Neuropsychologia 189 (2023) 108664

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Studying waves of prediction in the brain using narratives

Christopher Baldassano

Department of Psychology, Columbia University, 1190 Amsterdam Ave., New York, NY, 10027, USA

ARTICLE INFO

Keywords:
Narrative
Prediction
Memory
Event segmentation

ABSTRACT

Narrative stimuli offer a unique opportunity for research in cognitive neuroscience because they evoke cognitive processes that are difficult or impossible to study with traditional paradigms. An especially compelling feature of narratives is their temporal structure, which allows for meaningful predictions about upcoming events. As we proceed through a narrative, we can maintain a complex set of short- and long-term guesses about the future and continually refine our predictions as the story unfolds. Experiments using narratives can allow researchers to probe the ways in which memory systems are flexibly used during perception, including the mechanisms by which continuous experiences are segmented into discrete events. Despite the challenges of using narratives and other naturalistic stimuli in experimental research, these approaches offer a new window into critical components of real-world cognition.

The standard experimental approach in cognitive neuroscience is to present participants with a sequence of discrete trials, which are intentionally randomized to ensure that there are no consistent relationships between neighboring items. Although this decorrelated temporal structure has a number of benefits, including greatly simplifying statistical analyses, it is wholly unlike the real-world environments in which our brains typically function. Our sensory inputs at any given moment are tightly connected to those that will occur over the next few seconds and minutes, and can be related in more complex ways to experiences we will have over the coming days and years. Understanding the events of our everyday lives and making effective choices requires us to draw on our memories for specific past experiences as well as our general knowledge about the world.

The field of cognitive neuroscience has increasingly turned to more naturalistic paradigms in order to study these critical elements of real-world cognition, either by incorporating more realistic stimuli within controlled designs or by studying fully unconstrained perception during hours-long narratives. These experiments have uncovered new principles for how cognitive systems interact over time and the neural systems that support dynamic perception and memory. Remarkably, stimuli with long-timescale structure (like coherent narratives or movies) have been shown to engage a set of regions previously thought of as a "task-negative network," which robustly *deactivates* in response to standard randomized stimuli (Yeshurun et al., 2021). Although researchers are still untangling the many perceptual processes carried out in these high-level brain regions, I propose that a key feature of narratives is that they allow for the ongoing generation and continuous updating of predictions at multiple timescales. In this Viewpoint I describe the waves of predictability that are evoked by narratives (and perhaps other naturalistic stimuli) and how narratives can provide a new window into the dynamics of the mind and brain during everyday perception.

1. Waves of predictability

> The question mark looks like a hook for a reason. Readers are driven by the *need to know.* Chuck Wendig

A basic property of events in a narrative is that they are almost always related to future events, creating a "network" of connections between parts of the story (Lee and Chen, 2022). Observing a narrative stimulus at time t (S_t) generally reduces our uncertainty about upcoming information at any future time d timepoints later (S_{t+d}). This idea can be captured using the concept of mutual information, written as $I(S_t\,; S_{t+d}) > 0$ for all timepoints t and positive delays d. On average, this mutual information will be largest for short delays d, but there can be moments at which events from the distant past (e.g. many episodes ago on a television show) suddenly become relevant again. We could imagine quantitatively computing the strength of these relationships between timepoints, either using some "objective" model of events in the world (which has become more plausible due to recent advances in large language models), or by measuring the predictions of actual participants.

I argue, however, that this basic kind of predictability is not the only structure in narratives, and is in fact not what makes narratives engaging and compelling. A list of the numbers "1, 2, 3, 4, …" would be extremely

E-mail address: c.baldassano@columbia.edu.

https://doi.org/10.1016/j.neuropsychologia.2023.108664
Received 31 May 2023; Received in revised form 4 August 2023; Accepted 18 August 2023
Available online 19 August 2023
0028-3932/

predictable (with maximal mutual information between items) but is unlikely to be a bestseller. Instead, narratives generally have the property that predictability *builds over time*. Given everything we've heard in the story so far, the current moment S_t should add *new* information about upcoming events. Each clue in a (good) murder mystery should cause us to re-evaluate our predictions about who the killer is and what actions the characters will take next. We can capture this property using *conditional* mutual information as $I(S_t\ ;\ S_{t+d}\ |S_{1..t-1}) > 0$: even after we've already observed the narrative so far ($S_{1..t-1}$), the new information S_t further reduces our uncertainty about future events S_{t+d}. The value of this conditional mutual information need not be uniformly high for all timepoints t and positive delays d (and in fact fluctuations in predictability can provide important temporal structure to narratives, as discussed below), but must be consistently large enough to avoid a narrative becoming either chaotic (if prediction is never possible) or formulaic (if early events remove all uncertainty about later events).

In order for a stimulus to have this kind of predictability, predictions must always be both possible and imperfect. We should be able to make meaningful guesses about upcoming events, while leaving room for future information to refine or revise these guesses. This idea is consistent with previous proposals that engaging experiences should have an intermediate amount of predictability (Gold et al., 2019), but further hypothesizes that moderate predictability is not by itself sufficient for engagement. Instead, we need a "wave" of predictability, in which new information allows us to constantly update our expectations about the future (Fig. 1).

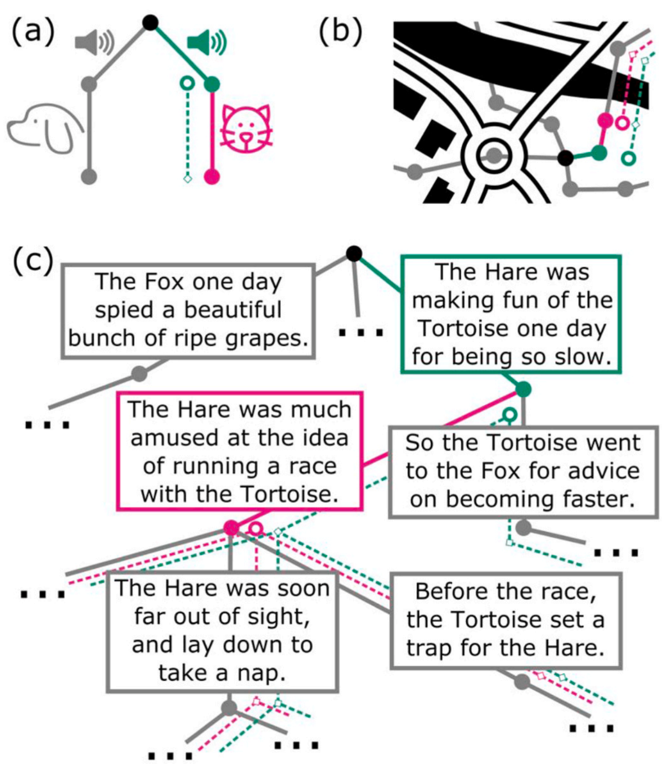

Fig. 1. Narratives allow for continual refining of multi-step predictions. In experiments with predictable structure, participants can make predications when observing an initial stimulus (green). In traditional associative learning (a), predictions (dotted lines) can be made only for the immediately-upcoming stimulus. Studies in which participants learn multi-step sequences (b) allow for anticipation father into the future, but these predictions are fully determined after seeing the first stimulus. Narratives (c) provide an open-ended but structured space in which participants can maintain probabilistic distributions over possible future events, and meaningfully update these predictions when observing a subsequent stimulus (purple).

2. Cognitively connecting events across time

Any literary narrative of depth asks your brain to pull threads across the whole, so, in a sense, all complex narratives are networks; your experience moving through them is never purely linear, but volumetric or spatial as your thoughts bounce across passages.

Jane Alison, *Meander, Spiral, Explode*

What mental processes are evoked by a stimulus with this kind of predictable structure? Experimental participants could anticipate future events by drawing on one or more specific episodic memories; my lab has found that repeated exposure to the same narrative movie clip shifts brain responses earlier in time, with smaller shifts in visual regions and larger shifts in prefrontal regions (Lee et al., 2021). Predictions could alternatively be generated based on general schematic knowledge about event sequences in the world, especially as people build expertise in a domain (Huang et al., 2023). In addition to forward prediction, we can also connect events backward in time to their prior causes, reactivating the memory of a previous moment that is critically related to the current stimulus.

There are a number of intriguing questions about how these processes take place. The neural systems that drive episode-based prediction, schematic prediction, and reinstatement of related events are likely all distinct, and may rely differentially on the hippocampus versus the neocortex. Measuring anticipatory signals in a region does not necessarily establish that this region is itself generating predictions, since these representations could be driven by top-down input; studies of patient populations or experiments using brain stimulation may be necessary for understanding where predictions originate. There are also cognitive questions about the features of narratives that support anticipation: for example, narratives for children often provide extensive structuring to facilitate prediction, but this may make stories *too* predictable to be enjoyable for adults.

3. Boundaries between the past and the future

It was as if a blade had fallen onto time itself, cleaving it into halves, that which came before and that which came after.

Justin Cronin, *The Passage*

The densely interconnected nature of narratives presents a major challenge for our memory systems. Because working memory does not have the capacity to maintain all the details encountered since the beginning of the story, we must strategically decide which information is currently the most relevant for understanding the story and should therefore be kept readily-available. The best moment to clear out our working memory, from a predictive perspective, is at a timepoint b when $I(S_{b-d}\ ;\ S_{b+d}\ |S_b)$ is very small for short delays d; that is, when the future and the past are largely independent given the current stimulus. At these moments, we have little reason to continue using valuable working memory resources to maintain past information S_{b-d}.

I propose that this fundamental working memory constraint is what gives rise to *event boundaries*, key moments in narratives that mark the start of new mental event models (Radvansky and Zacks, 2017) and evoke rapid changes in neocortical neural representations (Baldassano et al., 2017). When a new scene of a movie begins, details of the previous scene (such as the specific words used in the dialog) suddenly become much less likely to be important, and should cue us to release these outdated details from our memory buffer. Before performing this flushing, we must also ensure that critical information about the past is encoded into long-term memory; this encoding operation likely corresponds to the hippocampal activation observed at event offsets, which is related to subsequent memory (Ben-Yakov et al., 2014).

This hypothesis about event boundaries leaves open the question of exactly which cues in the narrative should trigger event boundaries, since we cannot know for sure whether recent details will be relevant again in the near future. A spike in prediction error could be one indication that a new event has begun (Radvansky and Zacks, 2017), but even an entirely expected moment (e.g. the countdown for a rocket launch reaching zero) could prompt us to begin a fresh event model focused on newly incoming information. It is also unclear whether working memory is *entirely* flushed at event boundaries; it is possible that only the least relevant information is removed while some portion is retained, or that only some of the brain regions along the cortical hierarchy are flushed at more minor event boundaries (Baldassano et al., 2017).

4. Beyond narratives

> There was the life you lived, which consisted of the choices you made. And then, there was the other life, the one that was the things you hadn't chosen.
>
> Gabrielle Zevin, *Tomorrow, and Tomorrow, and Tomorrow*

Narratives allow us to study the process by which people make and then continually update predictions, leveraging the complex real-world knowledge that we have developed over decades to enable sophisticated inferences about possible outcomes. Beyond narratives, there are other kinds of naturalistic paradigms that could allow for the same kind of iterative building of predictions. Perhaps the purest form occurs in music, which can engage our attention purely through auditory dynamics without inherent semantic content. David Huron's book *Sweet Anticipation: Music and the Psychology of Expectation* (2006) argues that generating predictions is a key psychological aspect of music, driven by mechanisms including "veridical familiarity" (the exact repetition of a familiar passage) and "schematic predictability" (expectations from learned musical conventions). Many interactive games also have complex temporal trajectories, allowing a skilled player to anticipate potential upcoming events (Huang et al., 2023). Prediction may be especially critical in this case since players must make active choices, testing their ability to maintain and act on relevant information about the state of the game. Whether all of these paradigms engage the same neural mechanisms for prediction is still an open question; in a recent study we found that music evoked structured neural dynamics in regions similar to those implicated in narrative perception (Williams et al., 2022), suggesting that there may be common mechanisms for generating predictions across domains.

Narratives and other naturalistic stimuli can provide a rare opportunity to study a key process in cognition: developing nuanced predictions that are repeatedly tested and updated. Perhaps we are innately drawn to these kinds of stimuli (and choose to spend our free time pursuing them) precisely *because* they allow us to exercise our cognitive abilities to imagine potential futures and think backward to connect related events. Our daily lives take place largely in familiar places, with familiar people, in familiar sequences of events, and we continually use this knowledge to create, refine, and reconstruct our experiences; a full scientific understanding of cognition therefore requires us to use experimental tools that can evoke these same processes.

CRediT authorship contribution statement

Christopher Baldassano: Conceptualization, Writing – original draft, Writing – review & editing, Visualization.

Data availability

No data was used for the research described in the article.

Acknowledgements

Thank you to Janice Chen for invaluable feedback on a draft of this manuscript, and to the Dynamic Perception and Memory Lab for many stimulating conversations about these ideas.

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. In: Neuron, vol. 95. Elsevier BV, pp. 709–721. https://doi.org/10.1016/j.neuron.2017.06.041. Issue 3 e5.

Ben-Yakov, A., Rubinson, M., Dudai, Y., 2014. Shifting gears in Hippocampus: temporal dissociation between familiarity and novelty signatures in a single event. In: The Journal of Neuroscience, vol. 34. Society for Neuroscience, pp. 12973–12981. https://doi.org/10.1523/jneurosci.1892-14.2014. Issue 39.

Gold, B.P., Pearce, M.T., Mas-Herrero, E., Dagher, A., Zatorre, R.J., 2019. Predictability and uncertainty in the pleasure of music: a reward for learning?. In: The Journal of Neuroscience, vol. 39. Society for Neuroscience, pp. 9397–9409. https://doi.org/10.1523/jneurosci.0428-19.2019. Issue 47.

Huang, J., Velarde, I., Ma, W.J., Baldassano, C., 2023. Schema-based predictive eye movements support sequential memory encoding. In: eLife, vol. 12. eLife Sciences Publications, Ltd. https://doi.org/10.7554/elife.82599.

Huron, D., 2006. Sweet Anticipation. The MIT Press. https://doi.org/10.7551/mitpress/6575.001.0001.

Lee, C.S., Aly, M., Baldassano, C., 2021. Anticipation of temporally structured events in the brain. In: eLife, vol. 10. eLife Sciences Publications, Ltd. https://doi.org/10.7554/elife.64972.

Lee, H., Chen, J., 2022. Predicting memory from the network structure of naturalistic events. In: Nature Communications, vol. 13. Springer Science and Business Media LLC. https://doi.org/10.1038/s41467-022-31965-2. Issue 1.

Radvansky, G.A., Zacks, J.M., 2017. Event boundaries in memory and cognition. In: Current Opinion in Behavioral Sciences, vol. 17. Elsevier BV, pp. 133–140. https://doi.org/10.1016/j.cobeha.2017.08.006.

Williams, J.A., Margulis, E.H., Nastase, S.A., Chen, J., Hasson, U., Norman, K.A., Baldassano, C., 2022. High-order areas and auditory cortex both represent the high-level event structure of music. J. Cognit. Neurosci. 34 (4), 699–714. https://doi.org/10.1162/jocn_a_01815. MIT Press - Journals.

Yeshurun, Y., Nguyen, M., Hasson, U., 2021. The default mode network: where the idiosyncratic self meets the shared social world. In: Nature Reviews Neuroscience, vol. 22. Springer Science and Business Media LLC, pp. 181–192. https://doi.org/10.1038/s41583-020-00420-w. Issue 3.