# Top-down attention shifts behavioral and neural event boundaries in narratives with overlapping event scripts

Alexandra De Soares[1], Tony Kim[1], Franck Mugisho[2], Elen Zhu[1],

Allison Lin[1], Chen Zheng[3], Christopher Baldassano[1]

[1] Department of Psychology, Columbia University

[2] Department of Computer Science, Columbia University

[3] Department of Human Development, Teachers College, Columbia University

## Summary

Understanding and remembering the complex experiences of everyday life is only possible because we can draw on prior knowledge about how events in our world unfold over time. How does the brain construct event representations using our library of schematic scripts, and what are the behavioral and neural consequences of activating a specific script during perception? We developed a novel set of 16 stories, each of which combines one of four location-relevant event scripts (restaurant, airport, grocery store, lecture hall) with one of four socially-relevant event scripts (breakup, proposal, business deal, meet cute). Participants in an fMRI study and a separate online study listened to these stories after being primed to attend to one of the two scripts or without priming. Responses in angular gyrus, parahippocampal gyrus, and subregions of medial prefrontal cortex (mPFC) were driven by both location and social script information, showing that these regions can track schematic sequences from multiple domains. Comparing priming conditions, we found that activating a location-related event script shifted the timing of subjective event boundaries to align with script-relevant changes in the narratives. This behavioral shift was mirrored by shifts in the timing of neural responses, with mPFC event boundaries (identified using a Hidden Markov Model) aligning to location-relevant rather than socially-relevant boundaries when participants were location primed. Our findings demonstrate that neural event dynamics are actively modulated by top-down goals, and provide new insight into how narrative event representations are constructed through the activation of temporally-structured prior knowledge.

# Introduction

A key challenge in human neuroscience is to understand how the complex, continuous stream of information we experience during our everyday lives is used to construct stable and meaningful neural representations. Despite constantly-changing neural responses in early sensory regions, work over the past decade has shown that stimuli with meaningful temporal structure (such as narratives or film clips) evoke representations in higher-order brain regions that are stable over tens of seconds [1,2]. These stable periods are punctuated by moments at which representations rapidly shift, which correspond to boundaries between different events in the narrative [3–5]. This stability is thought to be scaffolded by the activation of schematic event scripts [6–8] that reflect prior knowledge about the sequential temporal structure of events we encounter in the world, allowing us to strategically direct our attention and make predictions [9–11]. For example, when catching a flight, experienced travelers will have clear expectations about the sequence of events that will occur at an airport and the critical pieces of information they will need to track at each stage. Stimuli that activate prior schematic knowledge drive responses in higher-order brain regions, including the medial prefrontal cortex (mPFC) [12–16].

Prior neuroimaging studies of event scripts, however, have had two major limitations. First, these studies have assumed that there is only a single relevant script active during a story, while realistic events are much more likely to be a combination of multiple, overlapping scripts, such as celebrating a birthday at a restaurant. Even the earliest work on story scripts noted that “the concurrent activation of more than one script creates rather complex problems” [6]; multiple scripts must be simultaneously tracked, scripts can compete for incoming pieces of information and for attention, and different scripts may generate conflicting event boundaries. One possibility for how the brain could track multiple scripts simultaneously is by representing different kinds of scripts in separate domain-specific regions, as has been proposed for other cognitive processes [17]. For example, scripts for social interactions and scripts related to the

physical location an event takes place could be supported by different brain regions, reducing interference. Alternatively, knowledge from multiple scripts could contribute to a single high-level event model maintained in the Default Mode Network (DMN) [12,13,18]. If this event model is sufficiently high-dimensional, then information from various scripts (as well as unique episode-specific details) could be combined to create a representation of the current event.

The other major shortcoming of previous paradigms for studying event scripts is that they cannot address the central theoretical claim that script activation changes how neural responses are organized into events [11,19,20], since brain responses are compared across different stimuli with varying degrees of schematicity or varying event scripts. Narratives with multiple simultaneous scripts instead provide an intriguing opportunity to examine the causal impact of activating an event script on neural responses, since the level of attention to each script can be manipulated for the same narrative. Priming participants with different schemas can change how narratives are perceived and remembered, with better memory for details relevant to the primed schema [21]. Neurally, participants who interpret a story using different schemas also exhibit different representations in higher-order brain areas [22–24]. These studies have used only static schemas, however; event scripts consist of a sequence of temporal units in which the script-relevant actions and goals change over the course of the narrative, providing a potential scaffold for the temporal event structure of neural responses. Top-down processes such as switching between task sets have been shown to generate event boundaries [25], but whether activating a specific event script has the power to shift the event boundaries present in a narrative has not been tested.

Our study aims to address these questions about how multiple scripts are tracked during narrative perception and how directing attention to a specific script can influence perceptual dynamics. We collected fMRI data while participants listened to a novel set of narrative stimuli in which each story was built from two distinct event scripts (Fig 1). One of these scripts was

determined by the *location* that the story took place (e.g. an airport), and the other described the *social* interaction taking place (e.g. a marriage proposal). There were four location scripts and four social scripts, and each of the sixteen stories had a unique combination of a location and social script. Every script consisted of a sequence of four schematic events, interleaved such that the event boundaries for each script occurred at different sentences during the story. We found that regions throughout the DMN exhibited script-specific neural responses for both location and social scripts, supporting a domain-general role for this network in tracking event scripts. We also manipulated attention during some of the stories (in our fMRI study and in an online study), priming participants to pay attention to either the location script or social script. We found that activating the location-relevant script shifted subjective event boundaries to align with location-relevant boundaries in the narrative, and that there was a corresponding shift in the timing of neurally-identified event boundaries in mPFC. These results provide a key link connecting neural dynamics during naturalistic perception to decades of work on schematic and event perception in cognitive psychology, demonstrating that event boundaries are actively constructed and not simply produced by stimulus dynamics.

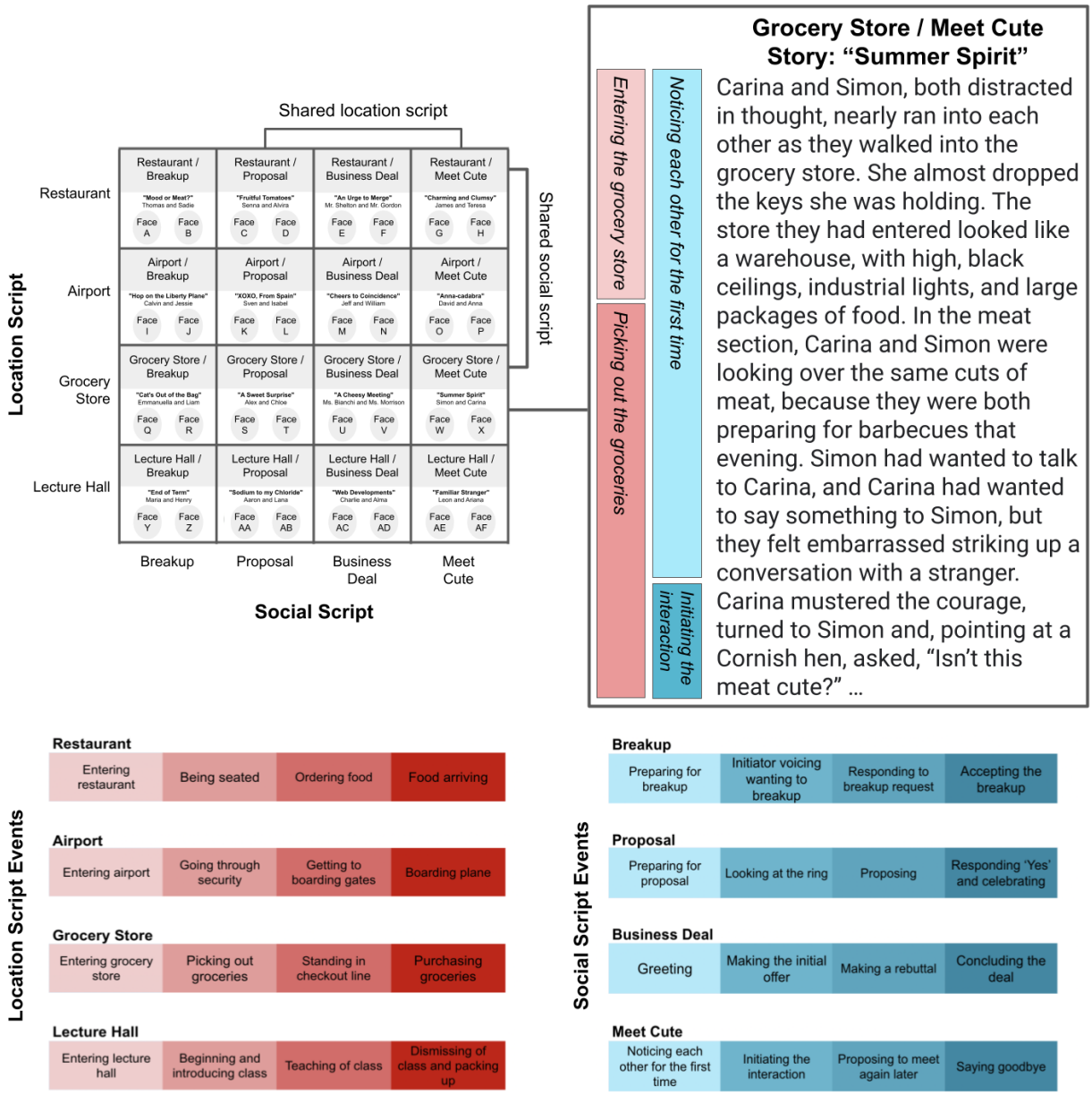


**Figure 1: Narrative stimulus set with overlapping scripts.** We selected four location-based scripts and four social-interaction scripts, each of which consisted of a stereotyped four-stage sequence of events. Each story was written based on a unique combination of a location and social script, resulting in four narratives for each script. The events in both scripts occurred simultaneously during the narrative, with event boundaries interleaved such that script-relevant transitions occurred at different sentences for the two scripts. The story was presented as spoken audio, while a static image of the story title and the two characters' faces and names was displayed. (Faces have been removed from this figure for this preprint, per bioRxiv policy.)

# Results

## Location and social scripts are represented in overlapping regions

To identify brain regions (in searchlights and predefined ROIs) that represented information about the location and social scripts, we measured the across-participant similarity for event patterns evoked by stories with shared scripts (Fig 2). Both location-related and social-related scripts were represented in overlapping regions throughout the DMN, particularly in the superior temporal sulcus (STS), the right angular gyrus, parahippocampal cortex (PHC), and a portion of mPFC. We also saw script-specific event representations for both script types in parts of visual cortex, despite there being no dynamic visual content in these stimuli. There were also some regions that represented only a single category of scripts, with selectivity for location-script effects in a broader set of visual regions and left insula and social-script effects in the left temporoparietal junction, left posterior medial cortex (PMC), and broader mPFC. Measuring the location and social script effects in predefined ROIs (Fig 2b), we found overlapping representations in the angular gyrus (location script effect: $p = 0.008$; social script effect: $p < 0.001$) and PHC (location $p = 0.004$ ; social $p = 0.012$). Significant social script effects (but not significant location script effects) were observed in the STS (location $p = 0.17$; social $p = 0.013$), superior frontal gyrus (SFG; location $p = 0.51$; social $p = 0.002$), mPFC (location $p = 0.128$; social $p = 0.018$), and hippocampus (location $p = 0.153$; social $p = 0.010$). This finding indicates that location and social script information is encoded in both shared and distinct regions, with notable overlap in the DMN, suggesting that abstract event representations in these high-level regions are constructed from multiple kinds of scripts in a domain-general way.

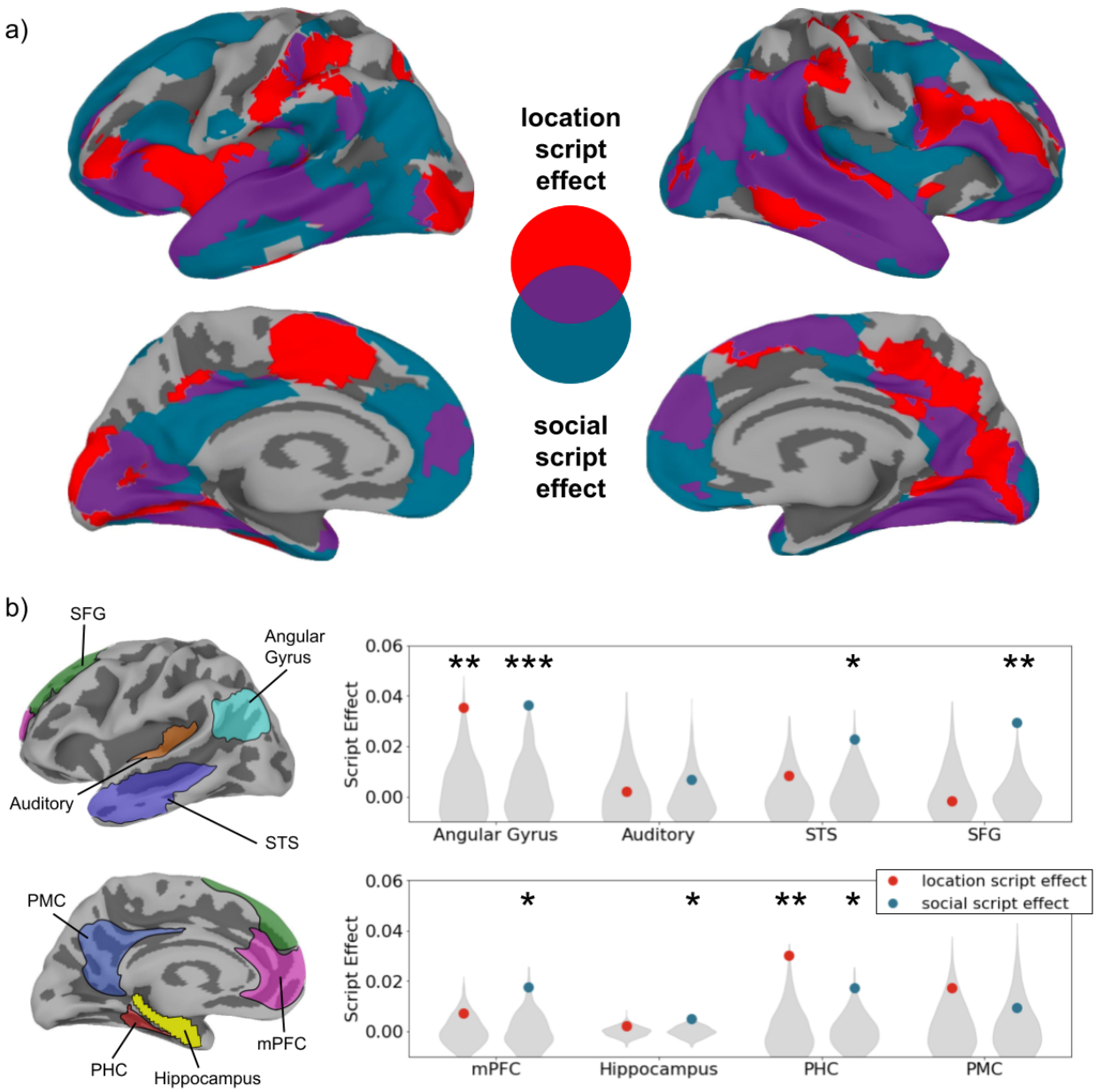


**Figure 2: Location and social script effects in searchlights and ROIs.** We compared event patterns across stories with shared scripts to identify regions that encode script-specific information. a) Performing a searchlight analysis on the cortical surface, we identified a set of high-level regions (purple) in which representations were influenced by both kinds of scripts, as well as some regions (red and blue) that tracked only one category of scripts. Both maps are thresholded at $q < 0.05$. b) Computing script effects in predefined ROIs [12], we found that angular gyrus and PHC significantly encoded both kinds of scripts, while the STS, SFG, mPFC, and Hippocampus significantly encoded the social script only.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

## Biasing attention to specific event scripts

To what extent does the activation of scripts depend on top-down attentional mechanisms, and what are the behavioral consequences of script activation? We conducted a separate online behavioral study (N=308) in which we primed participants to attend to either the location or social script in one of our sixteen stories (or a baseline condition, in which they did not receive either prime). We asked them to take on the perspective of a role associated with the schema (e.g. restaurant critic role for the restaurant script), and directed attention to each of the four script events by providing a sequence of four questions they should aim to answer while reading the story (Fig 3). The choice of which story was presented and the assignment to one of the three priming conditions was randomized across participants.

Using a memory test at the end of the story, we confirmed that script priming impacted the episodic details that participants were able to recall. Participants were asked questions about episodic details in the story they had just heard and were scored on a scale of 0-3 for how precisely they could recall the correct answer (Fig 4a). A mixed-effects model showed that social-script priming improved performance specifically on socially-relevant questions ($\beta = 0.393$, $t_{372.00} = 4.096$, $p < 0.001$), and there was also better overall performance for location-primed participants ($\beta = 0.262$, $t_{628.29} = 2.964$, $p = 0.003$) and on socially-relevant questions ($\beta = 0.266$, $t_{372.00} = 4.107$, $p < 0.001$). Directly comparing only the location- vs. socially-primed participants with an independent-samples t-test, we found superior performance on details relevant to the location script for location-primed participants ($\Delta\mu = 0.234$, $t_{231.34} = 2.620$, $p = 0.009$) and on details relevant to the social script for social-primed participants ($\Delta\mu = 0.277$, $t_{227.76} = 2.620$, $p = 0.007$). We replicated these memory results in our fMRI participants (Fig 4b), finding improved scores on socially-relevant details with social priming ($\beta = 0.638$, $t_{823} = 5.700$, $p < 0.001$) and improved performance on location-relevant details for location-primed participants ($\beta = 0.461$, $t_{823} = 4.122$, $p < 0.001$), with better overall performance for

location-primed participants ($\beta = 0.559$, $t_{823} = 7.066$, $p < 0.001$) and better scores for socially-relevant questions ($\beta = 0.343$, $t_{823} = 4.338$, $p < 0.001$). Directly comparing the location and social priming conditions only, we found significantly better performance on location questions for location-primed participants ($\Delta\mu = 0.517$, $t_{251} = 6.686$, $p < 0.001$) and on social questions for socially-primed participants ($\Delta\mu = 0.582$, $t_{251} = 6.767$, $p < 0.001$).

We also asked online participants to identify event boundaries while reading the story, indicating whether each sentence was the start of a new event (Fig 4c). Participants in all priming groups showed substantial agreement about which sentences were event boundaries, relative to a null distribution of boundaries with a matched distribution of event lengths ($p < 0.001$ for all groups). To determine whether the priming caused a systematic shift in boundary judgments toward the script-related boundaries we had created in the stories, we compared participants' event boundaries to the putative event boundaries corresponding to each script (Fig 4d). We found that the priming groups differed in their alignment to the location boundaries ($F_{2,305} = 3.53$, $p = 0.030$), with significantly better alignment for participants who were location-primed rather than socially-primed ($p = 0.023$ by post-hoc Tukey test). We did not observe differences across priming groups for alignment to the social-script boundaries ($F_{2,305} = 0.26$, $p = 0.77$). We also tested whether the event boundaries of participants who were primed in the same way became more similar to each other (regardless of whether they matched our putative script-related boundaries in the stimulus). We found (Fig 4e) that location-primed participants had boundaries that were significantly more similar than for random pairs of participants ($p < 0.001$), while social-script priming did not significantly improve boundary similarity ($p = 0.860$). These results show that manipulating script attention can influence the episodic details remembered in a narrative, and can (for location-related scripts) align subjective event boundaries to script-relevant event changes.

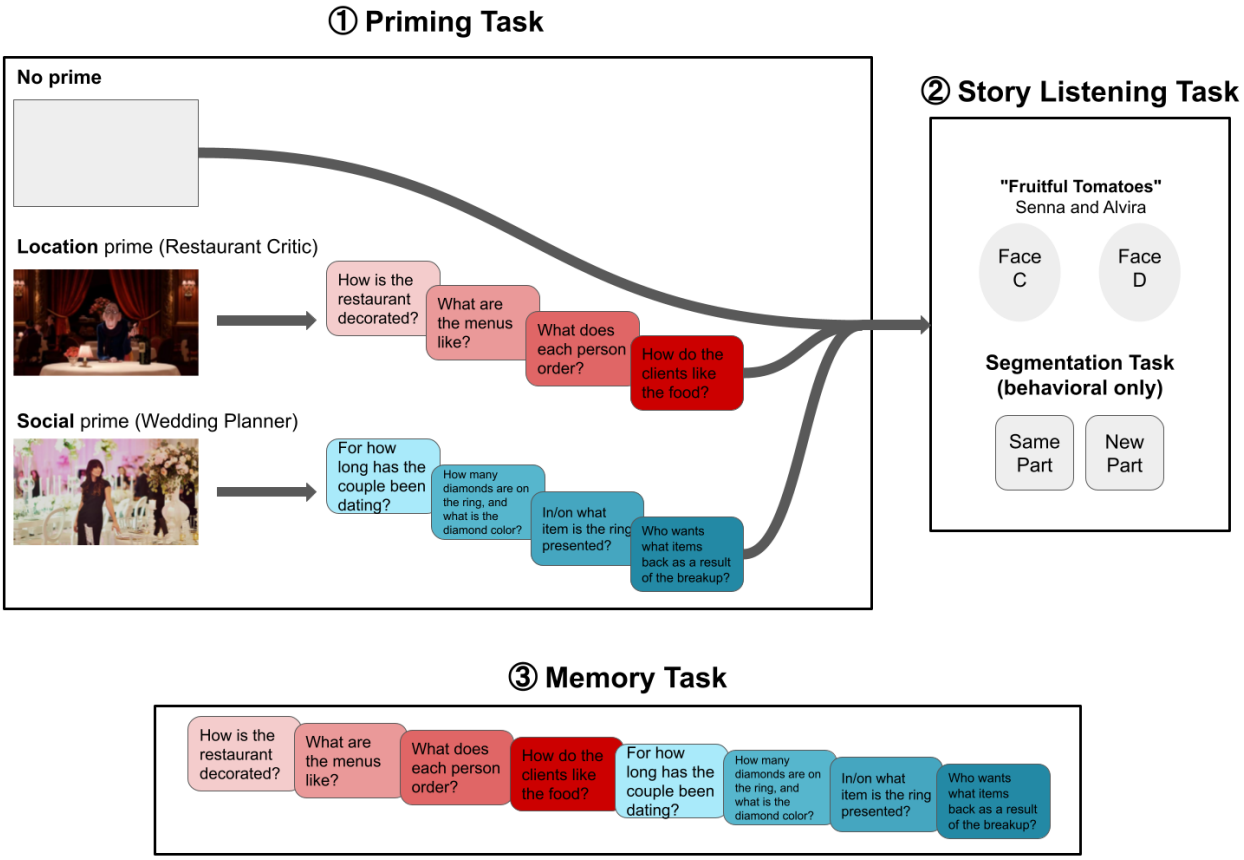


**Figure 3: Event script priming and behavioral segmentation task.** In both our fMRI experiment and a separate behavioral experiment, we placed participants in one of three priming conditions: no prime, location prime, or social prime. Participants in the location or social prime condition were told to take on the perspective of a role associated with the corresponding event script. We specifically directed their attention to each event in the primed script by providing them with a sequence of 4 questions they should aim to answer while listening to the story. Participants had to correctly identify and order the questions in the priming task in order to move on to the Story Listening Task, which was identical for all priming groups. In the online study only, participants completed a Segmentation Task while they listened to the story, indicating after listening to each sentence if that sentence was an event boundary. After participants listened to the story and completed a distractor task, participants answered eight questions about story-specific details, four related to each script. (Faces have been removed from this figure for this preprint, per bioRxiv policy.)

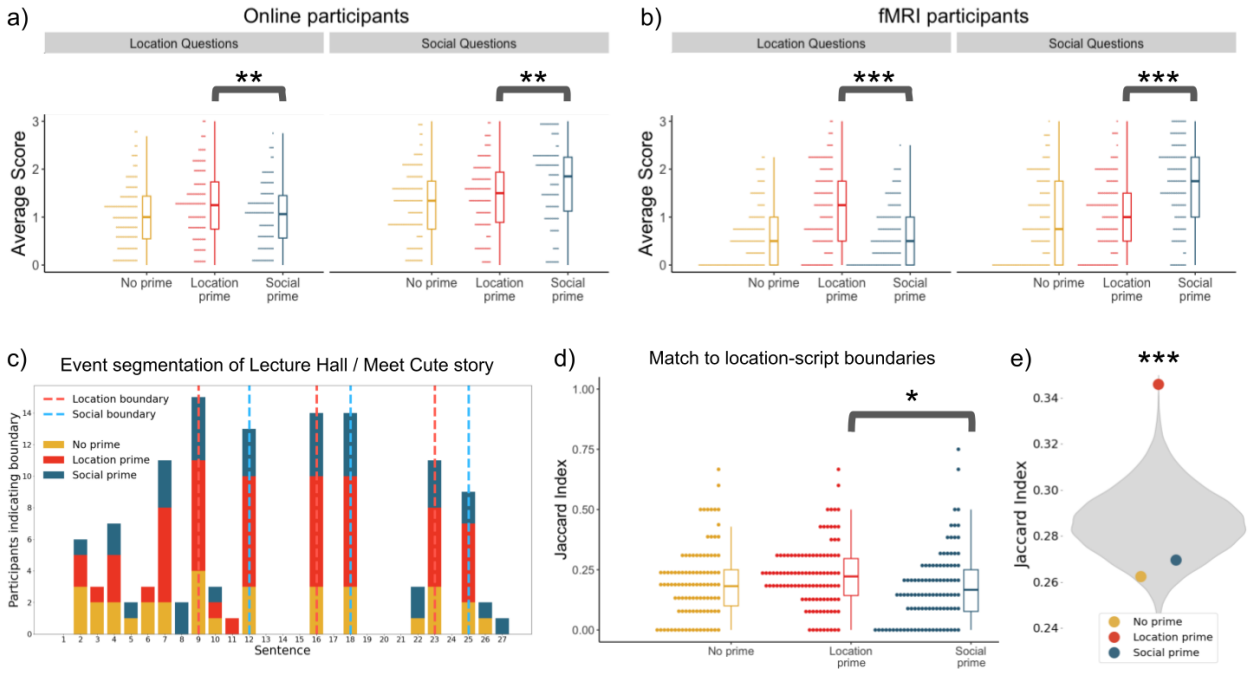


**Figure 4: Behavioral effects of biasing attention to event scripts.** a) Online participants were scored on their recall of story-specific details relevant to the location and social scripts. Priming had a significant effect on how participants remembered story details, with better memory for details relevant to the primed script. b) This effect replicated in our fMRI participants, who also showed significantly better memory for prime-relevant details. c) Online participants in all priming groups showed substantial agreement about which sentences were event boundaries, shown here for one example story. Colored bars indicate the number of participants in each condition that labeled a sentence as starting a new event, with dotted lines denoting the event transitions for each script. d) Boundaries annotated by location-primed participants were significantly more aligned to the location-related transitions in the narrative, compared to social-primed participants. e) Location-primed participants were significantly more similar to each other compared to random pairs of participants (gray).
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

## Priming-driven shifts in neural boundaries and response timecourses

Since our online study showed that biasing attention toward a location-related script shifted subjective event boundaries to align with script-relevant event boundaries in the stimulus, we tested whether there was a corresponding shift in neural event boundaries. We fit a Hidden Markov Model (HMM) with four latent events to the group average story-evoked activation for each priming group, temporally clustering multivoxel responses into stable events punctuated by rapid transitions. We performed this analysis in three regions of interest: mPFC, based on extensive prior work implicating it in schematic processing [12–16], and angular gyrus and PHC, since these areas exhibited significant script-specific patterns for both script types. The HMM produced an estimate of the timing for each group's transition through four events, allowing us to derive a measure of neural event change at each timepoint (Fig 5a). We measured the average neural event change at the putative location and social event boundaries to produce a measure of alignment between neural event boundaries and putative event boundaries.

The event boundaries in mPFC shifted substantially in response to location priming (Fig 5b). No-prime participants were neurally aligned with social boundaries (match to location boundaries, $p = 0.36$; social boundaries, $p = 0.043$), with similar (but only marginally significant) alignment for social-primed participants (location boundaries, $p = 0.40$; social boundaries, $p = 0.069$). Location priming, however, led to significant alignment with location-related boundaries (location boundaries, $p = 0.022$; social boundaries, $p = 0.44$), which marked a significant increase in alignment to location-related boundaries compared to the other conditions ($p = 0.045$) and a marginally-significant decrease in alignment to the social-related boundaries ($p = 0.083$). These shifts were consistent with our behavioral results, showing that location script

priming has an effect on the timing of neural event boundaries and that these boundaries become more aligned with event transitions relevant to the location-related script.

In the angular gyrus, for the no-prime baseline, neural event boundaries were significantly matched to the event boundaries of the social script ($p = 0.027$) but not the location-script boundaries ($p = 0.53$). Social priming preserved this temporal structure (location boundaries, $p = 0.26$; social boundaries, $p = 0.043$) while location priming disrupted it (location boundaries, $p = 0.37$; social boundaries, $p = 0.67$). The alignment of the location-primed group to social boundaries was significantly worse than for the no-prime and social-prime groups ($p = 0.016$). In the parahippocampal cortex, we observed only marginal alignment without priming (location boundaries, $p = 0.058$; social boundaries, $p = 0.156$) and no significant alignment for social priming (location boundaries, $p = 0.33$; social boundaries, $p = 0.32$), but the location-primed participants did show significant neural event changes at the location-related narrative boundaries (location boundaries, $p = 0.024$; social boundaries, $p = 0.60$).

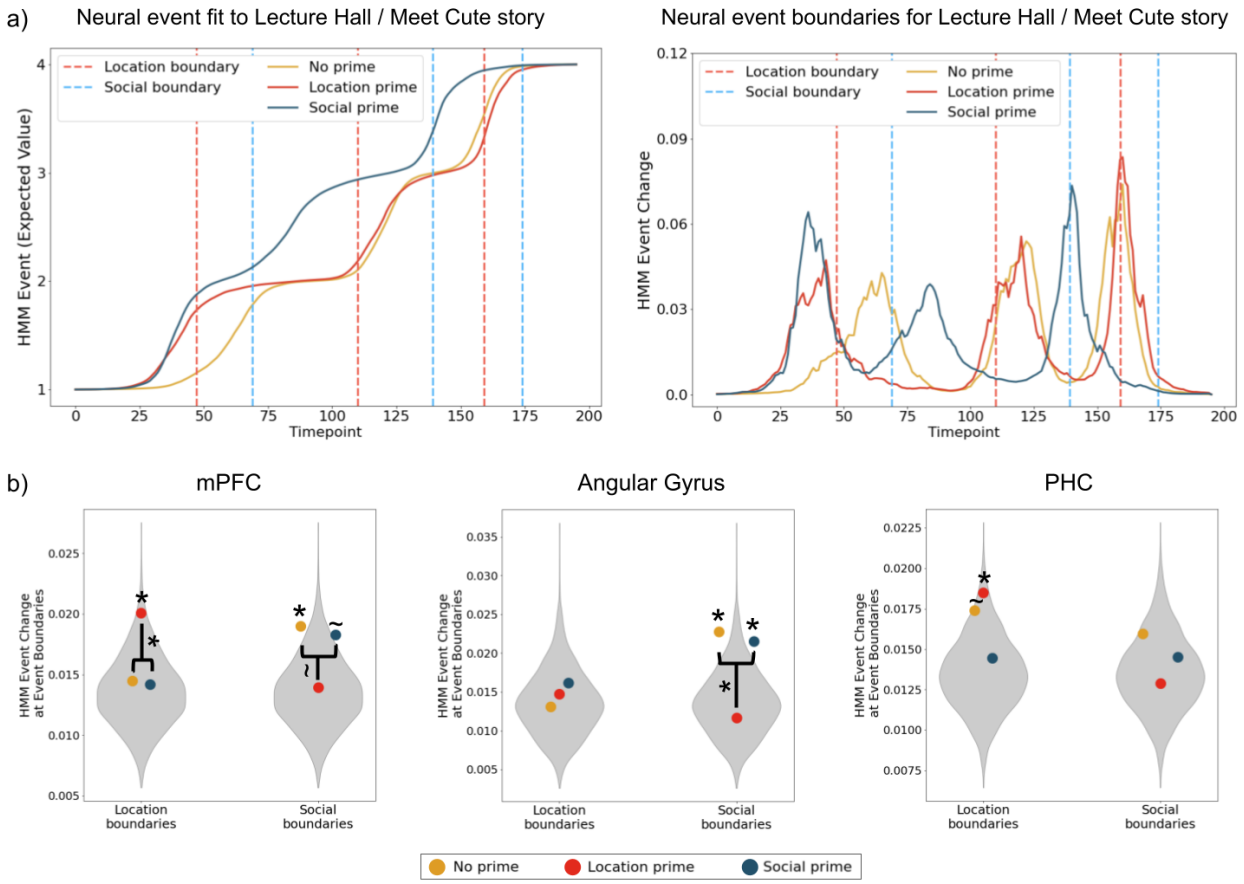


**Figure 5: Event boundaries across priming groups in the angular gyrus, PHC, and mPFC.** a) HMM fits for the three priming conditions in PHC for an example story. The left plot shows how each group transitions through four events during the story, and the right plot shows the rate of change of this event index. Spikes in this second plot indicate moments at which neural responses are rapidly transitioning to a new stable pattern. We compared the alignment between these neural transitions and the event boundaries in the narrative relevant to each of the two scripts. b) In the mPFC, the no-prime group was significantly aligned to the social boundaries, and the location-prime group was significantly aligned to the location boundaries. In the angular gyrus, we observed significant alignment to social boundaries in the no prime and social prime conditions, but not when participants were location primed. In the PHC, there was significant alignment only to the location boundaries in the location-primed group.
~ $p < 0.10$, * $p < 0.05$

Finally, we conducted a model-free exploratory analysis to test for any systematic changes in neural responses across priming groups. We computed a normalized measure of across-group Response Similarity [5] that compares the inter-group correlations to within-group correlations, testing whether there was a consistent difference in the timecourse of activation due to priming (while controlling for possible differences in across-participant variability). Although this approach provides more limited statistical power than our targeted tests of event boundary shifts above, it allows us to identify brain regions whose responses were strongly impacted by attentional priming.

We first compared the no-prime group to the average of both priming groups, and identified numerous areas that showed significant differences (significantly diminished Response Similarity) including regions in the intraparietal sulcus, the STS, mPFC, and early visual cortex (Fig 6a). We additionally performed a small volume correction to look for an effect specifically in our three regions of interest. We found that there were significant differences in Response Similarity for primed vs. unprimed participants only in the mPFC (Fig 6b).

We next tested for prime-specific responses during story-listening. We computed the Response Similarity between the participants who received location vs. social priming for each story, and averaged across stories. No voxels in the searchlight analysis survived FDR correction when considering the whole cortical surface, but after performing a small volume correction in our three regions of interest we observed significant effects within angular gyrus and PHC (Fig 6c). The narrative responses in these areas were therefore sensitive to which prime was activated, further demonstrating a causal impact of event scripts on neural representations.

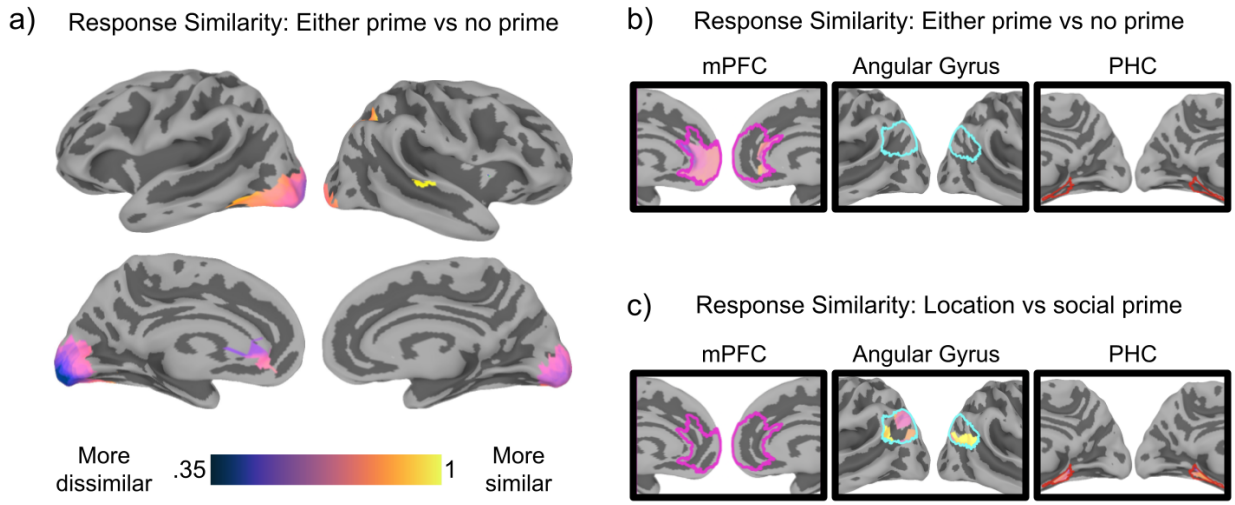


**Figure 6: Neural activation during story listening differs with prime.** a) Computing the Response Similarity between story-evoked activation in the no prime group and both prime groups revealed significant dissimilarity in parts of the intraparietal sulcus, STS, mPFC, and early visual cortex. Strong priming-related effects correspond to lower values in this analysis, indicating a lower amount of similarity between conditions; colored regions are significant at $q < 0.05$. b) Small volume corrections in the angular gyrus, PHC, and mPFC ($q < 0.05$ within each region) revealed significant dissimilarities in the mPFC between listeners without a prime vs. either a location or social script prime. c) When we computed Response Similarity between participants that received two different script primes for a story, we found that angular gyrus and PHC were both sensitive to specific script activated during narrative perception ($q < 0.05$ within each region).

# Discussion

Using narratives that combined multiple kinds of event scripts, we found that scripts related to both physical location and social interactions influenced neural event representations of core event-perception regions in the Default Mode Network (DMN). Additionally, we used multi-script narratives as a tool to understand the top-down impact of script activation on perception and memory. By biasing attention to one of the scripts in the narrative, we showed that script activation changes behavioral outcomes, influencing subjective judgements of event boundaries and changing which episodic details are remembered. We also observed impacts of script activation in the brain, including event boundaries shifting to prime-relevant timepoints in regions including mPFC. These results argue for a view of event representation and segmentation as an active process, in which schematic event knowledge plays a key role in the dynamics of mental and neural representations during narrative perception.

## The Default Mode Network flexibly represents multiple script types

While we found that script information was broadly distributed throughout the cortex, representations of *both* location and social scripts were focused specifically on regions in the DMN, which includes mPFC, PHC, angular gyrus, and STS [26]. Though originally understood as a network engaged only for internally-driven cognition [27], recent work has instead framed the DMN as integrating stimulus-driven signals with idiosyncratic internal knowledge about the world such as memories, schemas, and goals [18]. Our results support this view of the DMN as a convergence zone where high-level event representations are constructed using both bottom-up and top-down information.

Earlier work found that two location-based scripts (restaurants and airports) were represented in the DMN [12], but it was unclear whether the DMN represented scripts in a domain-general way; some parts of the DMN overlap with regions known to represent scene

content and geometry [28], which might support tracking of location-based scripts only. Using a stimulus set with four times as many script types, from two different categories, our current findings instead argue that the DMN can support generalized information about event sequences. Among our *a priori* ROIs, we saw significant representation of both script types in the angular gyrus, which has been proposed as a domain-general buffer for maintaining event information over time [29], and in the PHC, which has been shown to perform contextual associative processing in both spatial and nonspatial tasks [30].

Using this library of "building blocks" from prior experience (and combining them with episode-specific details), the DMN is able to create a nuanced representation of an event during perception and reconstruct this representation during later retrieval [31]. Flexibly combining and remixing knowledge across domains allows us to rapidly comprehend complex events, even when the information in the stimulus is highly restricted [24]. Of course, the event scripts used in this study are still only a small fraction of the full set of schemas employed in narratives, and were selected so that they would likely to be shared across the individuals in our sample; future work could examine a broader basis set of event schemas and account for idiosyncratic knowledge and expertise of specific participants.

## The social script effect in the Hippocampus

We found that the hippocampus was sensitive to the type of social script in a narrative, suggesting that it plays a role in tracking and/or encoding information about sequential social interactions. While the hippocampus has traditionally been thought to be primarily oriented toward representations of physical space [32], it has also been hypothesized to represent more abstract structures like social information [33] and sequence order [34]. The most general version of this hypothesis is that the computations in the hippocampus could support any kind of experience with sequential relational parts, compressing such instances to represent their latent structure [35]. Our results support this kind of domain-general function of the hippocampus, with

even abstract social sequences (such as the stages of a business negotiation) being tracked in a latent cognitive map. Note, however, that in both this study and in our previous work [12], we did not observe significant tracking of location-related scripts in the hippocampus; this null finding could be due to difficulties in measuring hippocampal activity with fMRI, due to acquisition difficulties in the medial temporal lobe [36] or highly sparse representations [37], or could indicate that the hippocampus monitors only a subset of event scripts.

Another possibility is that the hippocampus itself does not track trajectories through social scripts, but instead reflects goal-directed attentional states induced by other brain regions. Manipulating top-down attention has been shown to influence activity patterns in the hippocampus, stabilizing patterns of activation that are specific to particular goals [38]. These hippocampal states could be driven by interactions with retrosplenial cortex [38] and mPFC [39], both of which were sensitive to social script events in our study. A previous study also found largely intact narrative responses in a hippocampal amnesic patient [40], suggesting that the hippocampus is not required for activating narrative schemas.

## Priming-driven changes in neural responses and memory for details

Previous work using naturalistic stimuli has shown that attentional priming can impact representations during encoding. Priming participants with perspectives to elicit a social schema or a non-social schema impacted activation in angular gyrus and PHC [22], and manipulating participant's initial beliefs before presenting an ambiguous story led to interpretation-specific responses in the DMN [23]. Our findings extend this work by including a larger variety of priming conditions (eight possible scripts, as well as a no-prime baseline), and, critically, by using multi-stage temporal scripts rather than static perspectives. Consistent with these prior studies, we observed the biggest impact of priming in angular gyrus and PHC, with significantly dissimilar responses between participants attending to the location versus social script of a narrative. We also detected priming-related impacts in mPFC, with an attentional prime (of

either script type) significantly shifting the baseline no-prime responses in this region. These results show that script-related responses previously observed in the DMN [12,13] do not simply reflect features present in the stimulus, but can also be activated or adjusted by top-down goals.

Script priming also influenced the story-specific details that participants could later recall. In both our online study and our fMRI study, participants were significantly more accurate at responding to questions about details that were related to their attended script. This replicates classic work on schema priming, showing that an activated schema can help direct attention to relevant information and provide an organizational scaffold for forming memories [21]. Although schemas can also play a role during recall by providing retrieval cues [41], in our design we prompted participants with the specific questions from both scripts during the memory test; the observed differences are therefore likely driven only by encoding-time processes.

## Scripts impact subjective boundary judgments and neural event boundaries

We found that biasing attention to the location scripts in a multi-script narrative caused significant shifts in subjective event boundaries. Participants who were location primed had segmentations that were significantly more aligned to the moments in the narratives where there was a transition in the location-script event sequence. Additionally, location primed participants segmented the story more similarly to each other compared to random pairs of participants. These results are consistent with previous work showing that an attentional focus on spatial information can impact annotated boundaries [42], demonstrating that this shift can also be driven by multi-stage location scripts.

Our results did not show a significant impact of social-script priming on annotated boundaries, and other work using the classic "burglar" and "home buyer" social scripts [41] similarly did not find evidence for greater alignment between annotated boundaries and prime-relevant story boundaries [43]. This suggests that boundaries may be particularly sensitive to attentional manipulations of spatial or location-related information, though the reasons for this

are unclear. One possibility is that the event boundaries in location scripts tend to be unambiguous (e.g. entering the checkout line at the grocery store) and therefore easier for participants to consistently align to, compared to event shifts in social situations which could be more graded. Another possibility is that there is a default orientation toward social events which occurs across all priming conditions, and the script priming therefore does not increase the salience of these socially-relevant boundaries. This second possibility is supported by our finding (in both the online and fMRI experiments) that participants performed better in general when answering socially-relevant questions, across priming conditions.

Using a Hidden Markov Model to identify neural transitions between stable states of activity [3], we found that location-script priming caused shifts in neural event boundaries that mirrored those seen in our online experiment. Location-primed participants exhibited boundaries that were significantly more aligned to location-relevant narrative boundaries (in mPFC) and less aligned to socially-relevant boundaries (in angular gyrus). These results provide the first evidence that activating an event script can have a top-down impact on the timing of neural pattern shifts in the DMN. Although event boundaries are often thought of as inherent features of a stimulus, and therefore predictable from the stimulus content alone [44,45], our findings instead provide evidence that event boundaries are actively constructed in the mind and depend on the prior knowledge and current goals of an individual. Rather than simply inheriting the temporal dynamics of the environment, event segmentation is a cognitive process that optimizes the organization of continuous experience into the most relevant units. Individual differences in the segmentation process can have impacts on what we remember [46] and how information is temporally organized [47].

Our findings shed new light on how event representations are constructed in the Default Mode Network from a broad set of prior knowledge, and how the activation of a schema for an event sequence can impact neural responses, perceived event boundaries, and subsequent memory for details.  Our unique stimulus set allowed us to investigate how overlapping scripts

can interact to support narrative perception and how attention can restructure the temporal dynamics of mental representations. These results identify mechanisms by which past experiences, distilled into schematic event scripts, change the way that we construct our present perceptions for realistic experiences.

## Acknowledgments

Thank you to Will Reblando for assistance in graphical design of the figures, Max Rosmarin for providing the voice recordings for the story stimuli, and to all the members of the Dynamic Perception and Memory Lab and Aly Lab for their suggestions and support. This work was supported by a Columbia University Lenfest Junior Faculty Development Award, to C.B.

## Author contributions

Alexandra De Soares: Conceptualization, Data curation, Investigation, Visualization, Software, Writing - original draft

Tony Kim: Data curation, Investigation, Writing - review & editing

Franck Mugisho: Data curation, Investigation, Writing - review & editing

Elen Zhu: Data curation, Investigation, Writing - review & editing

Allison Lin: Data curation, Investigation, Writing - review & editing

Chen Zheng: Data curation, Investigation, Writing - review & editing

Christopher Baldassano: Conceptualization, Funding acquisition, Methodology, Project Administration, Supervision, Visualization, Software, Writing - review & editing

# Methods

## fMRI experiment

### Participant details

We collected data from a total of 38 subjects (21 female, ages 19 - 32 years). All participants were fluent in English, and 32 were native English speakers. All participants had normal or corrected-to-normal vision. The experimental protocol was approved by the Institutional Review Board of Columbia University, and all participants gave their written informed consent. Two participants were excluded due to excessive motion in the scanner (more than 60% of recall timepoints identified as motion outliers, as described below), yielding a final sample of 36 (our target sample size). This sample size was chosen based on prior experiments examining event boundaries with fMRI [4,12] and to ensure that each of the 36 possible combinations of priming conditions (see "Experimental design" below) was assigned to one participant.

### Stimuli

Each narrative stimulus included events reflecting one “location script” and one “social script”. Location scripts were all common sequences of activities that occur in a specific kind of place to accomplish a specific goal: eating at a restaurant, catching a flight at an airport, shopping at a grocery store, and attending a lecture. Social schemas were all common sequences of events for changes in relationships and were not constrained to a particular location: initiating a breakup, proposing marriage, conducting a business deal, and experiencing a “meet cute”. We expected participants to be familiar with these schemas either from their everyday experiences or from media exposure. Each script consisted of a temporal sequence of four events (see Table 1). We wrote 16 stories, each consisting of a unique combination of one

of the four location scripts and one of the four social scripts. Each story proceeded through the four events of each script in parallel. To allow us to behaviorally and neurally distinguish between responses related to location event boundaries versus social event boundaries, the event onsets for the two scripts were interleaved and never coincided in the same sentence (except for the very first sentence, which started the first event for both scripts). Each story had a unique set of characters and a unique setting, such that the only similarities between stories came from overlap in their event scripts.

Stories were presented as audio narratives and were on average 3 minutes and 45 seconds long. All stories were read by the same professional voice actor. While participants listened to the story they were presented with a still screen displaying the title of the story and the pictures (from the 10k US Adult Faces Database [48]) and names of the two main characters for that story.

| | Event 1 | Event 2 | Event 3 | Event 4 |
|---|---|---|---|---|
| **Location Scripts** | | | | |
| **Restaurant** | Entering restaurant | Being seated | Ordering food | Food arriving |
| **Airport** | Entering airport | Going through security | Getting to boarding gates | Boarding plane |
| **Grocery Store** | Entering grocery store | Picking out food | Standing in checkout line | Purchasing groceries |
| **Lecture Hall** | Entering lecture hall | Introducing the lecture | Teaching class | Dismissal and packing up |
| **Social Scripts** | | | | |
| **Breakup** | Preparing for breakup | One party asks to break up | Initial response to the breakup | Accepting the breakup |
| **Proposal** | Preparing for the proposal | Looking at the ring | Proposing | Accepting and celebrating |
| **Business Deal** | Greeting | Making the initial offer | Making a rebuttal | Concluding the deal |
| **Meet Cute** | Initially noticing each other | Initiating the interaction | Proposing to meet again later | Saying goodbye |

**Table 1: Location and social scripts and their events.** Each location and social script has a stereotyped four-event sequence which was consistent in each story that contained that script.

## Script priming

In order to prime participants, we asked them to take on a perspective pertaining to the script, inspired by previous work on schema priming [21,22,41]. To encourage attention to the specific four-event sequence of each script, we provided participants with a sequence of four script-specific questions they would need to answer at the end of the story. Each of these questions corresponded to one of the script events, as shown in Table 2.

Participants were primed with two of the location scripts and two of the social scripts. Before the scanning session, each participant was shown the four perspectives they would be asked to take on. Then for each perspective, they were taught the four script-specific questions. Participants then had to correctly identify and order the questions for each script from six possible answers, with the correct sequence being re-shown after each attempt. After completing the training on all four perspectives, they were asked to reorder the questions for each perspective again without being shown the correct sequence. When participants were in the scanner and about to listen to a story relevant to one of their primes, they would be told which perspective to take on, and had to correctly identify and order the four script-specific questions before moving onto the story.

| | Question 1 | Question 2 | Question 3 | Question 4 |
|---|---|---|---|---|
| **Location Scripts** | | | | |
| **Restaurant** | How is the restaurant decorated? | What are the menus like? | What does each person order? | How do the clients like the food? |
| **Airport** | When the clients arrive at the airport, how much time do they have until their flight departs? | What is the reason for the hold-up at security? | Toward which gate are the clients walking? | What section and seat does each client sit in on the plane? |
| **Grocery Store** | What is the grocery store like upon entering? | What items do the clients pick out to buy? | How many checkout lanes are open and which one do the clients step into? | How much are the groceries and what method of payment do the clients use? |
| **Lecture Hall** | What is the lecture hall like? | What class are the students in and what is the day's lecture about? | What is something taught in lecture? | What is the next assessment/ assignment for the class, and when is it scheduled/due? |
| **Social Scripts** | | | | |
| **Breakup** | For how long has the initiator of the breakup been thinking about breaking up with his/her partner? | What is the initial reason stated by the initiator for why he/she is breaking up? | Does the person who is being broken up with want to break up, and what's the reason stated by them for this? | Who wants what items back as a result of the breakup? |
| **Proposal** | For how long has the couple been dating? | How many diamonds are on the ring, and | In/on what item is the ring presented? | Who does the new fiancee text first? |

| | | | | |
|---|---|---|---|---|
| | | what is the diamond color? | | |
| **Business Deal** | What is each business person's job title? | How much money is at stake in the initial offer made? | What is the name of the other industry competitor? | With what gesture do the business partners secure the deal? |
| **Meet Cute** | What object is the initiator of the interaction holding when he/she first notices the other person? | What is the initial question that begins the conversation between the couple? | When will be the next time the couple meets and for what occasion? | What time is it when the couple parts? |

**Table 2: Location and social scripts questions.** Priming for each location and social script involved learning a sequence of four questions that corresponded to each event in the script.

## Data collection and processing

Data were collected on a 3T Siemens Prisma scanner with a 64-channel head/neck coil. Functional images were obtained with an interleaved multiband EPI sequence (TE = 30 ms, flip angle = 62º, multiband = 3, FOV = 205 mm x 205 mm, 48 oblique axial slices), resulting in a voxel size of 2.5 mm isotropic and a TR of 1s. Whole-brain high resolution (1.0 mm isotropic) T-1 weighted structural images were acquired with an MPRAGE sequence.

Results included in this manuscript come from preprocessing performed using *fMRIPrep* 1.5.6 [49,50] (RRID:SCR_016216), which is based on *Nipype* 1.4.0 [51,52] (RRID:SCR_002502).

## Anatomical data preprocessing

The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection [53], distributed with ANTs 2.2.0 [54] (RRID:SCR_004757), and used as

T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a *Nipype* implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR_002823) [55]. Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, RRID:SCR_001847) [56], and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (RRID:SCR_002438) [57]. Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.2.0), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: *ICBM 152 Nonlinear Asymmetrical template version 2009c* [58] (RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym).

### Functional data preprocessing

For each BOLD run (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. A deformation field to correct for susceptibility distortions was estimated based on *fMRIPrep*'s *fieldmap-less* approach. The deformation field is that resulting from co-registering the BOLD reference to the same-subject T1w-reference with its intensity inverted [59,60]. Registration is performed with antsRegistration (ANTs 2.2.0), and the process regularized by constraining deformation to be nonzero only along the phase-encoding direction, and modulated with an average fieldmap template [61]. Based on the estimated susceptibility distortion, a corrected EPI (echo-planar imaging) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration

[62]. Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9) [63]. The BOLD time-series were resampled to surfaces on the following spaces: *fsaverage6*. The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as *preprocessed BOLD in original space*, or just *preprocessed BOLD*. The BOLD time-series were resampled into standard space, generating a *preprocessed BOLD run in ['MNI152NLin2009cAsym'] space*. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* [64]. The three global signals are extracted within the CSF, the WM, and the whole-brain masks. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives for each [65]. Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels [66]. Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

Many internal operations of *fMRIPrep* use *Nilearn* 0.6.1 [67] (RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIPrep*'s documentation.

After *fMRIPrep*, the data (now in fsaverage6 and MNI152 space) was further preprocessed by a custom python script that: removed from the data (via linear regression) any variance related to the six degrees of freedom motion correction estimate and their derivatives, mean signals in the CSF and white matter, motion outlier timepoints (defined above), and a cosine basis set for high-pass filtering w/ 0.008 Hz (125s) cut-off; z scored each run to have a zero mean and SD of 1; and divided the runs into the portions corresponding to each stimulus. All subsequent analyses, described below, were performed using custom python scripts and the Brain Imaging Analysis Kit (http://brainiak.org/) [68].

### ROI and searchlight definition

We used ROIs in the default mode network previously found to be responsive to schematic content [12]: angular gyrus (1868 vertices), superior temporal sulcus (STS, 2118 vertices), superior frontal gyrus (SFG; 2461 vertices), medial prefrontal cortex (mPFC; 2069 vertices), parahippocampal cortex (882 vertices), and posterior medial cortex (PMC; 2495 vertices). These ROIs were originally derived from a resting-state network atlas on the fsaverage6 surface [69]. We also used, as a control region, a previously-defined auditory cortex region (1589 vertices), which we expected to be largely insensitive to semantic properties of narratives [12]. Additionally, to extract the hippocampus as an ROI, we used the freesurfer subcortical parcellation provided in the *fmriprep* outputs, and included voxels which were common to at least 80% of participants (785 voxels).

Searchlight ROIs were defined as circular regions on the cortical surface, by identifying all vertices within 11 edges of a center vertex along the fsaverage6 mesh. Since the average edge length between vertices is 1.4mm, searchlights had a radius of approximately 15mm. We defined a circular searchlight around every vertex on a hemisphere, and then iteratively removed the most redundant searchlights (i.e. those whose vertices were covered by the most other searchlights). We stopped removing searchlights when doing so would cause some vertices to be covered by fewer than six searchlights. This yielded approximately 1000 searchlights on each hemisphere.

## Experimental design

We first tested the audio presentation and recording equipment by playing a short audio clip (unrelated to the narratives) to verify that the volume level was set correctly and by asking participants to talk about their breakfast that day into the microphone. Subjects were then presented with 6 stories in each of 2 experimental runs, using PsychoPy (RRID:SCR_006571) [70]. Participants were each primed with two of the four location scripts and two of the four social scripts; this produced C(4,2)*C(4,2)=36 possible priming conditions, with each condition assigned to a different participant. Each participant listened in total to 12 stories: 4 stories that used the primed location scripts but not the primed social scripts; 4 stories that used the primed social scripts but not the primed location scripts; 4 stories that used neither of the primed location or social schemas. Participants did not hear the remaining 4 stories for which both the location and social scripts had been primed. Story order and priming were shuffled in each block. If a story contained one of the primed event scripts, participants were reminded before the story of the perspective they were supposed to take on and took a quiz to identify and correctly order the four script-relevant questions. Participants would not proceed to the story until they correctly ordered the questions. At the onset of the story, the story image would

appear on a gray background followed by 5 seconds of silence. When the story finished, the story image would disappear followed by an additional 5 seconds of silence. Subjects were not explicitly told that they would be asked to recall the stories. After participants listened to the 12 stories, they were then asked to recall stories in the same order as they listened (data not reported in this paper).

After recall, participants were asked outside of the scanner to answer 8 short answer questions about each story. These questions were the script-specific questions for the social and location perspectives relevant to that story.

### Short Answer Priming Effect Analysis

We scored the short answers of all participants on a scale from 0-3 based on a detailed rubric, where 0 is either no answer or a completely wrong answer and 3 is a perfectly correct answer. We then computed the average location question score by averaging the ratings of all location questions, and in the same way we computed an average social question score. We ran a mixed effects model with prime, question type, and their interaction as fixed effects, and a subject-specific random intercept. For each question type, we also tested whether there was a significant difference in average score between the two types of priming by running a mixed effects model using only the social-primed and location-primed conditions.

### Location and Social Script Effect Analysis

To measure the extent to which event patterns generalized across stories with the same location script in a brain region, we first computed for each participant the average spatial activity pattern for each of the four events of each story (defined according to the putative location-related event boundaries). For all pairs of stories, we correlated all four event patterns

in one story for one participant with the four event patterns in the second story averaged across all other participants. We averaged the four correlations between corresponding events (e.g. event 1 in story 1 with event 1 in story 2) and subtracted the average of all other event correlations as a baseline. We iteratively held out each participant, recomputing the 16 story x 16 story event correlation matrix for each held-out participant, and then averaged these matrices to produce a single event correlation matrix. To compute the location script effect, we compared the average event correlation for pairs of stories with shared location scripts to the event correlation for pairs of stories with no shared scripts. If the same-script correlations are larger, this indicates that the location script of a story influences the event patterns in a brain region. We then repeated this entire process for the social scripts, computing event patterns according to the social-relevant boundaries and comparing event similarity for stories with shared social scripts versus no shared scripts.

We used permutation testing to assess statistical significance. After computing the location and script effects, we repeated the process 1000 times, randomly shuffling the identities of the 16 stories each time. For our *a priori* ROIs, we computed p values directly as the fraction of the 1001 effect values (including the real value) that were at least as large as the real value. For the searchlight analysis, in order to better estimate small p values, we approximated the distribution of null values using a Normal distribution, and converted the real effect value into a z score relative to this distribution (i.e. the number of standard deviations by which the real value exceeded the mean of the null distribution). A z value was computed for each surface vertex as the average of z values from all searchlights that included that vertex. The vertex z values were then converted to p values using the survival function of the Normal distribution, and finally converted into a map of q values using the same false discovery rate correction that is used in AFNI [71].

## Hidden Markov Model Analysis

In order to test if priming had an effect on neural event boundaries, we used a Hidden Markov Model (HMM) approach [3] to compare participants' neural event boundaries to the putative location and social event boundaries in the narratives. For a given brain region and story, we first averaged the spatiotemporal pattern of brain activation within each of the three priming groups. We then fit an HMM with 4 events to the average pattern of activation of each group, producing a probabilistic estimate of which of the four events was active at each timepoint. We used this probability distribution to compute a measure of event change at each timepoint [72]; we computed the expected value of the active event at each timepoint, then took its temporal derivative. To assess the match between the neural event shifts and the putative event boundaries of the location and social scripts in each story, we computed the average event change at the location boundaries and the average at the social boundaries, and finally averaged each of these values across all stories for each priming group. We created a null distribution for this alignment by randomly assigning the location and social boundaries to timepoints and recomputing the average event change for each story, 1,000 times. The p value for each priming group was computed as the fraction of the 1,001 null and real values that were at least as large as the real value. We tested whether there were significant differences in boundary alignment between the location and non-location priming groups by measuring the difference between the location priming group and the average of the social and no prime groups, producing a p value in the same way (comparing the difference for the real values to the differences in the null distribution).

## Response Similarity Analysis

To test whether priming produced a significant shift in the timecourse of activation during story listening, we first divided the 27 participants that heard a story into their respective priming

conditions (‘no prime’, ‘location prime’, and ‘social prime’), yielding 9 participants in each condition for each story. We then split each priming group into a group 1 of four participants, and a group 2 of four participants, leaving out one participant. We averaged the timecourse of activation within each four-participant group, and correlated the group 1 timecourse in each priming condition with the group 2 timecourse in each priming condition. We averaged these results across 10 different random splits of the groups, yielding a 3 x 3 matrix of correlations between the three priming conditions, and then symmetrized this matrix by averaging it with its transpose. We repeated this full process for all stories and averaged the matrix across stories.

If priming produces a systematic shift in responses, we would expect that the diagonal correlations in this matrix (between groups with the same prime) should be larger than the off-diagonal correlations (between groups with different primes). We mathematically compared these values using a Response Similarity approach derived in our previous work [5]; specifically, we can estimate the similarity between two conditions while accounting for the within-condition variance by taking the across-group correlation and dividing by the geometric mean of the within-group correlations. If this ratio is significantly smaller than 1, this indicates that there are systematic differences between the groups' mean responses.

We first computed the Response Similarity for the no-prime group compared to both priming groups, averaging across the location and social priming groups. We also computed the Response Similarity for the location-primed group compared to the social-prime group. We performed permutation tests by shuffling the priming conditions of the 27 participants in a story 1,000 times. For computational efficiency in the searchlight analysis, we first performed only 100 permutations in each searchlight, and then performed the rest of the 1,000 searchlights only if at least 95% of the null Response Similarities were less than the real Response Similarity. To generate the searchlight map, real values and 1,000 permutation values were computed for each vertex by averaging those values from all searchlights that included that vertex. A p value was computed for each vertex as the fraction of the 1,001 Response Similarity values (including

the real value) that were at least as small as the real value. These were converted to *q* values using the same false discovery rate in AFNI [71], across all vertices on the cortical surface or as a small volume correction within selected *a priori* ROIs (mPFC, PHC, angular gyrus).

## Online behavioral experiment

### Participant Details

Data were collected online using Prolific (www.prolific.co). We collected data from 387 participants (181 females, age 18 - 75) who were 96% native English speaking and we compensated them $3.17- $5. We excluded 10 participants who did not answer the short answer questions following the story, resulting in a final sample of n=377. For the short-answer analyses, we additionally removed 2 participants whose short-answer data was not correctly recorded (yielding n=375), and for the event-boundary analyses, we additionally excluded 28 participants who did not mark any event boundaries and 41 participants who said that more than half of the sentences were event boundaries (yielding n=308). This sample size was chosen to ensure that at least 5 participants were assigned to each of the 3 possible primings of each of the 16 stories. The experimental protocol was approved by the Institutional Review Board of Columbia University.

### Stimuli

Stimuli were presented using Pavlovia (pavlovia.org). The stimuli in the online experiment were the same stories as in the fMRI experiment. However, rather than listening to the story continuously, participants listened to each sentence individually. When each sentence had finished, participants indicated if the just-presented sentence was the beginning of a new

“part” of the story or not, the story then proceeded to the next sentence. After participants listened to the story and made judgments on all sentences, they performed a brief distractor task where they were asked to remember difficult-to-spell words, and then they were asked the same 8 script-specific questions for the two scripts relevant to the story in the same manner as in the fMRI study. Finally participants typed what they recalled about the story (data not analyzed in this paper), and filled out a demographics questionnaire about in Qualtrics (www.qualtrics.com).

### Priming

Participants were randomly assigned to hear one story, and were randomly assigned to either be not primed, primed with the location script, or primed with the social script. Priming was done in the same manner as in the fMRI study, with participants asked to “take on” a perspective while listening to the story. Participants were first shown the four script-specific questions (Table 2) in the correct order, and were then given the quiz where they had to select the questions in the right order. If they made an incorrect choice they would be shown the correct sequence again and then had to re-take the quiz. Once they had correctly ordered all the questions, they would be asked to perform the ordering correctly a second time; if they made a mistake, they were re-presented with the quiz until they produced the correct ordering.

### Short Answer Priming Effect Analysis

Four raters scored the short answers of all participants, using the same scoring system as for the fMRI participants. We fit a mixed effects model as for the fMRI participants, and tested for a significant difference in average score between priming groups by running a Welch independent-samples t-test between groups for each question type.

### Comparison to putative event boundaries

We computed the Jaccard similarity between each participant's segmentation and the putative location and social boundaries of each story. These putative boundaries were the sentences that were designed to indicate a shift to the next event in the location or social script (see Table 1). In order to test whether participants in each priming group had segmentations that aligned with these event boundaries, we computed the Jaccard similarity index between each participant's segmentations and the putative boundaries from both scripts. We computed the average similarity in each priming group, and compared to a null in which between-boundary durations were preserved but the order of event durations was shuffled 10,000 times. We computed a p value as the fraction of agreements in the 10,001 null and real values that were at least as large as in the real agreement.

To test for differences in alignment between priming conditions for each type of event boundary, we used a one-way ANOVA to test for differences in mean agreement to the story boundaries between priming groups, and if there was a significant difference we ran Tukey's test to see which means were significantly different.

### Within-group similarity of event boundaries

In order to determine if priming caused participants to have more similar event boundary judgments, we computed the Jaccard similarity index between all pairs of participants within each priming group for each story, and averaged across stories. Permutation testing was used for assessing statistical significance; the priming conditions for the participants in each story were shuffled 10,000 times and within-group similarity was computed for each null dataset, and the p value was computed as the fraction of the 10,001 null and real similarities that were at least as large as the real within-group similarity.

## Data availability

Audio recordings of the story stimuli (separated by sentence) are available at https://figshare.com/projects/Script_Combination_Stories/168656 and fMRI data is available at https://doi.org/10.18112/openneuro.ds004631.v1.0.0. The behavioral data and analysis scripts will be shared upon acceptance for publication.

# References


1. Sonkusare, S., Breakspear, M., and Guo, C. (2019). Naturalistic Stimuli in Neuroscience: Critically Acclaimed. Trends Cogn. Sci., 10.1016/j.tics.2019.05.004.

2. Stephens, G.J., Honey, C.J., and Hasson, U. (2013). A place for time: the spatiotemporal structure of neural dynamics during natural audition. J. Neurophysiol. *110*, 2019–2026.

3. Baldassano, C., Chen, J., Zadbood, A., Pillow, J.W., Hasson, U., and Norman, K.A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. Neuron *95*, 709–721.e5.

4. Williams, J.A., Margulis, E.H., Nastase, S.A., Chen, J., Hasson, U., Norman, K.A., and Baldassano, C. (2022). High-Order Areas and Auditory Cortex Both Represent the High-Level Event Structure of Music. J. Cogn. Neurosci. *34*, 699–714.

5. Cohen, S.S., Tottenham, N., and Baldassano, C. (2022). Developmental changes in story-evoked responses in the neocortex and hippocampus. Elife *11*, 10.7554/eLife.69430.

6. Schank, R.C., and Abelson, R.P. (1977). Scripts, plans, goals and understanding: An inquiry into human knowledge structures (Lawrence Erlbaum Associates, Inc.).

7. Bower, G.H., Black, J.B., and Turner, T.J. (1979). Scripts in memory for text. Cogn. Psychol. *11*, 177–220.

8. Alba, J.W., and Hasher, L. (1983). Is memory schematic? Psychol. Bull. *93*, 203–231.

9. Bransford, J.D., and Johnson, M.K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. Journal of Verbal Learning and Verbal Behavior *11*, 717–726.

10. Huang, J., Velarde, I., Ma, W.J., and Baldassano, C. (2023). Schema-based predictive eye movements support sequential memory encoding. Elife *12*.

11. Zacks, J.M., Speer, N.K., Swallow, K.M., Braver, T.S., and Reynolds, J.R. (2007). Event perception: a mind-brain perspective. Psychol. Bull. *133*, 273–293.

12. Baldassano, C., Hasson, U., and Norman, K.A. (2018). Representation of Real-World Event Schemas during Narrative Perception. J. Neurosci. *38*, 9689–9699.

13. Masís-Obando, R., Norman, K.A., and Baldassano, C. (2022). Schema representations in distinct brain networks support narrative memory during encoding and retrieval. Elife *11*, 10.7554/eLife.70445.

14. Raykov, P.P., Keidel, J.L., Oakhill, J., and Bird, C.M. (2021). Activation of Person Knowledge in Medial Prefrontal Cortex during the Encoding of New Lifelike Events. Cereb. Cortex *31*, 3494–3505.

15. van Kesteren, M.T.R., Beul, S.F., Takashima, A., Henson, R.N., Ruiter, D.J., and Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: from congruent to incongruent. Neuropsychologia *51*, 2352–2359.

16. Tottenham, N. (2020). Neural meaning making, prediction, and prefrontal-subcortical development following early adverse caregiving. Dev. Psychopathol. *32*, 1563–1578.

17. Kanwisher, N. (2010). Functional specificity in the human brain: a window into the functional architecture of the mind. Proc. Natl. Acad. Sci. U. S. A. *107*, 11163–11170.

18. Yeshurun, Y., Nguyen, M., and Hasson, U. (2021). The default mode network: where the idiosyncratic self meets the shared social world. Nat. Rev. Neurosci. *22*, 181–192.

19. Shin, Y.S., and DuBrow, S. (2021). Structuring Memory Through Inference-Based Event Segmentation. Top. Cogn. Sci. *13*, 106–127.

20. Yates, T.S., Sherman, B.E., and Yousif, S.R. (2023). More than a moment: What does it mean to call something an “event”? Psychon. Bull. Rev.

21. Pichert, J.W., and Anderson, R.C. (1977). Taking different perspectives on a story. J. Educ. Psychol. *69*, 309–315.

22. Lahnakoski, J.M., Glerean, E., Jääskeläinen, I.P., Hyönä, J., Hari, R., Sams, M., and Nummenmaa, L. (2014). Synchronous brain activity across individuals underlies shared psychological perspectives. Neuroimage *100*, 316–324.

23. Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C.J., and Hasson, U. (2017). Same Story, Different Story. Psychol. Sci. *28*, 307–319.

24. Nguyen, M., Vanderwal, T., and Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. Neuroimage *184*, 161–170.

25. Wang, Y.C., and Egner, T. (2022). Switching task sets creates event boundaries in memory. Cognition *221*, 104992.

26. Andrews-Hanna, J.R., Smallwood, J., and Spreng, R.N. (2014). The default network and self-generated thought: component processes, dynamic control, and clinical relevance. Ann. N. Y. Acad. Sci. *1316*, 29–52.

27. Raichle, M.E. (2015). The brain’s default mode network. Annu. Rev. Neurosci. *38*, 433–447.

28. Dilks, D.D., Kamps, F.S., and Persichetti, A.S. (2022). Three cortical scene systems and their development. Trends Cogn. Sci. *26*, 117–127.

29. Humphreys, G.F., Lambon Ralph, M.A., and Simons, J.S. (2021). A Unifying Account of Angular Gyrus Contributions to Episodic and Semantic Cognition. Trends Neurosci. *44*, 452–463.

30. Aminoff, E.M., Kveraga, K., and Bar, M. (2013). The role of the parahippocampal cortex in cognition. Trends Cogn. Sci. *17*, 379–390.

31. Chen, J., Leong, Y.C., Honey, C.J., Yong, C.H., Norman, K.A., and Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. Nat. Neurosci. *20*, 115–125.

32. Hartley, T., Lever, C., Burgess, N., and O’Keefe, J. (2014). Space in the brain: how the hippocampal formation supports spatial cognition. Philos. Trans. R. Soc. Lond. B Biol. Sci.

*369*, 20120510.

33. Schafer, M., and Schiller, D. (2018). Navigating Social Space. Neuron *100*, 476–489.

34. Davachi, L., and DuBrow, S. (2015). How the hippocampus preserves order: the role of prediction and context. Trends Cogn. Sci. *19*, 92–99.

35. Raju, R.V., Swaroop Guntupalli, J., Zhou, G., Lázaro-Gredilla, M., and George, D. (2022). Space is a latent sequence: Structured sequence learning as a unified theory of representation in the hippocampus. arXiv [q-bio.NC].

36. Olman, C.A., Davachi, L., and Inati, S. (2009). Distortion and signal loss in medial temporal lobe. PLoS One *4*, e8160.

37. Diamantaki, M., Frey, M., Berens, P., Preston-Ferrer, P., and Burgalossi, A. (2016). Sparse activity of identified dentate granule cells during spatial exploration. Elife *5*. 10.7554/eLife.20252.

38. Aly, M., and Turk-Browne, N.B. (2016). Attention Stabilizes Representations in the Human Hippocampus. Cereb. Cortex *26*, 783–796.

39. Günseli, E., and Aly, M. (2020). Preparation for upcoming attentional states in the hippocampus and medial prefrontal cortex. Elife *9*, 10.7554/eLife.53191.

40. Zuo, X., Honey, C.J., Barense, M.D., Crombie, D., Norman, K.A., Hasson, U., and Chen, J. (2020). Temporal integration of narrative information in a hippocampal amnesic patient. Neuroimage *213*, 116658.

41. Anderson, R.C., and Pichert, J.W. (1978). Recall of previously unrecallable information following a shift in perspective. Journal of Verbal Learning and Verbal Behavior *17*, 1–12.

42. Bailey, H.R., Kurby, C.A., Sargent, J.Q., and Zacks, J.M. (2017). Attentional focus affects how events are segmented and updated in narrative reading. Mem. Cognit. *45*, 940–955.

43. Newberry, K.M., and Bailey, H.R. (2019). Does semantic knowledge influence event segmentation and recall of text? Mem. Cognit. *47*, 1173–1187.

44. Kumar, M., Goldstein, A., Michelmann, S., Zacks, J.M., Hasson, U., and Norman, K.A. (2022). Bayesian Surprise Predicts Human Event Segmentation in Story Listening. 10.31234/osf.io/qd2ra.

45. Michelmann, S., Kumar, M., Norman, K.A., and Toneva, M. (2023). Large language models can segment narrative events similarly to humans. arXiv [cs.CL].

46. Sargent, J.Q., Zacks, J.M., Hambrick, D.Z., Zacks, R.T., Kurby, C.A., Bailey, H.R., Eisenberg, M.L., and Beck, T.M. (2013). Event segmentation ability uniquely predicts event memory. Cognition *129*, 241–255.

47. Jafarpour, A., Buffalo, E.A., Knight, R.T., and Collins, A.G.E. (2022). Event segmentation reveals working memory forgetting rate. iScience *25*, 103902.

48. Bainbridge, W.A., Isola, P., and Oliva, A. (2013). The intrinsic memorability of face photographs. J. Exp. Psychol. Gen. *142*, 1323–1334.

49. Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., Snyder, M., et al. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat. Methods *16*, 111–116.

50. Esteban, O., Blair, R., Markiewicz, C.J., Berleant, S.L., Moodie, C., Ma, F., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., et al. (2018). fmriprep 10.5281/zenodo.1942006.

51. Gorgolewski, K., Burns, C.D., Madison, C., Clark, D., Halchenko, Y.O., Waskom, M.L., and Ghosh, S.S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front. Neuroinform. *5*, 13.

52. Esteban, O., Markiewicz, C.J., Burns, C., Goncalves, M., Jarecka, D., Ziegler, E., Berleant, S., Ellis, D.G., Pinsard, B., Madison, C., et al. (2022). nipy/nipype: 1.8.3 10.5281/zenodo.6834519.

53. Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., and Gee, J.C. (2010). N4ITK: improved N3 bias correction. IEEE Trans. Med. Imaging *29*, 1310–1320.

54. Avants, B.B., Epstein, C.L., Grossman, M., and Gee, J.C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. *12*, 26–41.

55. Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging *20*, 45–57.

56. Dale, A.M., Fischl, B., and Sereno, M.I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage *9*, 179–194.

57. Klein, A., Ghosh, S.S., Bao, F.S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Chaibub Neto, E., et al. (2017). Mindboggling morphometry of human brains. PLoS Comput. Biol. *13*, e1005350.

58. Fonov, V.S., Evans, A.C., McKinstry, R.C., Almli, C.R., and Collins, D.L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. Neuroimage *47*, S102.

59. Wang, S., Peterson, D.J., Gatenby, J.C., Li, W., Grabowski, T.J., and Madhyastha, T.M. (2017). Evaluation of Field Map and Nonlinear Registration Methods for Correction of Susceptibility Artifacts in Diffusion MRI. Front. Neuroinform. *11*, 17.

60. Huntenburg, J. (2014). Evaluating nonlinear coregistration of BOLD EPI and T1w images.

61. Treiber, J.M., White, N.S., Steed, T.C., Bartsch, H., Holland, D., Farid, N., McDonald, C.R., Carter, B.S., Dale, A.M., and Chen, C.C. (2016). Characterization and Correction of Geometric Distortions in 814 Diffusion Weighted Images. PLoS One *11*, e0152472.

62. Greve, D.N., and Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. Neuroimage *48*, 63–72.

63. Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images.

Neuroimage *17*, 825–841.

64. Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., and Petersen, S.E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage *84*, 320–341.

65. Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., et al. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. Neuroimage *64*, 240–256.

66. Lanczos, C. (1964). Evaluation of Noisy Data. Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis *1*, 76–85.

67. Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., and Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. Front. Neuroinform. *8*, 14.

68. Kumar, M., Anderson, M.J., Antony, J.W., Baldassano, C., Brooks, P.P., Cai, M.B., Chen, P.-H.C., Ellis, C.T., Henselman-Petrusek, G., Huberdeau, D., et al. (2021). BrainIAK: The Brain Imaging Analysis Kit. Apert Neuro *1*. 10.52294/31bb5b68-2184-411b-8c00-a1dacb61e1da.

69. Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. *106*, 1125–1165.

70. Peirce, J.W. (2007). PsychoPy--Psychophysics software in Python. J. Neurosci. Methods *162*, 8–13.

71. Cox, R.W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. *29*, 162–173.

72. Lee, C.S., Aly, M., and Baldassano, C. (2021). Anticipation of temporally structured events in the brain. ELife *10*, 10.7554/eLife.64972.