Mnemonic content and hippocampal patterns shape judgments of time

Brynn E. Sherman[1]*, Sarah DuBrow[2]*, Jonathan Winawer[3], Lila Davachi[4,5]

* denotes equal contributions

[1]Department of Psychology, Yale University
[2]Department of Psychology, University of Oregon
[3]Department of Psychology and Center for Neural Science, New York University
[4]Department of Psychology, Columbia University
[5]Nathan Kline Institute for Psychiatric Research

Abstract

Our experience of time can feel dilated or compressed, rather than reflecting true "clock time." Although many contextual factors influence the subjective perception of time, it is unclear how memory accessibility plays a role in constructing our experience of and memory for time. Here, we used a combination of behavioral and fMRI measures to ask the question of how memory is incorporated into temporal duration judgments. Behaviorally, we found that event boundaries, which have been shown to disrupt ongoing memory integration processes, result in the temporal compression of duration judgments. Additionally, using a multivoxel pattern similarity analysis of fMRI data, we found that greater temporal pattern change in the left hippocampus within individual trials was associated with longer duration judgments. Together, these data suggest that mnemonic processes play a role in constructing representations of time.

Statement of Relevance

Our everyday experiences convey a powerful truth: That our *perception* of time often diverges from the *reality* of time. When enjoying an active vacation with family, time moves quickly: hours go by in minutes. When sitting through an unnecessary meeting, time moves slowly: minutes go by in hours. What is the origin of these phenomenologically compelling illusions of time perception? Past research has examined how a range of specific factors, from emotions to blinking, contribute to the distortion of time. Here, in contrast, we evaluate how the content and accessibility of our memories shapes time perception. We show that context shifts, known to disrupt memory processing, also lead to robust contractions of perceived time. We discuss how both effects — memory disruptions and time distortions — may be linked via the hippocampus.

“An hour, once it lodges in the queer element of the human spirit, may be stretched to fifty or a hundred times its clock length; on the other hand, an hour may be accurately represented on the timepiece of the mind by one second.” — Virginia Woolf

Introduction:

On a busy vacation, time may escape you — by the time you go to the museum, grab lunch in the park, souvenir shop, and visit a historical site, the day may seem to have flown by. Yet, when recalling the trip to a friend, that same day may feel like a week; all of those events could not have possibly occurred within the same few hours. This puzzle raises a fundamental question of how the structure of experience can paradoxically influence subjective impressions of time in experience and in reflection.

A great deal of work has focused on the latter: how the structure of experience influences how we *remember* elapsed time. In particular, abrupt shifts in context, or event boundaries, influence memory for time. Memory for the temporal order of events is disrupted across event boundaries (DuBrow & Davachi, 2013; Horner et al., 2016; Heusser et al., 2018). Further, intervals which contain an event boundary are remembered as *longer* than equivalently-timed intervals without a boundary, (Ezzyat & Davachi, 2014; Clewett et al., 2020) and mnemonic duration judgments scale with the number of events (Faber & Gennari, 2015; Losistsky et al., 2016; Faber & Gennari, 2017). Such findings converge with the intuition that busy days feel long in memory: events may serve to dilate time in memory. How, though, do events also result in time feeling subjectively shorter *in the moment*?

Event boundaries also influence memory on the more immediate time-scale of working memory by reducing access to information prior to the boundary (Morrow et al., 1989; Zwaan, 1996; Rinck & Bower, 2000; Zacks & Tversky, 2001; Speer & Zacks, 2005; Ezzyat & Davachi, 2011; Radvanksy, 2012). Thus, perhaps this decreased accessibility to mnemonic information translates into a *contraction* of time for intervals containing event boundaries. Models of time perception posit an “integrator” or “accumulator” which sums across experience to determine how much time has passed (e.g., Droit-Volet & Meck, 2007; Wittmann, 2013). By reducing the contents of working memory, event boundaries may reduce the information integrated into the duration judgment, leading to an underestimation of time. Consistently, there is some evidence that in-the-moment time judgments are compressed after one or more discrete boundaries (Liverence & Scholl, 2012; Bangert et al., 2019; Yousif & Scholl, 2019). However, whether this effect of event boundaries on more immediate time judgments is indeed due to a decreased accessibility to the prior event remains an open question.

A separate body of literature has examined the interplay of time and memory by asking whether the hippocampus — a structure critical for episodic memory (Scoville & Milner, 1957) — plays a role in tracking time. In one seminal study addressing the question of how memory is incorporated into duration judgments, Meck and colleagues found that although disrupting hippocampal function did not result in impairments in perceiving duration, it critically impaired temporal working memory and led to underestimation of the “reference memories” (Meck et al., 1984). This study suggests that the hippocampus may be involved in the process of making

duration judgments, but leaves open the question of how, precisely, mnemonic information is incorporated into representations of time.

A plethora of more recent work has established that the medial temporal lobe — in particular the hippocampus — indeed is sensitive to temporal duration information (Eichenbaum, 2013; Davachi & DuBrow, 2015; Lee et al., 2020). Independent of memory demands per se, populations of hippocampal neurons reflect the passage of time (Manns et al., 2007; Mankin et al., 2012; Rubin et al., 2017). Individual hippocampal neurons are sensitive to temporal information, firing during delays (MacDonald et al., 2011; 2013; Naya & Suzuki, 2011; Sakon et al., 2014; Umbach et al., 2020; Reddy et al., 2021) or at specific temporal moments (Terada et al., 2017; Sun et al., 2020). Furthermore, memories acquired close in time are coded by overlapping populations of neurons (Cai et al., 2016). Such overlapping representations may have consequences for the subjective representation of time in *long-term memory:* events remembered as further apart in time are associated with greater fMRI pattern change in the medial temporal lobe (Ezzyat & Davachi, 2014; Nielson et al., 2015; Deuker et al., 2016; Losistsky et al., 2016). However, whether and how such hippocampal representations influence more immediate judgments of time remains unclear.

In the present study we examine the role of memory accessibility and hippocampal representations on subjective judgments of time. Critically, we induce context shifts by inserting event boundaries, allowing us to assess the role of memory representations while holding true duration constant. First, in three behavioral experiments, we extend prior work demonstrating that event boundaries reduce estimates of duration, and provide evidence that these reductions are specifically due to decreased mnemonic accessibility. In a final fMRI experiment, we use pattern similarity to track changes in representations within a single temporal interval, and find that pattern change in the left hippocampus supports duration judgments, suggesting that the hippocampus may carry behaviorally-relevant temporal information on the order of seconds.

Methods:

Experiment 1

Participants

Twenty-one individuals (13 female; age range 18-33; mean = 21.5) were recruited from New York University and the larger community and participated for either course credit or payment ($10/hr). Informed consent was obtained in a manner approved by the University Committee on Activities Involving Human Subjects at New York University. One participant was excluded for reporting that they explicitly counted the intervals during the task, leaving twenty usable participants.

Stimuli

Stimuli were colored blue (R = 0, G = 0, B = 255), green (R = 0, G = 255, B = 0), and yellow (R = 260, G = 200, B = 80) squares presented centrally on a mid-gray background. Each square was presented on-screen for a 0.5 – 5 seconds interval (sampled equally in increments of 0.5

seconds). For half of the trials, the color remained the same for the entire duration (continuous condition), and for the other half of the trials the color switched half-way through the total duration (boundary condition).

Each participant viewed each color/color-pair once for each time point. For each participant, color pairings were randomly assigned such that the pairing of the three colors in the boundary condition was fixed (for example, the boundary trials for one participant would always be blue-green, green-yellow, yellow-blue, and this pairing would be randomly assigned for each participant). The order of presentation was pseudo-randomized such that condition-duration combinations were not repeated back-to-back and no condition or duration appeared more than 4 times consecutively.

In total, the task consisted of 3 trials per condition and time point (60 trials total).

Duration Judgment Task
On each trial, participants were instructed to attend to the square and keep track of the time. Importantly, participants were explicitly verbally instructed not to count while the square was on the screen (Rattat & Droit-Volet, 2012), and debriefing questionnaires suggested that compliance with this instruction was high.

Participants subsequently were presented with a continuous timeline with the prompt “How long was the square on the screen?” This timeline was bounded by 0.5 sec and 5 sec. Participants were instructed to disregard any color changes and estimate the total duration of the square presentation, regardless of color. Participants responded using a computer mouse. Responses were self-paced.

Experiment 2

Participants
Twenty-nine individuals (18 female; age range 18-24; mean = 19.4) were recruited from New York University for course credit. Informed consent was obtained in a manner approved by the University Committee on Activities Involving Human Subjects at New York University. The sample size for this experiment was chosen to be larger than that of Experiment 1 to increase sensitivity, given that more conditions were present in this experiment.

Stimuli
Stimuli were colored blue (R = 0, G = 0, B = 200), green (R = 75, G = 205, B = 75), and yellow (R = 260, G = 200, B = 80) squares presented centrally on a mid-gray background. Each square was presented on-screen for a 0.5 – 5 seconds interval (sampled equally in increments of 0.5 seconds). For half of the trials, the color remained the same for the entire duration (continuous condition). The other half of the trials were one of three boundary conditions: for one third of the boundary trials (one-sixth of the total experiment), the color switched one-quarter of the way through the total duration (such that the second event was 75% of the total duration); for one third, the color switched one-half of the way through (second event was 50% of total duration);

for one third, the color switched three-quarters of the way through the total duration (second event was 25% of total duration).

Each participant viewed each color/color-pair an equal number of times for each time point (here, twice over the entire experiment). Color pairings and pseudorandomization of trial sequences was conducted using the same procedure as in experiment 1.

In total, the task consisted of 3 trials per condition and time point (120 trials total).

Duration Judgment Task
The duration judgment task was identical to that described in Experiment 1.

Experiment 3

Participants
Eighty individuals (62 female; age range 18-27; mean = 19.9) were recruited from New York University and the larger community and participated for either course credit or payment ($10/hr). Twenty individuals each were assigned to each of the four experimental groups (described below). Informed consent was obtained in a manner approved by the University Committee on Activities Involving Human Subjects at New York University.

Stimuli
Stimuli were colored squares presented centrally on a mid-gray background. Each square was presented on-screen for a 0.5 – 5 seconds interval (sampled equally in increments of 0.5 seconds).

For all participants, for half of the trials, the color remained the same for the entire duration (continuous condition). The other half of the trials were "boundary condition," which differed by experimental group. In the abrupt change groups, there were two color switches, which occurred one-third and two-thirds of the way through the total duration. To test whether the boundary duration compression effect was driven by the number of changes, or change of the final color specifically, we had an "ABC" group, in which the color switched to a new color at each switch, and an "ABA" group, in which the final segment of the trial returned to the color of the first segment. In the gradual change groups, the color slowly morphed from the initial color to the final color without an abrupt boundary. Because both rate of color change and ending color (if rate of change was held constant) could be cues to duration, we included two groups: one which had a constant rate of change and different end colors for different durations (rate-constant), and one which had constant end colors and different rates of change for different durations (rate-changing). See **Supplemental Figure 2** for a depiction of the four conditions.

For the abrupt change groups, square stimuli were blue (R = 0, G = 0, B = 255), green (R = 0, G = 255, B = 0), and yellow (R = 260, G = 200, B = 80). For each participant, color triplets were randomly assigned such that the order of the three colors in the boundary condition was fixed. Each participant viewed each color/color-triplet twice for each time point.

For the gradual change groups, square stimuli morphed between red (R = 250, G = 0, B = 0) and blue; blue (R = 0, G = 0, B = 250) and red; and blue and green. In the rate changing group, the step size between colors per screen refresh was calculated separately for each timepoint such that the end color was always the same, e.g., starting at [250 0 0] and ending at [0 0 250]. In the rate constant group, the step size between colors per screen refresh was fixed, such that 1 RGB value was subtracted from the start color and one added to the end color for each refresh.

Pseudorandomization of trial sequences was conducted using the same procedure as in experiment 1. In total, the task consisted of 6 trials per condition and time point (120 trials total), broken up into two runs.

Duration Judgment Task
The duration judgment task was identical to that described in Experiment 1.

Experiment 4

Participants
Eighteen right-handed native English speakers (7 female; age range 22-34; mean = 27.1) were recruited from New York University and the larger community and participated for payment ($25/hr). We planned to recruit twenty individuals, as in Experiment 1, but due to attrition only had eighteen usable participants. However, this experiment has more trials than the previous, and the neural analyses use a within-participant approach to brain-behavior correlations. Informed consent was obtained in a manner approved by the University Committee on Activities Involving Human Subjects at New York University. Due to equipment failure, the first eight out of fifteen runs were lost for one participant; thus, only the usable runs of this data set were included in analyses. Due to time constraints, the verbal temporal perception task and localizer task were not collected for two participants.

Stimuli
Stimuli were colored blue (R = 0, G = 0, B = 160), red (R = 160, G = 0, B = 0), and yellow (R = 195, G = 175, B = 35) squares presented centrally on a mid-gray background. Each square was presented on-screen for a 2, 4, 6, or 8 second interval. For half of the trials, the color remained the same for the entire duration (continuous condition), and for the other half of the trials the color switched half-way through the total duration (boundary condition).

Each participant viewed each color or color-pair an equal number of times for each time point (here, ten times over the entire experiment). Color pairings and pseudorandomization of trial sequences was conducted using the same procedure as in experiment 1.

During each delay period (see below), a colored noise mask was presented over the entire display. A single mask image was created by randomly sampling R, G, and B values independently for each pixel on a 13” MacBook Pro.

In total, the task consisted of 30 trials per condition and time point (240 trials total). These were broken into 15 fMRI runs of 16 trials each, such that each condition/duration combination was presented twice per run.

Duration Judgment Task

The duration judgment task was identical to that of Experiment 1, with the following changes. The timeline was bounded by 0 sec and 10 sec. Participants responded using an MRI-compatible trackball. Additionally, there was a variable delay interval (2, 4, or 6 seconds) between square presentation and time judgment, which was included to orthogonalize the fMRI regressors associated with the presentation and judgment periods. During this delay, a colored noise mask was presented over the entire screen to prevent any after-images that may be caused by a lingering percept of the square.

There was a variable inter-trial interval (4, 6, or 8 seconds) during which participants performed an odd/even task after the time judgment period. Each number was presented for a maximum of 1.9 seconds with an inter-stimulus interval of 0.1 seconds. Participants responded using the left and right buttons on the sides of trackball. The inter-trial periods were jittered in order to orthogonalize the fMRI regressors associated with the presentation and judgment periods, and this interval served as baseline in fMRI analyses.

Verbal duration judgment task

To address concerns of systematic biases in number line usage between participants, a behavioral session was run following completion of the fMRI study, in which participants – instead of indicating their responses on the number line – verbally reported how long they thought the square was on the screen for, in seconds (**Supplemental Figure 3C,D**). Participants were given no numerical bounds or constraints on how precise their responses had to be. No functional data were collected during this session.

Localizer

At the end of the fMRI session, participants engaged in a blocked color localizer in which each of the three colors was flashed repeatedly (0.9 seconds on, 0.1 seconds blank inter-stimulus interval) for 32 seconds. Each color was presented twice per run, and there were two runs of this task. The data from this run are not reported in this paper.

fMRI Parameters

Functional images were acquired on a Siemens Allegra head-only 3T scanner. Data were collected using an EPI pulse sequence (34 contiguous slices oriented parallel to the AC-PC axis; TR = 2000 ms; TE = 15 ms; flip angle = 82°; voxel size 3x3x3 mm). Additionally, a high-resolution T1-weighted anatomical scan (magnetization-prepared-rapid-acquisition gradient echo sequence, 1x1x1 mm) was obtained for each participant. During functional scans, stimuli were viewed through a mirror attached to the head coil.

Preprocessing of fMRI Data

fMRI data processing was carried out using FEAT (fMRI Expert Analysis Tool) Version 6.00, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). BOLD images were first skull-stripped using the Brain Extraction Tool (Smith, 2002). The first four volumes of each run were discarded to allow for T1 stabilization. Additionally, images underwent motion correction using MCFLIRT (Jenkinson et al., 2002), grand-mean intensity normalization by a single multiplicative factor, and high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 50.0s).

Univariate Whole-Brain ROI Definition

For univariate analyses to define ROIs, the functional data were spatially smoothed with a 8.0mm FWHM Gaussian kernel. Preprocessed data were registered to each participant's anatomical image, which was registered to an MNI 2mm standard space image using FLIRT (Jenkinson & Smith, 2001, Jenkinson et al., 2002); this registration was used to transform group-level contrasts into each participant's native space for creating the visual cortical ROIs.

Multivariate Pattern Analyses

For the multivariate (pattern similarity) analyses, no spatial smoothing was applied to the data. Each run was realigned to the final fMRI run of the session. Motion outliers were detected using FSL Motion Outliers; TRs containing motion outliers were excluded from pattern similarity analyses (see "Pattern Similarity Analysis" section below).

ROI Definition

Anatomical hippocampal ROIs (left and right) were defined for each participant via FSL's FIRST automated segmentation tool (Patenaude et al., 2011). Given prior findings that patterns specifically in the left hippocampus scale with subjective estimates of time (Ezzyat & Davachi, 2014; Nielson et al., 2015), we hypothesized that our brain-behavior relationships would be present in the left hippocampus. Thus, we report results from only the left hippocampus in the main text (but see **Supplemental Figure 4** for exploratory right hippocampal and bilateral results).

In defining visual cortex, we aimed to isolate voxels that were maximally sensitive to visual properties of the main task. We ran a series of univariate contrasts intended to identify voxels that were more responsive to the colored squares than to the visual mask. At the run level, we ran a GLM with 7 regressors of interest, for the 6 and 8 second trials only: onsets of continuous and boundary trials (modeled as a stick function), separately; midpoints of continuous and boundary trials (modeled as a stick function), separately; onsets of the colored noise mask (modeled as a stick function); duration judgment response window (boxcar with the length of the duration judgment period, which varied trial-by-trial); and odd/even ITI task (boxcar with the length of the ITI). Each regressor was convolved with a gamma hemodynamic response function.

First, to identify voxels that were sensitive to the onset of a color stimulus, we contrasted activation corresponding to the onsets of trials, as well as the boundary (onset of the second color), to the midpoints of a continuous trial. Next, to identify voxels that were most peripherally

responsive, we contrasted activation relating to the onset of the mask to the onset of the square stimulus (of both conditions). These contrasts were first defined at the run level and then modeled across runs as a fixed effect within participant. Group-level analyses were modeled using mixed-effects (FLAME 1), and were thresholded using clusters determined by $Z>2.3$ and a corrected cluster significance threshold of $p = 0.05$ (Worsley, 2001).

The resulting "peripherally-responsive" mask was then subtracted from the resulting "onset-sensitive" mask to isolate voxels that were maximally sensitive to colored square stimuli. Given prior work demonstrating patterns of activity in V1 and V4 carry information about color (Brouwer & Heeger, 2009), this region was masked to only include voxels within probabilistic V1 or hV4 (Wang et al., 2015), thresholded to 75%, and was then transformed into each participant's native space.

Anatomical caudate and putamen ROIs were defined for each participant via FSL's FIRST automated segmentation tool (Patenaude et al., 2011).

Prefrontal cortical regions were defined using the Harvard-Oxford Cortical Structural Atlas. Specifically, ROIs were created in standard space for Middle Frontal Gyrus; Inferior Frontal Gyrus, pars triangularis; and Inferior Frontal Gyrus, pars opercularis. Each of these three ROIs were then transformed into participants' functional spaces and thresholded at 75%. The Inferior Frontal Gyrus, pars triangularis and Inferior Frontal Gyrus, pars opercularis were then concatenated together to create one IFG ROI.

Entorhinal cortex was manually segmented on every participant's T1-weighted anatomical scans, using published anatomical landmarks (Insausti et al., 1998; Pruessner et al., 2002).

Pattern Similarity Analysis

Pattern similarity analyses were performed on 8 second trials only to maximize separation between the hemodynamic responses to the beginning and end of each trial. The data were modeled via an approach adapted from Turner et al (2012). This approach is well suited for within-trial analyses, as it uses a multi-parameter single-trial GLM, which allows for the un-mixing of temporally adjacent BOLD responses. Because our stimuli were temporally extended, we extracted the activation pattern across voxels for TRs 2-7 (4-14 seconds) after the true stimulus onset. A design matrix was constructed in which each of these "trial-related" TRs was a separate regressor of interest; in addition, separate nuisance regressors of no-interest included separate TRs for every other trial collapsed across each condition/color combination, as well as TRs from the ITI and delay periods. This design matrix was regressed against the activation of each voxel in each ROI to compute a separate beta estimate for each TR of the trial. To avoid making strong assumptions about where the true onset, midpoint, and offset of each trial fall, we estimated the beginning and end of the stimulus period by averaging the first two and last two betas, respectively. Specifically, from trial onset, the second and third TRs (corresponding to approximately 0-2 seconds after stimulus onset) were averaged as an estimate of the "beginning" and the fifth and sixth TRs (corresponding to approximately 6-8 seconds after stimulus onset) were averaged as an estimate of the "end" (see Fig 2).

TRs containing a motion outlier were excluded from subsequent analyses; furthermore, if there was more than one motion outlier within the "trial-related" TRs, the entire trial was excluded from subsequent analyses.

Event Similarity
Event similarity was computed by correlating the patterns corresponding to the "beginning" and "end" of every trial separately. Pearson correlation values were Fisher-transformed and averaged for the two conditions to produce a "similarity" score.

Event Similarity Predicting Behavior
To assess whether within-trial event similarity in visual cortex and hippocampus influence duration judgments, we performed a mixed-effects analysis implemented via the lme4 package in R (http://cran.r-project.org/web/packages/lme4). To account for individual differences in number line usage, we z-scored participants' duration judgments across all trials. Mixed-effects linear models were run, and significance was assessed using iterative model comparisons, which resulted in chi-squared values and corresponding *p*-values. Normalized duration judgment was the dependent measure; visual cortical similarity, hippocampal similarity, and condition, as well as interactions between them, were included as fixed-effect predictors. Our general approach consisted of running the simplest form of the model, and then only including significant predictors from these simple models in subsequent models. Model comparisons were also used to determine which variables were included as by-participant random factors; as a result of these comparisons, the condition factor was included as a by-participant random effect, such that a slope and intercept were calculated for each participant.

Color Similarity
To isolate responses to color, irrespective of temporal position information, we assessed similarity of the "end" of a boundary trial to the "beginning" of all other trials, conditionalized on whether the color at the end of the trial of interest was the same or different than the color at the beginning of the other trials. This across-trial color similarity metric was computed as the Fisher transformation of the Pearson correlation between the "end" of the trial of interest and the "beginning" of every other trial separately. Trials from the same fMRI run were excluded due to temporal autocorrelation. For each trial, the mean of the similarity values to all other trials was computed. The mean color similarity for "same" and "different" color comparisons was then calculated.

Results:

Accuracy
To assess accuracy of participants' duration judgments, a Spearman rank correlation between the actual duration of the trial and the participant's subjective judgment (response made on the number line) was calculated for each participant. All participants in all experiments exhibited a positive Spearman rank correlation value (See **Supplemental Figure 1**). For the purposes of statistical evaluation, rho values were Fisher-transformed, and were significantly above 0 at the

group level in all experiments (Experiment 1: Mean correlation, or z = 1.29, 95% CI = [1.15, 1.43], $t(19) = 19.14$, $p < 0.001$, $d = 4.28$. Experiment 2: Mean z = 1.18, 95% CI = [1.08, 1.27], $t(28) = 24.66$, $p < 0.001$, $d = 4.58$. Experiment 3: Mean z = 1.06, 95% CI = [0.99, 1.13], $t(79) = 29.31$, $p < 0.001$, $d = 3.28$. Experiment 4: Mean z = 1.08, 95% CI = [0.96, 1.20], $t(17) = 18.98$, $p < 0.001$, $d = 4.47$).

To confirm that each participant performed above chance relative to their own response distribution (i.e., irrespective of response bias), a permutation test was also computed for each participant: over 10,000 iterations, each participant's full distribution of responses was shuffled, and the rank correlation between the time of each trial and this shuffled response was computed. Significance was determined as being higher than 95% of this null distribution. All participants in all experiments demonstrated accuracy reliably above chance relative to their own distributions.

Influence of event structure on duration judgments

To examine how event boundaries affect subjective duration judgments, we designed a paradigm in which participants viewed a colored square (0.5 - 5s in Experiments 1-3; 2 - 8s in Experiment 4; **Figure 1A**), which remained a single color for the duration of the trial (continuous condition) or switched colors during the interval (boundary condition). Participants then judged how long the square, regardless of color, was presented on the screen using a continuous timeline (**Figure 1B**). We hypothesized that the color switches act as event boundaries, thus decreasing accessibility of the pre-boundary interval in memory. Critically, if memory for information from across the entire interval is integrated into a duration estimate, then reduced memory access for pre-boundary information would lead to shorter duration judgments in the boundary condition. Across experiments, we manipulated features of the boundary that we hypothesized would affect the accessibility of memory for the interval: position (i.e., where the boundary occurred in the interval) and sharpness (i.e., whether the change was instantaneous or gradual).

In Experiment 1, in which boundary trials switched colors half-way through (**Figure 1C**), we found that participants judged boundary trials to be reliably shorter than continuous trials (Mean difference, M = -0.24, 95% CI = [-0.34, -0.14]), $t(19) = -5.07$, $p < 0.001$, $d = -1.13$. (**Figure 1D**). This finding conceptually replicates prior findings demonstrating a compression of duration judgments for experiences which contain a boundary (Liverence & Scholl, 2012; Bangert et al., 2019; Yousif & Scholl, 2019).

After establishing an influence of event boundaries on duration judgments in this task, we next aimed to understand *why* boundaries lead to the compression of time. One reason why boundary trials might be judged as shorter is because the pre-boundary color is less accessible in working memory (Zwaan, 1996), leaving mostly the post-boundary memory to be integrated into the duration judgment, thus leading to a shorter estimate. If this "flushing" account is true, then judgments of time should be influenced not merely by the presence of a boundary, but should be sensitive to *when* in the interval the boundary occurred. In other words, boundaries which occur earlier in the trial will have more mnemonic content (as relatively less information

was flushed at the boundary), and thus may be judged as longer than trials in which the boundary occurred later in the sequence.

In Experiment 2, we manipulated the placement of the boundary, such that the last event comprised 100% (continuous condition), 75%, 50%, or 25% of the trial (**Figure 1E**). First, we replicated the result of Experiment 1, such that boundary trials were judged on average to be shorter than continuous trials (M = -0.12, 95% CI = [-0.17, -0.07]), $t(28) = -5.00$, $p < 0.001$, $d = -0.93$. To test our primary hypothesis that the extent of compression is modulated by the duration of the second color event, we examined whether the position of the boundary significantly modulated participants' duration judgments (See **Figure 1F**). We found a significant main effect of boundary placement (continuous (100%), 75%, 50%, 25%) on duration judgment ($F(3, 84) = 7.69$, $p < 0.001$, $\eta_p^2 = 0.22$). To assess whether this main effect reflected the scaling of duration judgments with the length of the post-boundary event, we asked whether duration judgments were linearly related to the length of the last event. For each participant, we computed the Pearson correlation between their duration judgments and the relative placement of the boundary, and found a reliable negative correlation across participants (Mean fisher-transformed correlation = -0.57, 95% CI = [-0.90, -0.25], $t(28) = -3.66$, $p = 0.001$, $d = -0.68$). To further probe this relationship, we performed planned pairwise comparisons between neighboring conditions. The 75% boundary trials were rated as significantly shorter than 100% (continuous) trials (M = -0.07, 95% CI = [-0.14, -0.02], $t(28) = -2.64$, $p = 0.013$, $d = -0.49$). The pairwise difference between 75% and 50% trials also reached significance, such that 50% trials were judged as significantly shorter than 75% trials (M = -0.08, 95% CI = [-0.16, -0.00], $t(28) = -2.09$, $p = 0.046$, $d = -0.39$). However, we did not find a significant difference between 50% and 25% trials (M = 0.03, 95% CI = [-0.05, 0.11], $t(28) = 0.77$, $p = 0.446$, $d = 0.14$). These data suggest that subjective time scales with the length of the most recent event up to the halfway point. Thus, once the most accessible information becomes too short (e.g, half the length of the trial), participants may rely on additional heuristic information (e.g., that multiple events occurred).

In Experiment 3, we examined whether abrupt changes (associated with flushing of working memory contents) were necessary to elicit boundary-related temporal compression. In a between-participants design, participants were tested on continuous and either "abrupt" boundary trials (with three color switches, equally spaced over the interval) or "gradual" boundary trials (in which the color smoothly morphed over the entire interval; **Figure 1G**). We hypothesized that if the context drifted gradually, rather than abruptly, over the course of the trial, we would observe a reduced effect of color change on duration judgments. Such gradual shifts may be associated with some forgetting of early information, but we would not necessarily expect a full "flushing" based on prior work showing that prior memories persist when change is gradual (Gershman et al., 2014).

First, we examined each of the four experimental groups separately. Because there were no differences in the boundary effect for the two abrupt change and two gradual change groups, respectively (See Experiment 3 Stimuli Methods; **Supplemental Figure 2**), these groups were collapsed. In the abrupt change group, we robustly replicated the boundary compression effect

(M = -0.17, 95% CI = [-0.24, -0.11]; $t(39) = -5.14$, $p < 0.001$, $d = -0.81$). In the gradual change group, we also replicated the effect (M = -0.08, 95% CI = [-0.15, -0.01]; $t(39) = -2.30$, $p = 0.027$, $d = -0.36$), though the effect was numerically weaker. Consistent with our hypothesis, an ANOVA revealed a marginal interaction between condition (continuous vs. boundary) and group (gradual vs. abrupt change), such that there was reduction in the boundary effect for the gradual group ($F(1, 78) = 3.63$, p = 0.060, $\eta_p^2 = 0.04$; **Figure 1H**). These data suggest that abrupt changes in context (associated with decreased memory accessibility) result in stronger boundary effects, and provide further support that the extent of duration distortion across events is modulated by accessibility to the event representations.

In Experiment 4, we sought to replicate the finding from Experiment 1, with longer (8 sec) trials in the fMRI study (**Figure 1I**). Indeed, we found that boundary trials were judged as significantly shorter than continuous trials (M = -0.31, 95% CI = [-0.61, -0.02]), $t(17) = -2.26$, $p = 0.037$, $d = -0.53$ (**Figure 1J**). We additionally ran this analysis for all trials, both within duration and collapsed across, and found a similar pattern of results **(Supplemental Figure 3A,B)**.

Event Similarity in the Hippocampus

In Experiment 4, our primary interest was how hippocampal representational change within individual trials is related to behavioral duration judgments. To first examine the influence of event structure on neural similarity measures, we examined differences in similarity across individual 8-second trials. Specifically, we extracted the pattern of activity across voxels at the beginning of the trial and at the end of the trial, and correlated those two patterns to measure the representational change across the interval. This was done separately for continuous and boundary trials (**Figure 2A)**. Thus, in the continuous condition, we examined changes in patterns of activity from the beginning to the end of a single event, whereas in the boundary condition, we probed changes in patterns across two events. Importantly, however, the interval length in the two conditions was matched. This analysis allows us to examine representations that occur at the event-level: if a region represents events as a period of neural stability (DuBrow and Davachi, 2014; Baldassano et al., 2017; Ezzyat & Davachi, 2021), then we would expect to see greater similarity for the continuous condition, which contains one event, versus the boundary condition, which contains two. However, if a region is sensitive to the temporal position information within an interval (e.g., the start of an event), then we may expect to see greater similarity in the boundary condition, akin to a “resetting” of the neural population at a boundary (Levy, 1989; Wallenstein et al., 1998; Ben-Yakov et al., 2013; Terada et al., 2017; Ben-Yakov & Henson, 2018; Sun et al., 2020).

We focused on two regions of interest: our primary region of interest, the left hippocampus, and the visual cortex as a control region (See ROI Definition Methods; **Figure 2B**). Specifically, we hypothesized that pattern similarity in the visual cortex would decrease in the boundary condition at a color switch, since it should be sensitive to perceptual similarity. Contrary to our hypothesis, we did not find differences in similarity by condition in the visual cortex, (M = 0.00, 95% CI = [-0.03, 0.04], $t(17) = 0.23$, $p = 0.823$, $d = 0.05$ **(Figure 2C, left)**. In the left hippocampus, however, there was significantly greater event similarity for the boundary condition, relative to continuous, (M = -0.04, 95% CI = [-0.07, -0.01], $t(17) = -2.50$, $p = 0.023$, d

= -0.59. (**Figure 2C, right**), consistent with a “resetting” mechanism by which the boundary recruits a similar neural population as the start of an event.

Color Similarity in Visual Cortex

The event similarity analysis provided evidence that hippocampal patterns, but not visual cortical patterns, were modulated by the presence of an event boundary. To ensure that the dissociation between hippocampus and visual cortex was not just due to differences in task sensitivity, we next asked whether patterns of brain activity were modulated by color, irrespective of events. To isolate color representations irrespective of event representations, we correlated the activity patterns from the second event of one (boundary) trial to the first event of all other trials (**Figure 2D**). We then sorted the data by whether the first and second events were the same color or different colors, and examined differences in pattern similarity as a function of color match. In the visual cortex, we found that pattern similarity was greater for two events of the same color, compared to two events of different colors (M = 0.01, 95% CI = [0.00, 0.01]), $t(17) = 2.50$, $p = 0.023$, $d = 0.59$ (**Figure 2E, left**), indicating visual cortex was sensitive to color information in this task. In contrast, the hippocampus showed no such effect, (M = 0.00, 95% CI = [-0.00, 0.00]), $t(17) = 0.48$, $p = 0.637$, $d = 0.11$ (**Figure 2E, right**).

Relationship between event dissimilarity and duration judgments

Finally, we asked whether pattern change over the course of a single 8-second trial could explain variance in duration judgments across trials. To examine how neural pattern similarity influences duration judgments, we ran a series of mixed-effects linear regressions. Specifically, we tested whether event dissimilarity (1 - ‘Event Similarity’ described above) in left hippocampus and in visual cortex predict duration judgments above and beyond the influence of event boundaries, as well as whether they interacted with condition.

First, we tested a model predicting duration judgment from both condition and left hippocampal pattern dissimilarity, and we found a main effect of hippocampal dissimilarity ($\beta = 0.40$; $\chi^2(1) = 9.68$, $p = 0.002$; **Figure 3A**), such that greater dissimilarity — or more pattern change — in the hippocampus predicted longer duration judgments. There was no significant interaction between hippocampal dissimilarity and condition ($\chi^2(1) = 2.48$, $p = 0.115$). Although this effect of hippocampal dissimilarity on duration judgments was only significant in the continuous condition ($\beta = 0.59$; $\chi^2(1)= 9.10$, $p = 0.003$), it was numerically in the same direction in the boundary condition ($\beta = 0.21$; $\chi^2(1)= 1.59$, $p = 0.207$).

Next, we constructed a model predicting duration judgment from visual cortical dissimilarity, in addition to condition. The model revealed no main effect of visual cortex dissimilarity on duration judgments ($\beta = 0.15$; $\chi^2(1)= 1.53$, $p = 0.216$; **Figure 3B**). However, a model predicting duration judgment from the main effect of condition as well as an interaction between visual cortical dissimilarity and condition, revealed significant interaction between visual cortex dissimilarity and condition ($\chi^2(2) = 8.60$, $p = 0.014$). Specifically, we found that visual cortical dissimilarity predicted duration judgments in the continuous condition ($\beta = 0.45$; $\chi^2(1)= 6.80$, $p = 0.009$) but not the boundary condition ($\beta = -0.19$; $\chi^2(1)= 1.29$, $p = 0.256$).

Given that both the hippocampus and visual cortex dissimilarity exhibited relations with duration judgments in the continuous condition, we asked whether they contribute independent variance. In a model predicting duration judgment in the continuous condition as a function of left hippocampal and visual cortical dissimiarity, we found that the hippocampus significantly contributed to duration judgment (β = 0.46; $\chi^2(1)$= 4.68, p = 0.030), but the visual cortex did not (β = 0.29; $\chi^2(1)$= 2.38, p = 0.123).

Given that many other brain regions have been implicated in time perception at this timescale (Buhusi & Meck, 2005; Umbach et al., 2020), we exploratorily performed this analysis in other regions. Specifically, we assessed whether pattern dissimilarity in the caudate, putamen, middle frontal gyrus (MFG), inferior frontal gyrus (IFG), and entorhinal cortex predicted duration judgments. We did not find main effects of dissimilarity on duration judgments, nor interactions with conditions, in any of these regions (caudate: main effect, β = -0.12, $\chi^2(1)$ =0.63, p = 0.427, interaction, $\chi^2(2)$ =0.89, p = 0.639; putamen: main effect, β = 0.24, $\chi^2(1)$ =0.68, p = 0.411, interaction, $\chi^2(2)$ =0.70, p = 0.703; MFG: β = 0.14, main effect, $\chi^2(1)$ =0.85, p = 0.355, interaction, $\chi^2(2)$ =1.42, p = 0.493; IFG: main effect, β = 0.27, $\chi^2(1)$ =3.38, p = 0.066, interaction, $\chi^2(2)$ = 3.83, p = 0.147; entorhinal cortex: main effect, β = 0.052, $\chi^2(1)$ =0.43, p = 0.510, interaction, $\chi^2(2)$ = 0.50, p = 0.778).

Discussion:

Using event segmentation to manipulate mnemonic content, we found that duration judgments are influenced by the accessibility of mnemonic representations. After demonstrating the event boundaries lead to reduced duration judgments in a novel task (Experiment 1), we found that duration judgments scale (to a limit) with the duration of the *most recent* event (Experiment 2) and that this reduction is attenuated when there is a gradual change in context that may maintain access to previous mnemonic information, rather than an abrupt shift (Experiment 3). Lastly, we identified that trial-by-trial neural pattern change in the left hippocampus predicts longer duration judgments (Experiment 4).

Relation to studies of event boundaries in short- and long-term duration judgments

Under a context change account of duration judgments, in which change is used to infer the passage of time, one might expect that boundary trials should be judged as longer than continuous trials (Poynter, 1983; Zakay et al., 1994) because the context across two events differs more than the context across a single event. Yet, our results show the opposite: intervals with a boundary are estimated as *shorter* than equivalent, continuous intervals. Rather, our results point to a memory-based account of duration judgments, in which active mnemonic information is used to infer the passage of time: event boundaries disrupt access to pre-boundary information (Radvanksy, 2012), leading to decreased time judgments despite greater context change. These results converge with prior findings that event boundaries lead to a compression of perceived time (Liverence & Scholl, 2012; Bangert et al., 2019; Yousif & Scholl, 2019).

In contrast, work at longer time scales shows precisely the opposite: intervals spanning an event boundary are remembered as *longer* than equivalent continuous intervals, and duration

judgments scale with the number of events (Poynter, 1983; Zakay et al., 1994; Ezzyat & Davachi, 2014; Faber & Gennari, 2015; Lositsky et al., 2016; Faber & Gennari, 2017). This discrepancy can be explained by the memory-based account, as event boundaries may affect mnemonic content differently in working versus long-term memory. That is, if event boundaries disrupt memory at encoding, there is less information in working memory, leading to an in-the-moment compression. However, event boundaries are often better remembered (Polyn et al., 2009; Swallow et al., 2009; Heusser et al., 2018; Rouhani et al., 2020), and may serve as anchor points to within-event content in long-term memory (DuBrow & Davachi, 2016; Heusser et al., 2018; Michelmann et al., 2019; see Shin & DuBrow, 2021 for review). If intervals containing event boundaries have greater mnemonic content in long-term memory, this may lead to the inference that more time has passed.

Event representations in the hippocampus

In examining the influence of event boundaries on hippocampal representations, we found increased pattern similarity across trials containing a boundary, relative to continuous trials. Given that events are thought to be a period of contextual stability (DuBrow et al., 2017), this finding of relatively less stability within an event may seem counterintuitive. However, our finding converges with theoretical and empirical work that some hippocampal neurons are sensitive to specific temporal positions within an event (Levy, 1989; Wallenstein et al., 1998; Ginther et al., 2011; Terada et al., 2017; Sun et al., 2020). If there are neurons that code for the “beginnings” of events, an event boundary might recruit the same neural population as the beginning of the trial, thus resulting in greater similarity across the two events within the trial; in contrast, for a continuous trial, the hippocampal representations would continue to grow dissimilar over time.

Alternatively, the increased pattern similarity across boundaries could be explained by enhanced pattern separation within events: such separation could orthogonalize experiences within an event and reduce within-context interference. Recent work finding that decreased pattern similarity for clips from the same, versus different, movies (Benear et al., 2020) is consistent with this hypothesis (but see Hsieh et al., 2014; Milivojevic et al., 2016). To distinguish between a position sensitivity vs. pattern separation account, it would be useful for future work with higher temporal precision to examine whether hippocampal pattern similarity across boundaries is driven by a change occurring specifically at the moment of the boundary, or continuous pattern change across the interval.

The role of the hippocampus in tracking time

Our finding that hippocampal pattern change correlates with subjective duration judgments is consistent with its proposed role in representing time (Eichenbaum, 2013). In humans, hippocampal pattern stability has been related to remembered temporal proximity, such that events remembered as “closer” together show greater pattern stability (Ezzyat & Davachi, 2014; Deuker et al., 2016), consistent with our current results. Together, these findings suggest that temporal dissimilarity in hippocampal representations may represent subjective distance and duration, across different timescales of both working and long-term memory.

Notably, a recent study found that individual sequences (item + timing information), but not time alone, could be decoded from hippocampal patterns (Thavabalasingam et al., 2019). Although these findings suggest that the hippocampus may not be able to represent temporal information in isolation, it is notable that participants' task was to encode the sequences, not to track time. This raises questions of whether hippocampal representations are shaped by task demands, rather than automatically representing temporal information. Additionally, given that we find that hippocampal patterns are related to *subjective* estimates of time, it is possible that the failure to decode actual time could be because the hippocampus codes for subjective, but not objective, duration.

Although we find that hippocampal patterns correlate with duration judgments, it is of note that hippocampal damage does not consistently affect subjective duration judgments (Richards, 1973; Perbal et al., 2001; Melgire et al., 2005; Noulhiane et al., 2007). Complementary studies using a similar paradigm in rodents (Jacobs et al., 2013) and humans (Palombo et al., 2016) found that hippocampal damage impairs duration judgments at long, but not short durations. However, the exact time scale at which processing becomes hippocampal-dependent is the subject of ongoing research, may be shorter than previously established (Sabariego et al., 2020), and may depend on the context of the task (Palombo et al., 2019). Our findings contribute to this growing evidence that the hippocampus contributes to temporal processing at short timescales.

Conclusions

One theoretical position is that brains, in fact, do not sense time; just because activity in a brain region is correlated with the passage of time does not mean that the region is clocking time per se (Ezzyat & Davachi, 2014; Davachi & DuBrow, 2015; Buzsáki & Llinás, 2017). Although here we report neural measurements that correlate with subjective duration judgments, we suggest that our neural results are a measure of event *memory* used to infer the passage of time, rather than time per se. Our findings thus raise questions about the extent to which our sense of time arises from pure timing signals in the brain, or to what extent time is a product of — or reconstruction from — our memories.

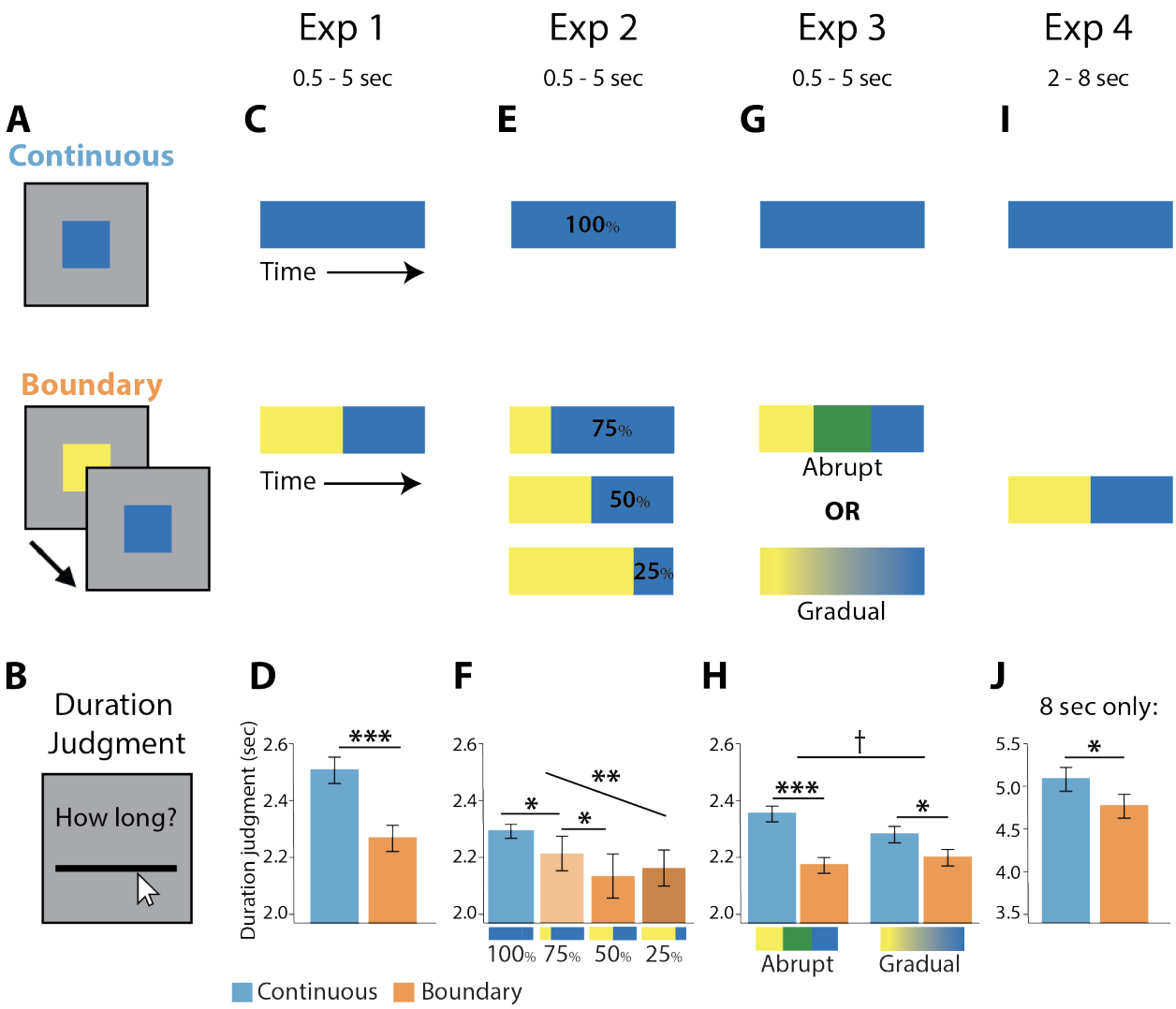

Figure 1. Behavioral Task & Data A) Participants viewed a colored square, which either stayed the same color for the total duration (continuous condition) or switched colors during the interval (boundary condition). B) Participants then estimated how long the square was on the screen for (regardless of color change). C) Schematic of trial structures. In experiment 1, boundaries always occurred halfway through the total 0.5-5sec interval. D) Experiment 1 results: boundary trials were judged as significantly shorter than continuous trials. E) In experiment 2, color switches in boundary conditions occurred either one-quarter, one-half, or three-quarters through the total 0.5-5sec interval, such that the second event was 75%, 50%, or 25% of the total duration, respectively. F) Experiment 2 results: there was a significant linear effect of boundary placement; pairwise comparisons showed the 75% trials were judged to be reliably shorter than 100% trials, and 50% trials were judged as reliably shorter than 75% trials. G) In experiment 3, boundary trials were either abrupt (two color switches, each color presented for one-third of the 0.5-5sec interval) or gradual (smoothly morphing through color space for the entire 0.5-5sec interval). H) Experiment 3 results: in both groups, boundary trials were judged as reliably shorter than continuous trials, though this effect was weaker in the gradual change group. I) As in experiment 1, boundaries occurred halfway through the interval, though the interval was extended to be 2-8 sec. J) Experiment 4 results: boundary trials were

judged as reliably shorter than continuous trials. *** p < .001, two-tailed; ** p < .005, two-tailed; *p < .05, two-tailed; † p < .07, two-tailed. Error bars denote the within-participant standard error of the mean.

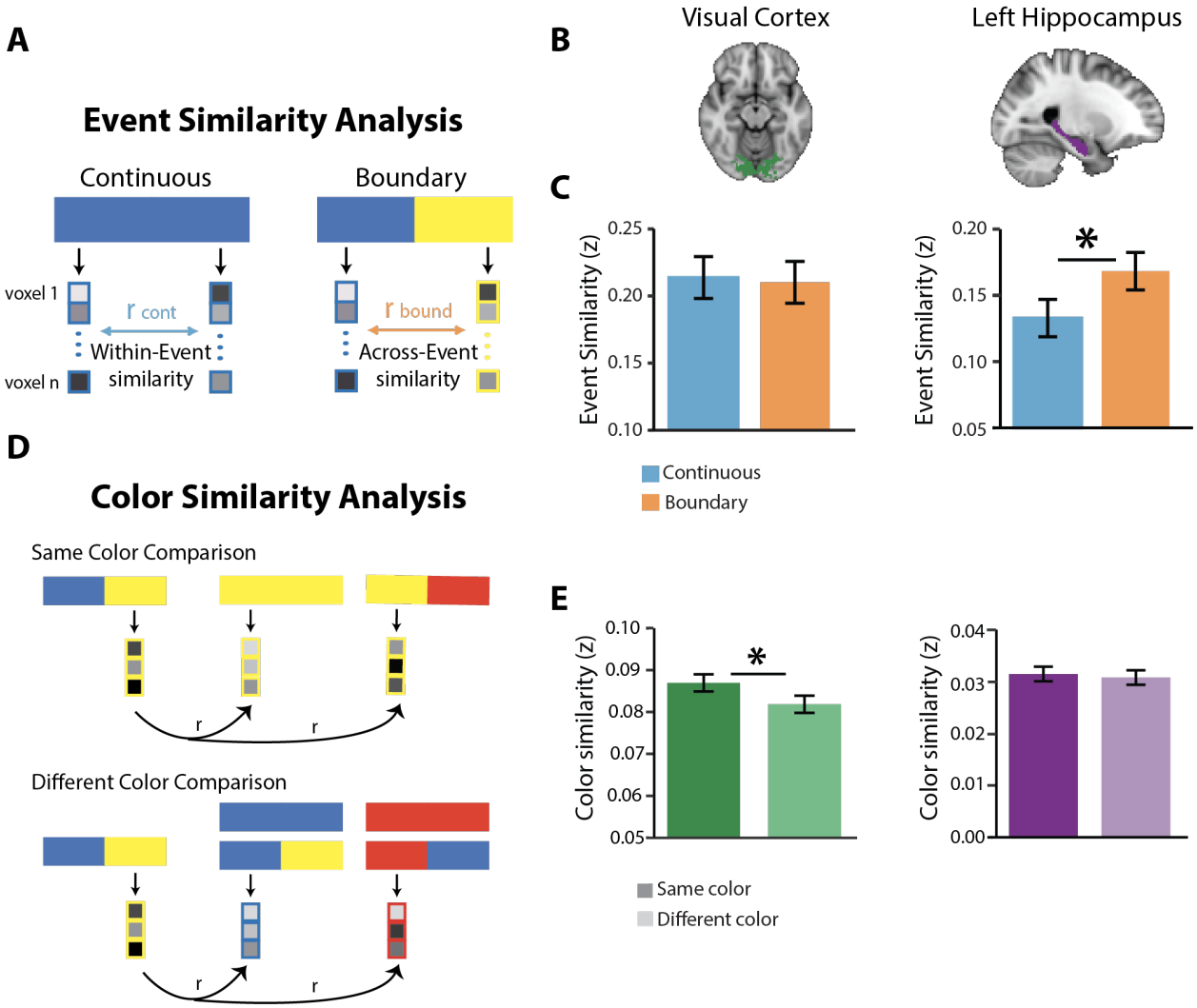

Figure 2. Pattern Similarity Approaches A) Event similarity was computed as the fisher-transformed Pearson correlation between the estimated evoked activity at the beginning of an 8 second trial and the estimated evoked activity at the end of that 8 second trial. Correlations were computed separately for continuous sand boundary trials. B) Visual cortex did not show differential event similarity by condition. C) The left hippocampus exhibited greater event similarity across boundary trials than across continuous trials. D) Color similarity was computed as the fisher-transformed Pearson correlation between the estimated evoked activity at the end of an 8 second trial, to the estimated evoked activity at the beginning of another trial that shares the same color. As a comparison, “different color” trials were computed as the correlation between the patterns at the end of one trial to the beginning of all other trials which did not share its color. E) The visual cortex exhibited greater similarity for trials of the same color than trials of a different color. F) Left hippocampus did not show differential similarity as a function of

trial color. * p < .05, two-tailed. Error bars denote the within-participant standard error of the mean.

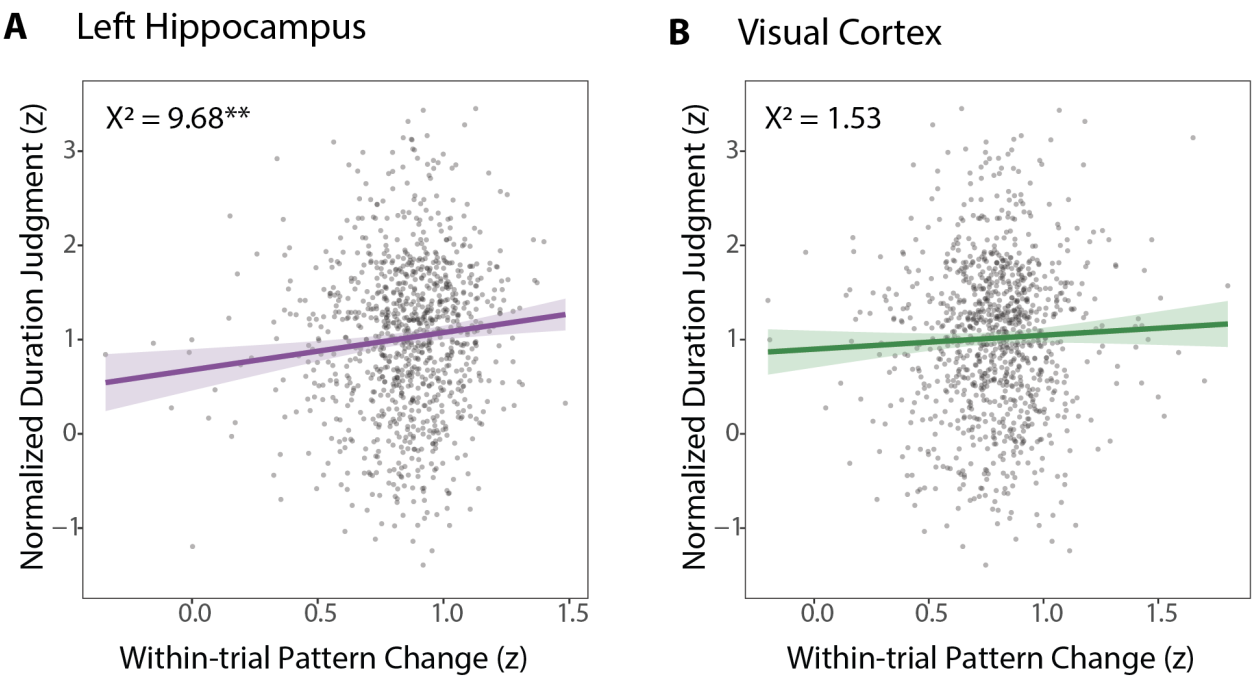

Figure 3. Results of mixed-effects linear model predicting time judgment by event dissimilarity. A) Event dissimilarity in the left hippocampus predicted duration judgments. Plotted from a model predicting time judgment as a function of: condition, a main effect of a main effect of hippocampal dissimilarity, and a by-participant random effect of condition. B) Event dissimilarity in visual cortex did not predict time judgment across conditions. Plotted from a model predicting time judgment as a function of: condition, a main effect of a main effect of visual cortical dissimilarity, and a by-participant random effect of condition.

Acknowledgments

We dedicate this work to Warren Meck and Howard Eichenbaum, pioneers in the study of time and the brain but, more importantly, both incredible people.

We thank Emily Cowan and Alexa Tompary for helpful conversations, and Sami Yousif for feedback on a previous version of the manuscript.

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.

Bangert, A. S., Kurby, C. A., Hughes, A. S., & Carrasco, O. (2019). Crossing event boundaries changes prospective perceptions of temporal length and proximity. *Attention, Perception, & Psychophysics*, 1-14.

Ben-Yakov, A., Eshel, N., & Dudai, Y. (2013). Hippocampal immediate poststimulus activity in the encoding of consecutive naturalistic episodes. *Journal of Experimental Psychology: General*, *142*(4), 1255.

Ben-Yakov, A., & Henson, R. N. (2018). The hippocampal film editor: Sensitivity and specificity to event boundaries in continuous experience. *Journal of Neuroscience*, *38*(47), 10057-10068.

Benear, S. L., Horwath, E. A., Cowan, E., Camacho, M. C., Ngo, C., Newcombe, N., ... & Murty, V. P. (2020).Children show adult-like hippocampal pattern similarity for familiar but not novel events. *psyArxiv.*

Brouwer, G. J., & Heeger, D. J. (2009). Decoding and reconstructing color from responses in human visual cortex. *Journal of Neuroscience*, *29*(44), 13992-14003.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature reviews neuroscience*, *6*(10), 755-765.

Buzsáki, G., & Llinás, R. (2017). Space and time in the brain. *Science*, *358*(6362), 482-485.

Cai, D. J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., ... & Flores, S. E. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. *Nature*, *534*(7605), 115-118.

Clewett, D., Gasser, C., & Davachi, L. (2020). Pupil-linked arousal signals track the temporal organization of events in memory. *Nature Communications*, *11*(1), 1-14.

Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: the role of prediction and context. *Trends in Cognitive Sciences*, *19*(2), 92-99.

Deuker, L., Bellmund, J. L., Schröder, T. N., & Doeller, C. F. (2016). An event map of memory space in the hippocampus. *Elife*, *5*, e16534.

DuBrow, S., & Davachi, L. (2014). Temporal memory is shaped by encoding stability and intervening item reactivation. *Journal of Neuroscience*, *34*(42), 13998-14005.

DuBrow, S., & Davachi, L. (2016). Temporal binding within and across events. *Neurobiology of learning and memory*, *134*, 107-114.

DuBrow, S., Rouhani, N., Niv, Y., & Norman, K. A. (2017). Does mental context drift or shift?. *Current Opinion in Behavioral Sciences*, *17*, 141-146.

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences*, *11*(12), 504-513.

Eichenbaum, H. (2013). Memory on time. *Trends in Cognitive Sciences*, *17*(2), 81-88.

Ezzyat, Y., & Davachi, L. (2011). What constitutes an episode in episodic memory? *Psychological Science*, *22*(2), 243-252.

Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron*, *81*(5), 1179-1189.

Ezzyat, Y., & Davachi, L. (2021). Neural evidence for representational persistence within events. Journal of Neuroscience.

Faber, M., & Gennari, S. P. (2015). In search of lost time: Reconstructing the unfolding of events from memory. *Cognition*, *143*, 193-202.

Faber, M., & Gennari, S. P. (2017). Effects of learned episodic event structure on prospective duration judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(8), 1203.

Gershman, S. J., Radulescu, A., Norman, K. A., & Niv, Y. (2014). Statistical computations underlying the dynamics of memory updating. *PLoS Comput Biol*, *10*(11), e1003939.

Ginther, M. R., Walsh, D. F., & Ramus, S. J. (2011). Hippocampal neurons encode different episodes in an overlapping sequence of odors task. *Journal of Neuroscience*, *31*(7), 2706-2711.

Heusser, A. C., Ezzyat, Y., Shiff, I., & Davachi, L. (2018). Perceptual boundaries cause mnemonic trade-offs between local boundary processing and across-trial associative binding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(7), 1075.

Horner, A. J., Bisby, J. A., Wang, A., Bogus, K., & Burgess, N. (2016). The role of spatial boundaries in shaping long-term event representations. *Cognition*, *154*, 151-164.

Hsieh, L. T., Gruber, M. J., Jenkins, L. J., & Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron*, *81*(5), 1165-1178.

Insausti, R., Juottonen, K., Soininen, H., Insausti, A. M., Partanen, K., Vainio, P., ... & Pitkänen, A. (1998). MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *American Journal of Neuroradiology*, *19*(4), 659-671.

Jacobs, N. S., Allen, T. A., Nguyen, N., & Fortin, N. J. (2013). Critical role of the hippocampus in memory for elapsed time. *Journal of Neuroscience*, *33*(34), 13888-13893.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143-156.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825-841.

Lee, A. C., Thavabalasingam, S., Alushaj, D., Çavdaroğlu, B., & Ito, R. (2020). The hippocampus contributes to temporal duration memory in the context of event sequences: A cross-species perspective. *Neuropsychologia*, *137*, 107300.

Levy, W. B. (1989). A computational approach to hippocampal function. In *Psychology of learning and motivation* (Vol. 23, pp. 243-305). Academic Press.

Liverence, B. M., & Scholl, B. J. (2012). Discrete events as units of perceived time. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 549.

Lositsky, O., Chen, J., Toker, D., Honey, C. J., Shvartsman, M., Poppenk, J. L., ... & Norman, K. A. (2016). Neural pattern change during encoding of a narrative predicts retrospective duration estimates. *Elife*, *5*, e16070.

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal “time cells” bridge the gap in memory for discontiguous events. *Neuron*, *71*(4), 737-749.

MacDonald, C. J., Carrow, S., Place, R., & Eichenbaum, H. (2013). Distinct hippocampal time cell sequences represent odor memories in immobilized rats. *Journal of Neuroscience*, *33*(36), 14607-14616.

Mankin, E. A., Sparks, F. T., Slayyeh, B., Sutherland, R. J., Leutgeb, S., & Leutgeb, J. K. (2012). Neuronal code for extended time in the hippocampus. *Proceedings of the National Academy of Sciences*, *109*(47), 19462-19467.

Manns, J. R., Howard, M. W., & Eichenbaum, H. (2007). Gradual changes in hippocampal activity support remembering the order of events. *Neuron*, *56*(3), 530-540.

Meck, W. H., Church, R. M., & Olton, D. S. (1984). Hippocampus, time, and memory. *Behavioral Neuroscience*, *98*(1), 3.

Melgire, M., Ragot, R., Samson, S., Penney, T. B., Meck, W. H., & Pouthas, V. (2005). Auditory/visual duration bisection in patients with left or right medial-temporal lobe resection. *Brain and Cognition*, *58*(1), 119-124.

Michelmann, S., Staresina, B. P., Bowman, H., & Hanslmayr, S. (2019). Speed of time-compressed forward replay flexibly changes in human episodic memory. *Nature Human Behaviour*, *3*(2), 143-154.

Milivojevic, B., Varadinov, M., Grabovetsky, A. V., Collin, S. H., & Doeller, C. F. (2016). Coding of event nodes and narrative context in the hippocampus. *Journal of Neuroscience*, *36*(49), 12412-12424.

Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language*, *28*(3), 292-312.

Naya, Y., & Suzuki, W. A. (2011). Integrating what and when across the primate medial temporal lobe. *Science*, *333*(6043), 773-776.

Nielson, D. M., Smith, T. A., Sreekumar, V., Dennis, S., & Sederberg, P. B. (2015). Human hippocampus represents space and time during retrieval of real-world memories. *Proceedings of the National Academy of Sciences*, *112*(35), 11078-11083.

Noulhiane, M., Pouthas, V., Hasboun, D., Baulac, M., & Samson, S. (2007). Role of the medial temporal lobe in time estimation in the range of minutes. *Neuroreport*, *18*(10), 1035-1038.

Palombo, D. J., Keane, M. M., & Verfaellie, M. (2016). Does the hippocampus keep track of time?. *Hippocampus*, *26*(3), 372-379.

Palombo, D. J., Di Lascio, J. M., Howard, M. W., & Verfaellie, M. (2019). Medial temporal lobe amnesia is associated with a deficit in recovering temporal context. *Journal of Cognitive Neuroscience*, *31*(2), 236-248.

Patenaude, B., Smith, S. M., Kennedy, D. N., & Jenkinson, M. (2011). A Bayesian model of shape and appearance for subcortical brain segmentation. *NeuroImage*, *56*(3), 907-922.

Perbal, S., Ehrlé, N., Samson, S., Baulac, M., & Pouthas, V. (2001). Time estimation in patients with right or left medial-temporal lobe resection. *Neuroreport*, *12*(5), 939-942.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). Task context and organization in free recall. *Neuropsychologia*, *47*(11), 2158-2163.

Poynter, W. D. (1983). Duration judgment and the segmentation of experience. *Memory & Cognition*, *11*(1), 77-82.

Pruessner, J. C., Köhler, S., Crane, J., Pruessner, M., Lord, C., Byrne, A., ... & Evans, A. C. (2002). Volumetry of temporopolar, perirhinal, entorhinal and parahippocampal cortex from high-resolution MR images: considering the variability of the collateral sulcus. *Cerebral Cortex*, *12*(12), 1342-1353.

Radvansky, G. A. (2012). Across the event horizon. *Current Directions in Psychological Science*, *21*(4), 269-272.

Rattat, A. C., & Droit-Volet, S. (2012). What is the best and easiest method of preventing counting in different temporal tasks?. *Behavior Research Methods*, *44*(1), 67-80.

Richards, W. (1973). Time reproductions by HM. *Acta Psychologica*, *37*(4), 279-282.

Rinck, M., & Bower, G. H. (2000). Temporal and spatial distance in situation models. *Memory & Cognition*, *28*(8), 1310-1320.

Rouhani, N., Norman, K. A., Niv, Y., & Bornstein, A. M. (2020). Reward prediction errors create event boundaries in memory. *Cognition*, *203*, 104269.

Rubin, A., Geva, N., Sheintuch, L., & Ziv, Y. (2015). Hippocampal ensemble dynamics timestamp events in long-term memory. *Elife*, *4*, e12247.

Sabariego, M., Tabrizi, N.S., Marshall, G.J., McLagan, A.N., Jawad, S., & Hales, J.B. (2020). In the temporal organization of episodic memory, the hippocampus supports the experience of elapsed time. *Hippocampus, 2020*, 1-10, https://doi.org/10.1002/hipo.23261.

Sakon, J. J., Naya, Y., Wirth, S., & Suzuki, W. A. (2014). Context-dependent incremental timing cells in the primate hippocampus. *Proceedings of the National Academy of Sciences*, *111*(51), 18351-18356.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of neurology, neurosurgery, and psychiatry*, *20*(1), 11.

Shin, Y. S., & DuBrow, S. (2021). Structuring memory through inference-based event segmentation. *Topics in Cognitive Science*, *13*(1), 106-127.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143-155.

Speer, N. K., & Zacks, J. M. (2005). Temporal changes as event boundaries: Processing and memory consequences of narrative time shifts. *Journal of Memory and Language*, *53*(1), 125-140.

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236.

Sun, C., Yang, W., Martin, J., & Tonegawa, S. (2020). Hippocampal neurons represent events as transferable units of experience. *Nature Neuroscience*, *23*(5), 651-663.

Terada, S., Sakurai, Y., Nakahara, H., & Fujisawa, S. (2017). Temporal and rate coding for discrete event sequences in the hippocampus. *Neuron*, *94*(6), 1248-1262.

Thavabalasingam, S., O'Neil, E. B., Tay, J., Nestor, A., & Lee, A. C. (2019). Evidence for the incorporation of temporal duration information in human hippocampal long-term memory sequence representations. *Proceedings of the National Academy of Sciences*, *116*(13), 6407-6414.

Turner, B. O., Mumford, J. A., Poldrack, R. A., & Ashby, F. G. (2012). Spatiotemporal activity estimation for multivoxel pattern analysis with rapid event-related designs. *NeuroImage*, *62*(3), 1429-1438.

Umbach, G., Kantak, P., Jacobs, J., Kahana, M. J., Pfeiffer, B. E., Sperling, M., & Lega, B. (2020). Time cells in the human hippocampus and entorhinal cortex support episodic memory. *Proceedings of the National Academy of Sciences, 117*(45), 28463–28474.

Wallenstein, G. V., Hasselmo, M. E., & Eichenbaum, H. (1998). The hippocampus as an associator of discontiguous events. *Trends in Neurosciences*, *21*(8), 317-323.

Wang, L., Mruczek, R. E., Arcaro, M. J., & Kastner, S. (2015). Probabilistic maps of visual topography in human cortex. *Cerebral Cortex*, *25*(10), 3911-3931.

Wittmann, M. (2013). The inner sense of time: how the brain creates a representation of duration. *Nature Reviews Neuroscience*, *14*(3), 217-223.

Worsley, K. J. (2001). Statistical analysis of activation images. *Functional MRI: An Introduction to Methods*, *14*(1), 251-70.

Yousif, S. R., & Scholl, B. J. (2019). The one-is-more illusion: Sets of discrete objects appear less extended than equivalent continuous entities in both space and time. *Cognition*, *185*, 121-130.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3.

Zakay, D., Tsal, Y., Moses, M., & Shahar, I. (1994). The role of segmentation in prospective and retrospective time estimation processes. *Memory & Cognition*, *22*(3), 344-351.

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1196.