Capturing the geometric structure of episodic memories for naturalistic experiences

Andrew C. Heusser (andrew.c.heusser@dartmouth.edu)
Department of Psychological and Brain Sciences, 6207 Moore Hall
Hanover, NH 03755 USA

Jeremy R. Manning (jeremy.r.manning@dartmouth.edu)
Department of Psychological and Brain Sciences, 6207 Moore Hall
Hanover, NH 03755 USA

Abstract

The human memory system is adept at cataloging the rich dynamics of ongoing experience. However, traditional trial-based memory experiments cannot capture these dynamics, and therefore cannot be used to study them. By constraining participants' experiences in an experiment to occur in temporally discrete trials, often arranged in a randomized order, the temporal, contextual, and emotional structure of those in-lab experiences necessarily differ from the naturalistic experiences we encounter in everyday life. Here we investigate how people verbally recall continuous videos by characterizing and relating the thematic dynamics, or "trajectories," of the stimulus and participants' recalls. Unlike trial-based studies of memory wherein participants attempt to recall the precise stimuli they encounter, naturalistic recall entails capturing the fundamental geometric components of the stimulus topic trajectory. The precise words participants use to describe the stimulus, and the level of detail and number of distinct events they recall, vary considerably across participants. Nevertheless, all of the participants' recall narratives captured the fundamental "shape" of the original stimulus. We view this work as providing a window into which aspects of naturalistic experiences must be preserved, and which might be more flexible, in considering whether and how those experiences are remembered.

Keywords: Episodic Memory; Naturalistic stimuli

Introduction

What does it mean to *remember* something? In traditional episodic memory experiments (e.g., list-learning or trial-based experiments; Murdock, 1962; Kahana, 1996), remembering is often cast as a binary operation: either an item is recalled or it isn't. More nuanced studies might incorporate self-reported confidence measures as a proxy for memory strength, or ask participants to discriminate between "recollecting" an experience or a feeling of "familiarity" (Yonelinas, 2002). However, characterizing and evaluating memory in more realistic contexts (e.g., telling a story to a friend about a recent vacation) is a fundamentally different task. Real-world recall is continuous, rather than binary. The specific words used to describe an experience have very little bearing on whether the experience is considered to have been "remembered." Further, one might remember the gist of an experience but forgot (or neglect to recount) particular details. Or different people who share an experience might recount the experience with a similar level of detail, but the specific details that were remembered might vary across people. Which aspects of those recollections should be considered fundamental, and which are extraneous to the main story?

Another major difference between traditional trial-based memory paradigms and real-world memory concerns the subjective experience of the participant. For example, consider an emotionally salient event in your life (e.g. marriage, the birth of a child, death of a loved one, etc.), or even a movie that you were especially impacted by. No list-learning paradigm (or similar) can hope to capture the nuance and depth of such experiences. Our everyday experiences feel "important" to us, whereas the words we study on a random word list do not. One component of naturalistic experiences that enhances their impact concerns the temporal structure with which they unfold. For example, our experiences in the real world are necessarily autocorrelated in space and time on short timescales. Further, or experiences are often characterized by longer timescale correlations that reflect the impact of our past actions and observations. These temporal dynamics are not typically present in traditional memory studies, but are important if we wish to understand how our memory systems remember our everyday experiences.

To study how people recall naturalistic experiences, we analyzed an open dataset which had 17 participants view and verbally recall (in order)an episode of the BBC series *Sherlock* (Chen et al., 2017). Although the original study included both behavioral and fMRI data, in this paper we have limited our analysis to only the behavioral data.

Characterizing memory for naturalistic stimuli

We developed a novel computational approach for studying memory for naturalistic stimuli. Our method uses Latent Dirichlet Allocation (Blei et al., 2003), a Bayesian model that transforms text into a probabilistic mixture of *topics*. We use this *topic model*, applied to text descriptions of each moment from the movie and transcripts of participants' verbal recalls of the movie, to represent the narrative structure. Specifically, we fit a topic model to overlapping windows of manually annotated text descriptions of scenes from an episode of Sherlock. The text description contained details of the scene such as the

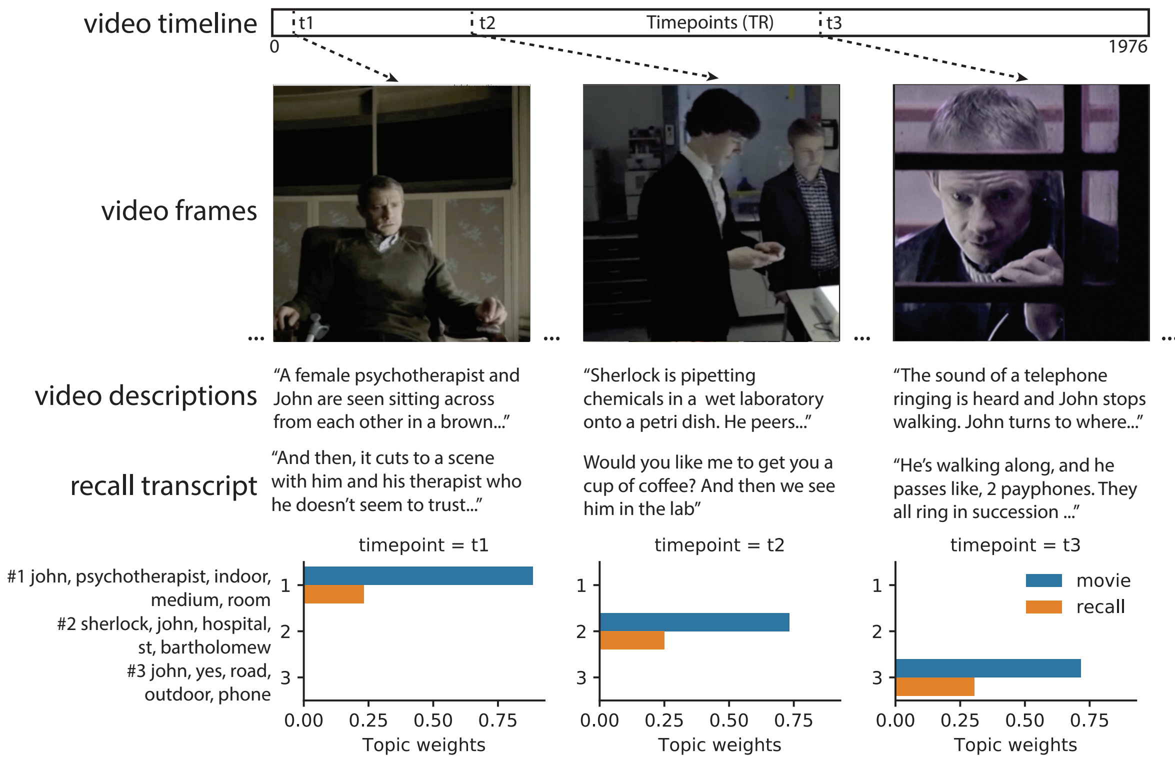

Figure 1: Schematic of the analysis approach. The top rectangle represents a timeline of the video stimulus. For each frame of the video, text descriptions were manually created. Three exemplary frames are displayed here. Below the video descriptions are text samples from an example participant's verbal recall transcript. We trained a topic model on the moment-by-moment video description text and transformed participant's recall transcripts using this same model. The bar charts display the resulting topic model weights for the video (in blue) and recall (in orange) for three example topic dimensions (to the left of the bar charts). The words represent the top 5 words for each example topic.

characters, location, and a short summary of the scene (see Fig. 1 for example text). We then transformed the text descriptions with the (same) topic model, resulting in a scenes (1000) by topics (100) matrix, where each row of the matrix reflects the mixture of topics reflected in that scene's description. We expanded this matrix from 1000 to 1976 timepoints by copying the vectors for scenes that spanned multiple timepoints. (This expansion was performed to match the timing of the fMRI data collected as participants viewed and recalled the movie.)

After watching the episode, participants verbally recalled (in order) as much of the episode as they could. We used the topic model, trained on the scene descriptions of the video, to transform participants' verbal recall transcripts (split by sentence in overlapping chunks of 10 sentences). This yielded a number-of-sentences by (100) topics matrix for each participant, where each row represented the estimated mixture of topics reflected by the corresponding sentences during recall. We resampled the rows of these matrices to ensure that each participant's recall topics matrix had 1976 rows (to match the number of rows of the video topics matrix).

We next computed timepoints (1976) by timepoints (1976) correlation matrices for the video topics matrix and each participant's recall topics matrix. All of these matrices exhibited strong block diagonal structure. In other words, the semantic content (reflected by the topic mixture proportions) exhibited periods of stability, followed by rapid changes, followed by stability. We interpreted the stable periods as *events* and the periods of rapid change as *event boundaries*. Following Baldassano et al. (2017), we used Hidden Markov Models to partition each matrices into k events, where the optimal value of k was determined from the data. Our algorithm found 34 events for the video topic matrix and a range of values (range: 8–27; mean: 15.41; std: 5.6) for the participants' recall topic matrices. The values of k across participants were highly correlated with hand-annotated recall accuracies ascribed to the participants by independent human raters ($r = 0.67, p = 0.003$), suggesting that the algorithm divided the models into meaningful units related to memory performance.

We used the event boundaries identified by the Hidden Markov Model fit to the video topic matrix to create a *video event topic matrix*. Specifically, we averaged the topic vectors within each event to compute an events (34) by topics (100) matrix. We performed the same procedure for the recall matrices to yield events by topics matrices for each participant. We computed the correlation between each event in the video and participant event topic matrices. This yielded a video events (34) by recall events (range: 8–27) correlation matrix for each participant. These participant-wise matrices represent the extent to which each recall event was similar in content to each video event. Finally, we matched each participants' events to the video event with the most correlated topic vector.

To visualize the relationship between the video and recall topics, we embedded the video and average recall model into a 2D space (using the UMAP dimensionality reduction algorithm (McInnes & Healy, 2018)) where the embedded points represent each event and the distances between points reflect the similarity between the corresponding events' topic mixture proportions. We computed the group average transition probabilities from each event to each other event and plotted the values as lines, where the transparency is proportional to the probability of the transition. Visual inspection reveals (and statistical analysis confirms) that the two models have a very similar geometric structure, and the transition probabilities tend to be strongest between events that occurred sequentially (or nearby) in time (Fig 2a). To explore individual variability (and consistency) in the recall event models, we embedded each participant's recall event model in the same 2D topic space described above (Fig 2b). Despite the variability in the number of events remembered, and the precise coordinates of individual recall events, the overall shapes of the participants' recall topic trajectories were strikingly similar to the video topic trajectory (Fig 2b). This reflects the participants' abilities to capture the fundamental conceptual geometry of the video narrative, despite individual differences in the wording and resolution with which that narrative was recounted.

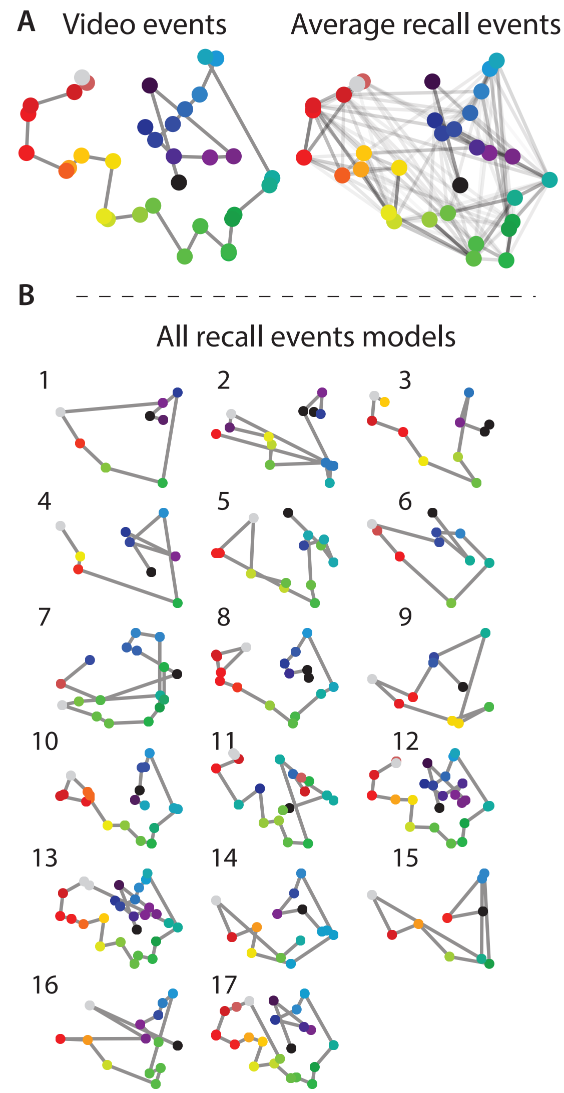

Figure 2: **A.** 2D embeddings (reduced with UMAP dimensionality reduction algorithm) of video and average recall event topic matrices. The colors represent distinct video events (unfolding from black to red). The lines connecting the dots in the average recall matrix represent average transition probabilities, where darker lines indicate a higher probability of transitioning. **B.** 2D embeddings of each participants' recall event topic matrices. Each dot represents a recall event and the connecting lines indicate the order of recall. Colors indicate the most likely video event that the participant was describing as determined by the video-recall matching model.

References

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron*, *95*(3), 709–721.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993 – 1022.

Chen, J., Leong, Y. C., Norman, K. A., & Hasson, U. (2017). Shared experience, shared memory: a common structure for brain activity during naturalistic recall shared experience, shared memory: a common structure for brain activity during naturalistic recall. *Nature Neuroscience*, *20*(1), 115–125.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*, 103–109.

McInnes, L., & Healy, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv e-prints*.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 482-488.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441-517.