perceptual segmentation of naturalistic events. *J Cogn Neurosci* 23:4057–4066.

Zacks JM, Speer NK, Reynolds JR (2009): Segmentation in Reading and Film Comprehension. *J Exp Psychol Gen* 138:307–327.

Zacks JM, Speer NK, Swallow KM, Braver TS, Reynolds JR (2007): Event perception: A mind-brain perspective. *Psychol Bull* 133:273–293.

Zacks JM, Speer NK, Swallow KM, Maley CJ (2010): The brain's cutting-room floor : segmentation of narrative cinema. *Front Aging Neurosci* 4:1–15.

Zacks JM, Speer NK, Vettel JM, Jacoby LL (2006): Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychol Aging* 21:466–482.

Zwaan RA, Radvansky G (1998): Situation models in language comprehension and memory. *Psychol Bull* 123:162–185.

Supplementary materials

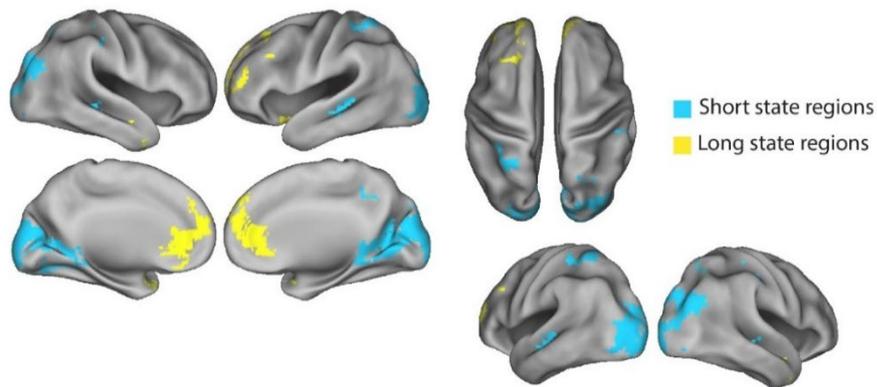
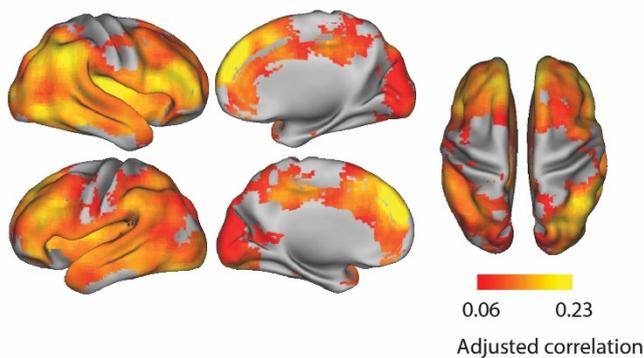


Figure S1: The regions with significantly shorter or longer median state lengths than average are depicted (FDR-corrected p -value < 0.05), indicating that differences in the median state length across brain regions are reliable.

A. Similarity between neural states and subjective event boundaries ($k=19$)



B. Delay between neural state boundary and event boundary

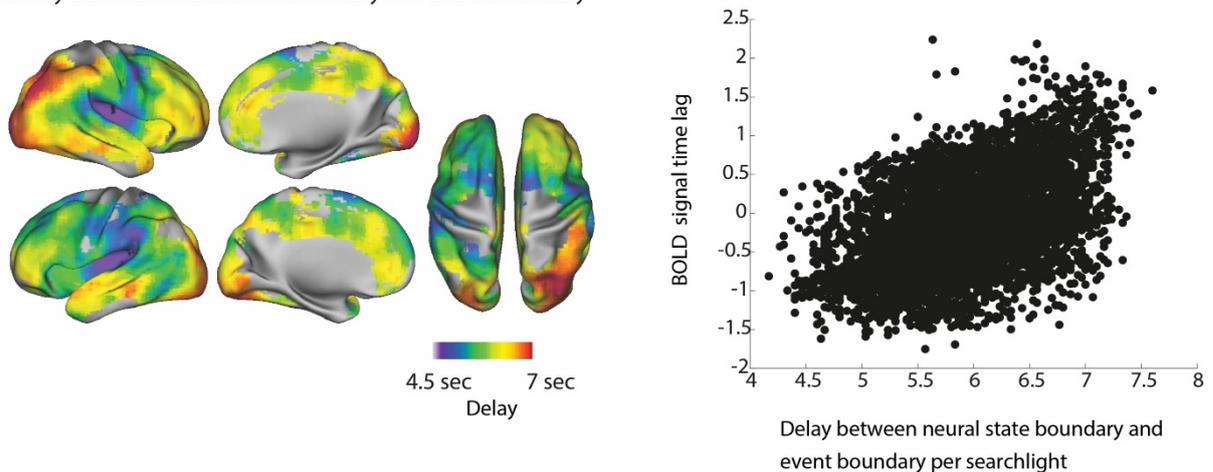


Figure S2. A. Association between the neural state boundaries identified in the fMRI data and the subjective event boundaries identified by participants outside the scanner, where the number of neural states is fixed at $K=19$ for each searchlight. Results are highly similar to the findings reported in the paper based on the optimal number of states for each brain region. B. The delay for which the optimal correspondence between neural state boundaries and event boundaries was observed. The

scatterplot shows that the observed delay in each searchlight was similar to the BOLD signal time lag that was previously identified based on the time lag of the global signal (Amemiya et al., 2020).

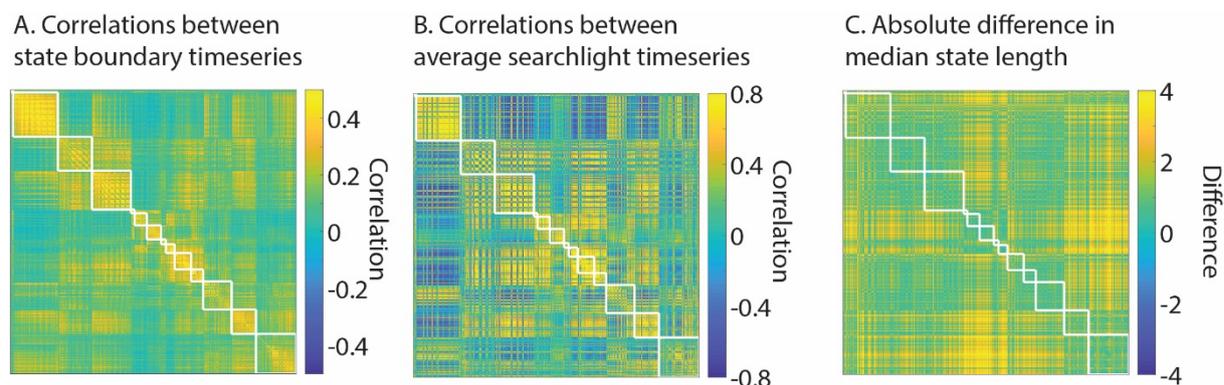


Figure S3. A. The correlation matrix between the state boundary vectors of each pair of searchlights. B. The correlation matrix based on the averaged brain activity timecourse in each searchlight (i.e., standard measure of functional connectivity). C. The difference between each pair of searchlights in median state length was markedly different from the correlation between state boundary vectors (S3A), showing that the correlation between different regions was not just due to regional differences in the optimal number of states.

Supplementary table 1: For each network, the table lists the network defined by Power et al (2011) that showed the highest overlap and the percentage of searchlights in the network that overlapped with that particular Power (2011) network.

Network name	Power network	Percentage of searchlights
Auditory	Auditory	24
Sensorimotor-medial	Sensorimotor (hand)	41
Sensorimotor-lateral	Sensorimotor (hand)	20
Visual medial	Visual	83
Visual medial ventral	Visual	53
Visual medial dorsal	Visual	73
Fusiform	Fusiform*	50
Visual lateral posterior	Visual	76
Visual lateral	Visual	34
Dorsal attention network	Dorsal attention network	25
Default mode network posterior	Default mode network	40
Fronto-parietal task control	Fronto-parietal task control	34
Default mode network anterior	Default mode network	52

* This network was not mentioned in the Power (2011) manuscript. However, it was part of the voxel-wise network labeled nifti image that was made available. We named this network in accordance with its overlap with the Fusiform gyrus in the automated anatomical labeling atlas (Tzourio-Mazoyer et al., 2002).