Exploring the evolving geometric structure of experiences and memories

Paxton C. Fitzpatrick, Andrew C. Heusser, and Jeremy R. Manning

Dartmouth College

Contextual Dynamics Laboratory

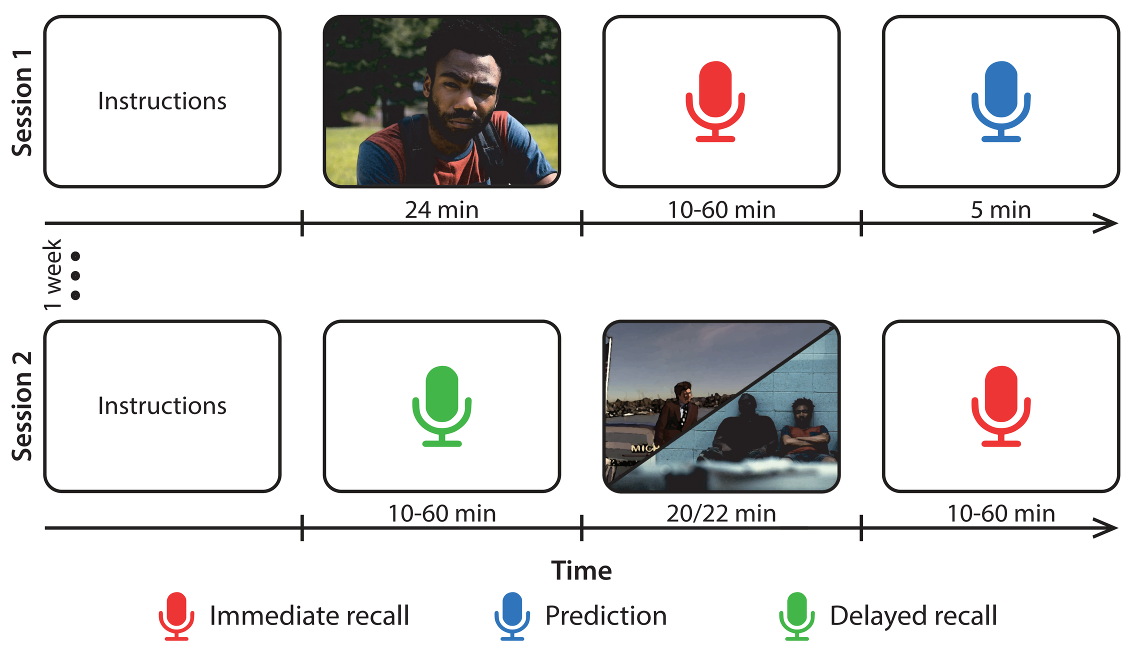
Introduction and overview

How do we transform real-life experiences into memories?
How do our memories change over time?
Why are some experiences more memorable than others?

We model the dynamic evolution of experiences and memories as high-dimensional **trajectories**, and present a novel approach for **predicting and decoding neural representations** of experienced and remembered content. Our findings suggest that:

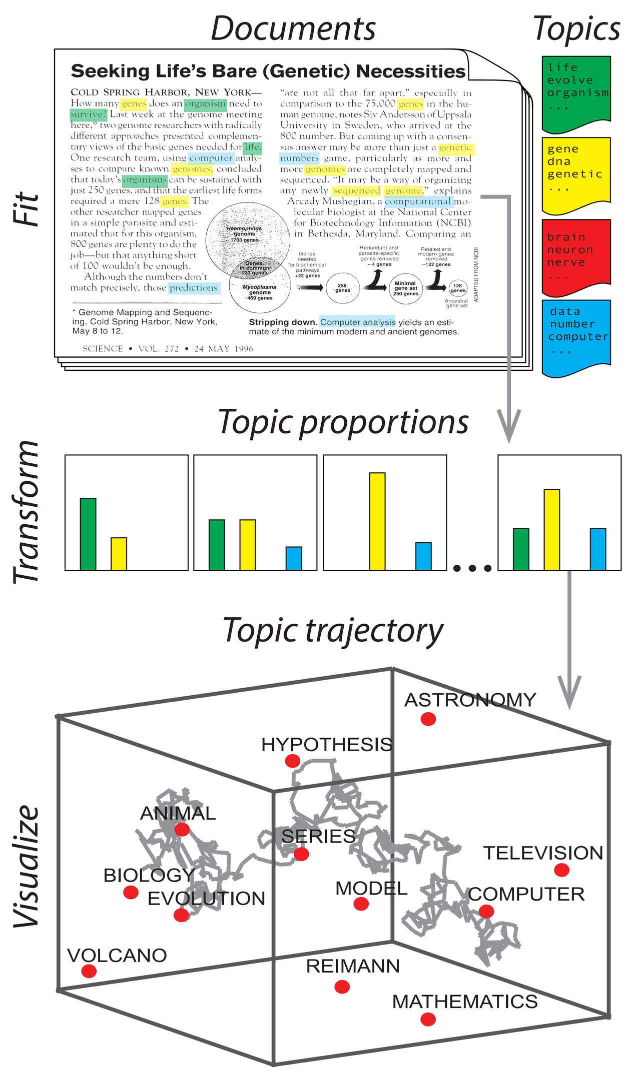
- While the broad **gist** of an experience is preserved well over time, **downsampling** its evolving content trajectory may underly the act of forgetting
- **Overlap** between the **semantic system** and other functional regions may affect our **inherent ability to remember** certain content of an experience.

A naturalistic memory experiment

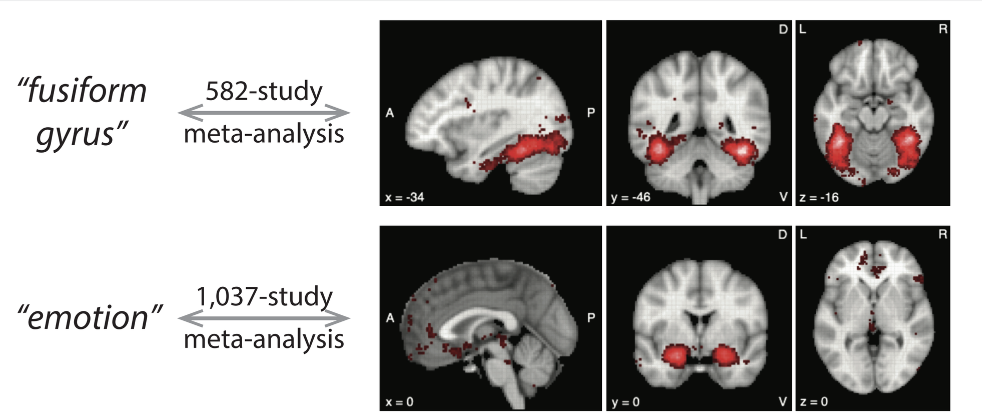


Participants (59) viewed the pilot episode of *Atlanta*, verbally recalled its content, and predicted what would happen in the next episode. A week later, they again recalled *Atlanta* episode 1 before viewing and recalling either *Atlanta* episode 2 or the pilot of *Arrested Development*.

Computational tools

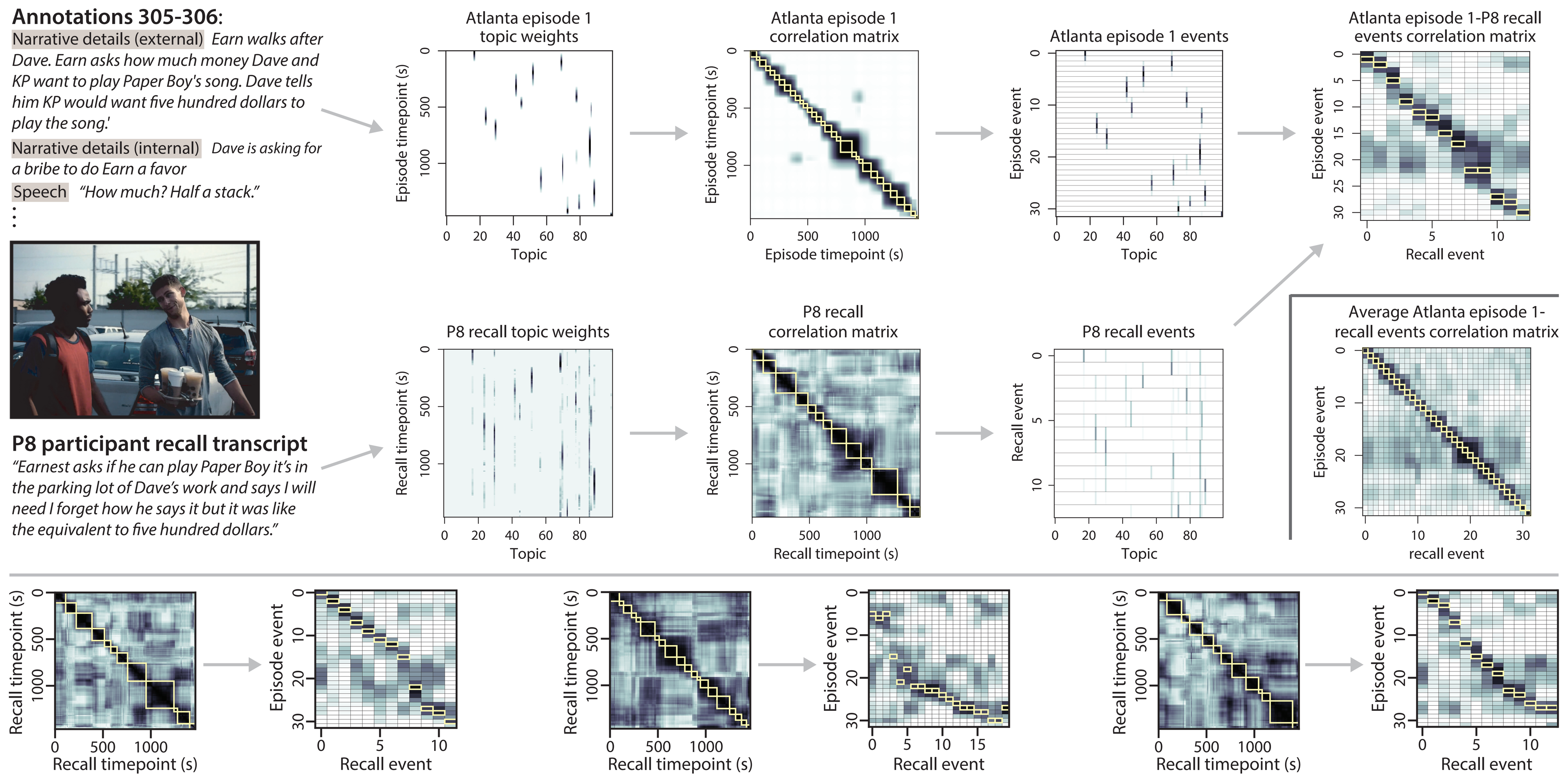


Topic models learn the latent themes in a set of documents and may then be used to transform text samples into a matrix describing these themes' presence in each sample. The evolving proportional weights of these themes (or "topics") over a length of text may be visualized as a trajectory through "topic space."



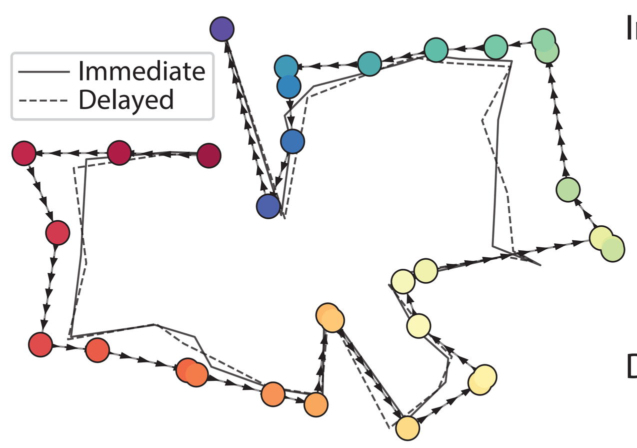
Neurosynth constructs meta-analysis images for key terms based on 500,000+ results reported in 14,000+ neuroimaging studies. It can then return the terms whose meta-analytic images are most similar to novel data.

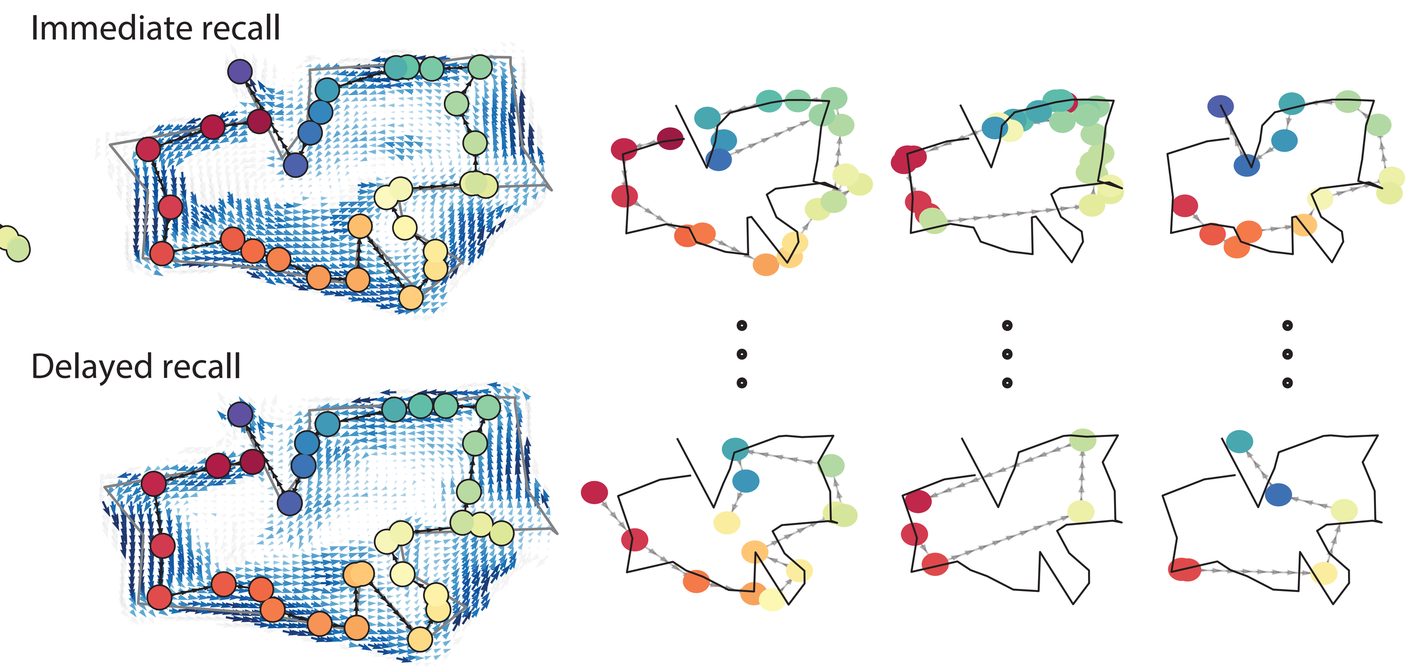
Modeling experience and memory

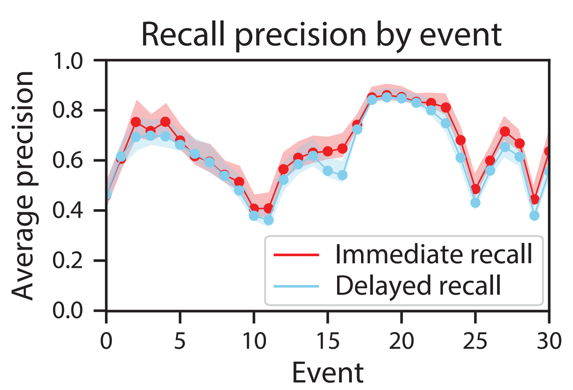


We fit a **topic model**[1] to hand-generated annotations of each episode and with it, transformed the episode and participants' recall transcripts into a common space. The **trajectories** formed in this space capture the moment-by-moment content of the experience and individuals' memories for it. We resampled the continuous trajectories to a standard length and segmented each into a set of discrete, maximally distinct events using **Hidden Markov Models**[2]. We then matched each recall event with the episode event to which it was most similar (i.e., whose average topic vector was most highly correlated). This also yielded the **precision** of each remembered event, given by the recovery of (correlation to) the corresponding episode event.

Changes in memories over time

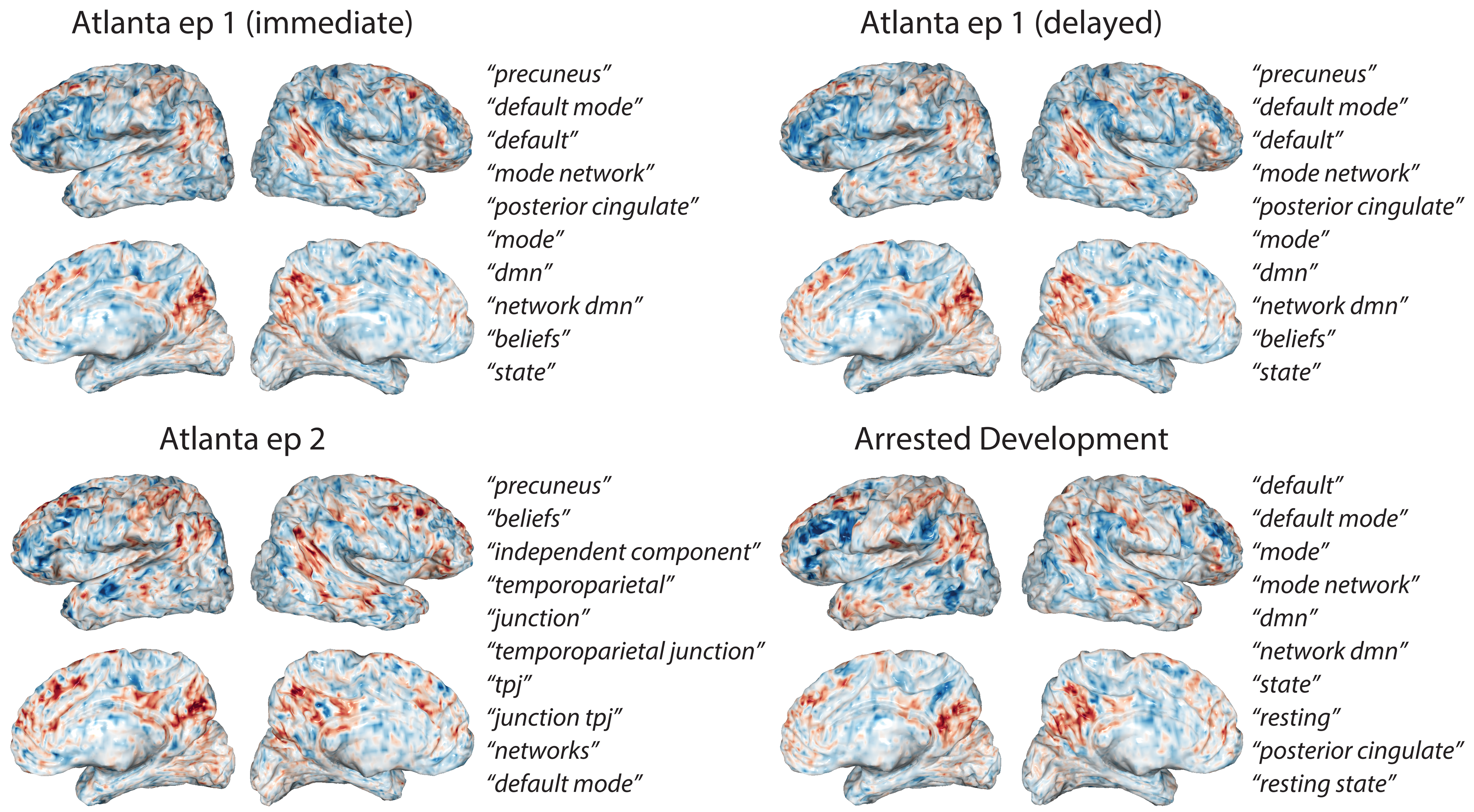




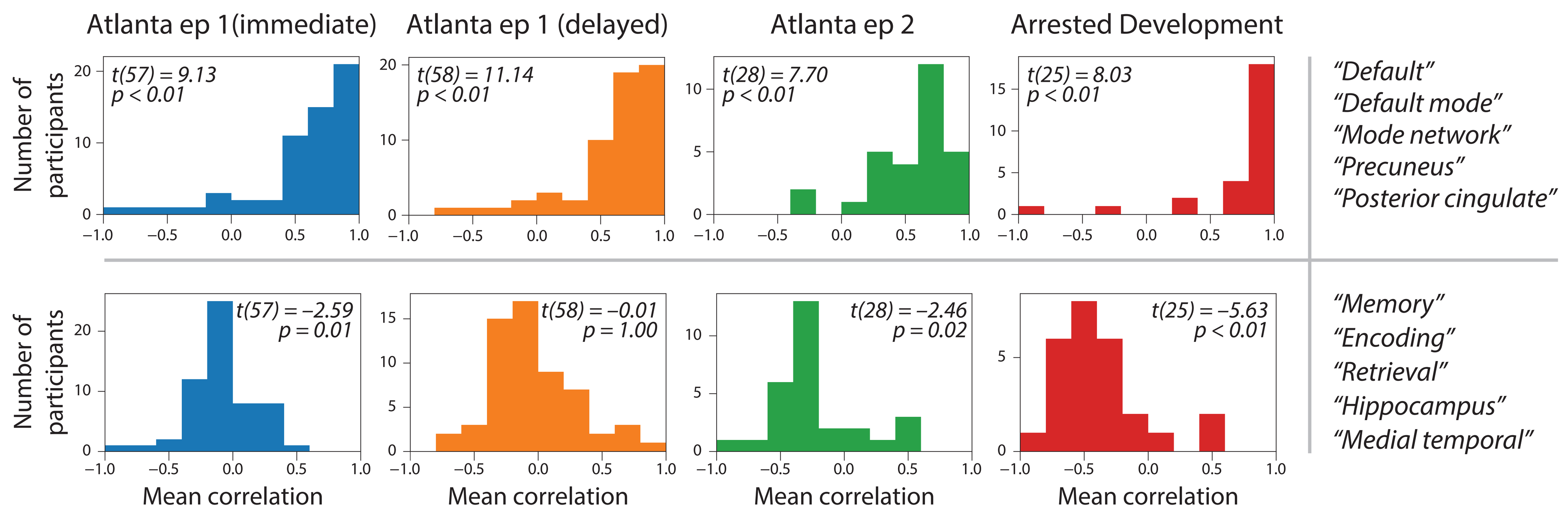


We projected the 100-dimensional episode and recall trajectories onto a 2D plane using UMAP[3,4]. While average precision and ability to recovery of the overall **shape** of the episode's unfolding remained extremely similar over the week-long delay, the **sampling rates** with which participants recovered the episode's dynamic content decreased sharply (Wilcoxon's T=543, p=0.02). This suggests that rather than discard episodic memories entirely, our memories **compress** representations of the external world over time, costing us the ability to distinguish between short-timescale segments of an experience.

A neural basis for content memorability



We projected each episode event's topic vector onto a mapping of the semantic system[5] and constructed recall precision-weighted averages of these activity maps across events. We then removed content-related activity and isolated signals related to memory quality. These brain maps show semantic system activation differences between events that participants recalled precisely and imprecisely. Decoding their activity with `Neurosynth`[6] revealed that these memorability differences are related to default mode network activity.



We projected each episode event onto the semanic system and decoded the degree of overlap between their neural representations and two functional networks: default mode and memory. Correlating each participant's event-wise recall precision with the resulting similarity scores revealed that content represented heavily in default mode regions tended to be remembered better, while content represented in memory areas was generally more difficult to recall.

CDL website: download our poster and preprint and learn more about our lab!

HyperTools: fit and apply text models, plot and manipulate high-dimensional data.

Quail: data analysis and plotting for list-learning and naturalistic memory experiments.

Referenes
1. DM Blei et al. (2003). Latent Dirichlet Allocation. JMLR, 3(1), 993-1022.
2. LR Rabiner (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2), 257-286.
3. L McInnes et al. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv*: 1802.03426.
4. AC Heusser et al. (2018). How is experience transformed into memory? *bioRxiv*: 409987.
5. AG Huth et al. (2016). Natural speech reveals the semantic maps that tile the human cerebral cortex. Nature, 532(7600), 453.
6. T Yarkoni et al. (2011). Large-scale automated synthesis of human functional neuroimaging data. Nature methods, 8(8), 665.

Supported in part by NSF EPSCoR Grant 1632738.