Structuring time: The hippocampus constructs sequence memories that generalize temporal relations across experiences

Jacob L. S. Bellmund[1], Lorena Deuker[2], Nicole D. Montijn[3], Christian F. Doeller[1,4]

1: Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
2: Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands
3: Department of Clinical Psychology, Utrecht University, Utrecht, the Netherlands
4: Kavli Institute for Systems Neuroscience, Centre for Neural Computation, The Egil and Pauline Braathen and Fred Kavli Centre for Cortical Microcircuits, Jebsen Centre for Alzheimer's Disease, Norwegian University of Science and Technology, Trondheim, Norway
Correspondence: bellmund@cbs.mpg.de

Abstract

The hippocampal-entorhinal region supports memory for episodic details, such as temporal relations of sequential events, and mnemonic constructions combining experiences for inferential reasoning. However, it is unclear whether hippocampal event representations reflect temporal relations derived from mnemonic constructions, event order, or elapsing time, and whether they generalize temporal relations across similar sequences. Here, participants mnemonically constructed times of events from multiple sequences using infrequent cues and their experience of passing time. After learning, event representations in the anterior hippocampus reflected sequence relations based on constructed times. These event representations generalized across sequences, revealing distinct representational formats for events from the same or different sequences. Structural knowledge about time patterns, abstracted from different sequences, biased the construction of specific event times. These findings demonstrate that the hippocampus reconciles representations of specific relations with the generalization across different episodes, consistent with memory-based constructions combining episodic details and general knowledge to simulate scenarios.

Introduction

Our memories are not veridical records, but constructions of our past (Bartlett, 1932). When constructing scenarios of the past or future, we often combine specific episodic details with general, semantic knowledge (Cheng et al., 2016; Hassabis and Maguire, 2007; Irish and Piguet, 2013; Schacter and Addis, 2007, 2020; Schacter et al., 2017). For example, we can infer the time when an event took place not only from episodic details but also from associative or contextual information and general knowledge (Friedman, 1993, 2004). To answer the question when you left for work yesterday, you may combine knowledge about usually departing from home around 8:30 a.m. with the specific sequence of events that unfolded – eating breakfast while listening to the 8 a.m. news and arriving at work a few minutes late for the 9 a.m. meeting despite good traffic conditions on your commute. You infer that you left later than usual, at around 8:40 a.m. Thus, constructive mnemonic processes allow you to estimate when this event occurred, even if a specific event time is not part of the original memory (Friedman, 1993, 2004). Event representations in the hippocampal-entorhinal region carry information about sequence relationships (Bellmund et al., 2020a; Ranganath and Hsieh, 2016), but whether this goes back to mnemonic construction is unclear. Next to its role in memory for specific sequences, the hippocampal-entorhinal region also generalizes across experiences via the abstraction of structural regularities and the recombination of information across episodes (Behrens et al., 2018; Zeithamova and Bowman, 2020), suggesting you may use knowledge about comparable mornings to recall your departure time. Here, we ask whether temporal event relations are generalized across sequences that share a similar structure and address the question how mnemonic construction and generalization combine in the hippocampus and in participants' memory for event times.

In line with its well-established role in episodic memory, the hippocampal-entorhinal region is centrally involved in processing and remembering specific event sequences (Bellmund et al., 2020a). For instance, learning sequences recruits the hippocampus and entorhinal cortex (Kumaran and Maguire, 2006a, 2006b), and hippocampal activity increases at event boundaries delineating sequences (Baldassano et al., 2017; Ben-Yakov and Dudai, 2011). Hippocampal multi-voxel patterns are sensitive to objects shown at learned sequence positions (Hsieh et al., 2014), and recent work suggests that the hippocampus incorporates the duration of intervals between elements in sequence representations

(Thavabalasingam et al., 2018, 2019). Further, pattern correlations in the hippocampus and entorhinal cortex relate to memory for temporal relations (Deuker et al., 2016; DuBrow and Davachi, 2014; Ezzyat and Davachi, 2014; Jenkins and Ranganath, 2016; Kyle et al., 2015; Lositsky et al., 2016).

Hippocampal and entorhinal representations of events occurring in sequence reflect the temporal relations of these events. In one experiment, participants learned the spatial and temporal relationships of events encountered in sequence along a route through a virtual city (Bellmund et al., 2019; Deuker et al., 2016). After relative to before learning, pattern similarity in the anterior hippocampus elicited by event images reflected the sequence relationships between pairs of events (Deuker et al., 2016). Pattern similarity correlated negatively with temporal distances, so that events close in time elicited more similar activity patterns relative to events separated by longer intervals. Within the entorhinal cortex, this effect was specific to the anterior-lateral subregion (Bellmund et al., 2019), consistent with the involvement of this area in precise temporal memory recall (Montchal et al., 2019; Evensmoen et al., 2020).

However, whether event representations in the anterior hippocampus and anterior-lateral entorhinal cortex reflect temporal distances based on actively constructed event times is unclear. Alternatively, these representations of temporal structure could go back to the order of events. For example, successive events could be linked together, resulting in representations of sequence order, where temporal distances are defined based on the number of associative links between events (Ebbinghaus, 1885; Lewandowsky and Murdock, 1989; Jensen and Lisman, 2005). Another possibility is that temporal structure representations arise through elapsing time more passively. For example, the firing of individual entorhinal neurons changes with varying time constants in rodents and non-human primates, allowing time to be decoded from population activity (Bright et al., 2020; Tsao et al., 2018). Slowly drifting activity patterns could be incorporated into event memories as temporal tags, providing a potential mechanism for temporal memory (Howard and Kahana, 2002). Here, we tested whether event representations reflect temporal relations based on mnemonically constructed event times, even when accounting for event order and objectively elapsing time.

In contrast to its role in memory for specific associations such as temporal relations of events in a sequence, the hippocampal-entorhinal region also integrates information across different episodes for inferential reasoning and generalization (Behrens et al., 2018; Kumaran and McClelland, 2012; Zeithamova and Bowman, 2020). Work in rodents and humans demonstrates that the hippocampus supports transitive inference, which requires inferring novel relations between stimulus pairs from knowledge about previously learned premise pairs (Bunsey and Eichenbaum, 1996; Heckers et al., 2004; Park et al., 2020). Further, it combines separately learned associations, enabling inferences about shared associations (Dusek and Eichenbaum, 1997; Koster et al., 2018; Preston et al., 2004; Schlichting et al., 2015; Shohamy and Wagner, 2008; Zeithamova and Preston, 2010; Zeithamova et al., 2012). How the hippocampus contributes to both memory specificity and generalization is a long-standing debate (McClelland et al., 1995; Kumaran and McClelland, 2012; Schapiro et al., 2017).

Recent work suggests a central role for the entorhinal cortex in the abstraction of structural knowledge that is linked to sensory experience in the hippocampus (Behrens et al., 2018; Whittington et al., 2020). Indeed, entorhinal activity patterns reflected structural similarities between choice options in a reinforcement learning task (Baram et al., 2020). Furthermore, in an associative inference task, hippocampal activity patterns carried information about the shared internal structure of image triads such that the hippocampal representational geometry was generalized across triads (Morton et al., 2020). Applying abstract structural knowledge enables flexible behavior through the generalization of relations to novel situations (Behrens et al., 2018; Whittington et al., 2020). In concert with representations of temporal relations between events in the hippocampal-entorhinal region (Bellmund et al., 2019; Deuker et al., 2016), this raises the question whether event representations generalize temporal relations across sequences sharing an underlying structure.

Notably, prior knowledge about structural regularities and semantic associations can bias mnemonic constructions. When estimating the size of studied images, participants' reconstructions were systematically distorted towards category averages (Hemmer and Steyvers, 2009; Hemmer et al., 2015). For relatively small fruits like strawberries, participants tended to overestimate the studied size, whereas they consistently underestimated sizes of large fruits like pineapples. This resulted in an overall bias towards the category mean of all fruits (Hemmer and Steyvers, 2009). Consistent with the notion that learned event structures contribute to event cognition (Franklin et al., 2020; Radvansky and Zacks, 2014; Zacks, 2020), external and semantic details are used to furnish past and future scenarios when few episodic details are generated (Irish et al., 2012; Devitt et al., 2017). When estimating the times of events from a movie, which was terminated prematurely, participants underestimated when events took place for events close to the end of the presented section, possibly due to prior knowledge about the typical structure of movie plots (Frisoni et al., 2021). These findings suggest that abstract knowledge about general patterns could systematically distort mnemonic constructions of

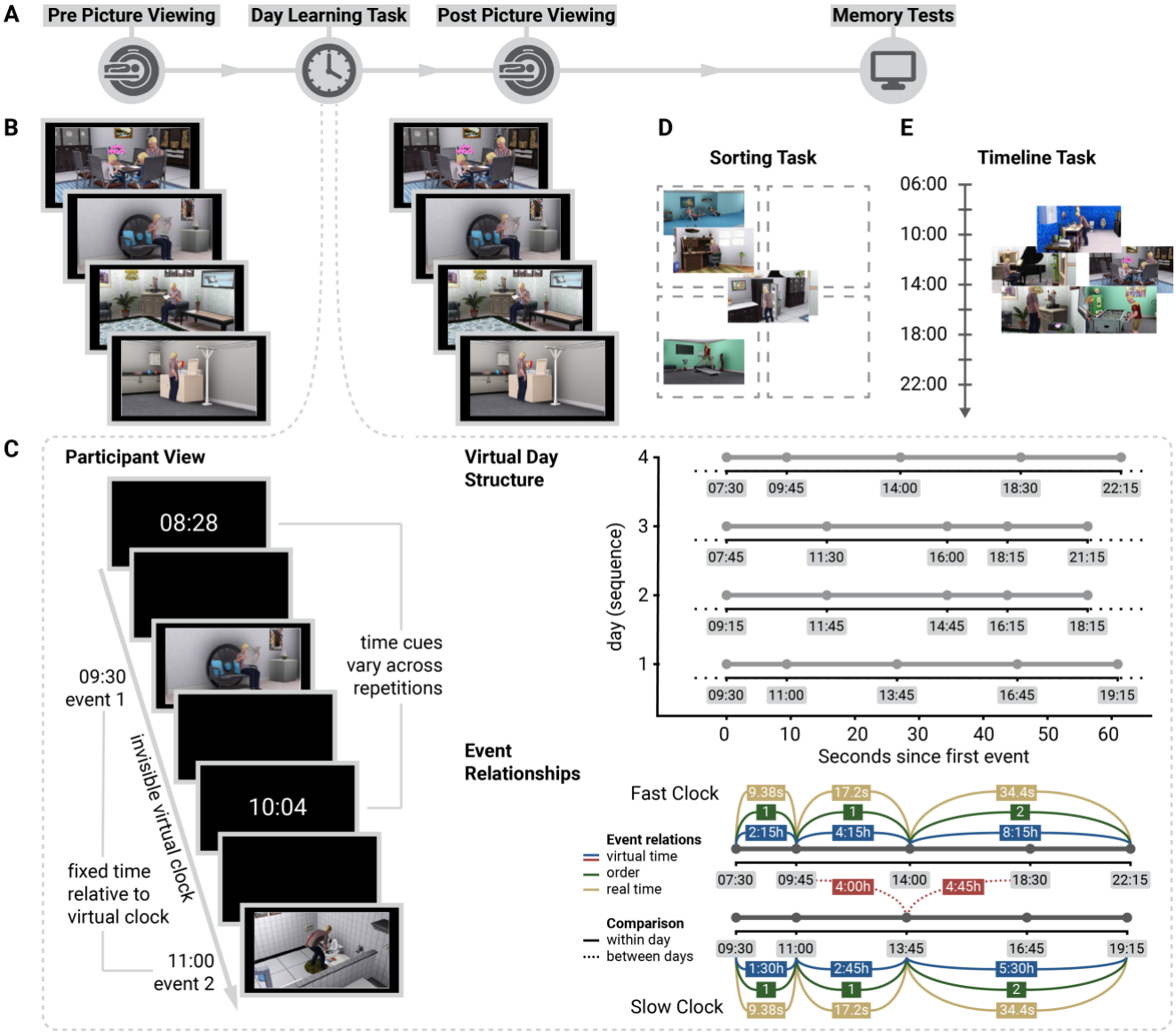


Figure 1. Experimental Design. A. Overview of the experiment. **B**. In the picture viewing tasks before and after learning, participants saw event images presented in the same random order and using identical stimulus timings. **C**. The day learning task took place in between the picture viewing tasks. Participants learned four sequences (virtual days) of five events each (Supplemental Figure 1) and inferred when events took place relative to a virtual clock. Left: The virtual clock ran hidden in the background for each sequence and was revealed only once in between successive events. These time cues varied across repetitions of a sequence, but events occurred at consistent points in virtual time. The duration of blank screen periods varied according to the interval between the indicated time and the event time. Thus, participants had to mentally construct event times by combining their experience of elapsing time with the time cues. Top right: The hidden clock ran at a fixed speed relative to real time for a given sequence, but its speed varied between sequences (Supplemental Figure 2). Bottom right: Different time metrics capture the temporal structure of the event sequences. Event relations can be quantified using temporal distances relative to the hidden clock (virtual time), sequence positions (order), and elapsed time in seconds (real time). While these metrics inevitably covary, they are partially dissociated by the clock speed manipulation. Virtual temporal distances can be quantified both within (solid lines) and across sequences (dotted lines). **D,E.** Participants' memory of the sequences was tested in two tasks. In the sorting task (**D**), participants sorted the scenes according to the four different sequences. In the timeline task (**E**), participants positioned the five event images of a given sequence next to a timeline to indicate constructed event times. **B-E**.The Sims 3 and screenshots of it are licensed property of Electronic Arts, Inc.

event times. If, as in the introductory example, you usually leave for work at 8:30 a.m., this may bias the estimate of your departure time on the specific day you arrived late towards this time.

Here, we combine functional magnetic resonance imaging (fMRI) with a sequence learning task requiring the active, memory-based construction of the times of events forming different sequences. We show that event representations in the anterior hippocampus change through learning to reflect actively constructed event times rather than sequence order or passively elapsing time. Furthermore, the anterior hippocampus generalizes temporal relations across sequences, and structural knowledge about other sequences systematically biases the construction of specific event times. While within- and across-sequence relations are detected in anatomically overlapping

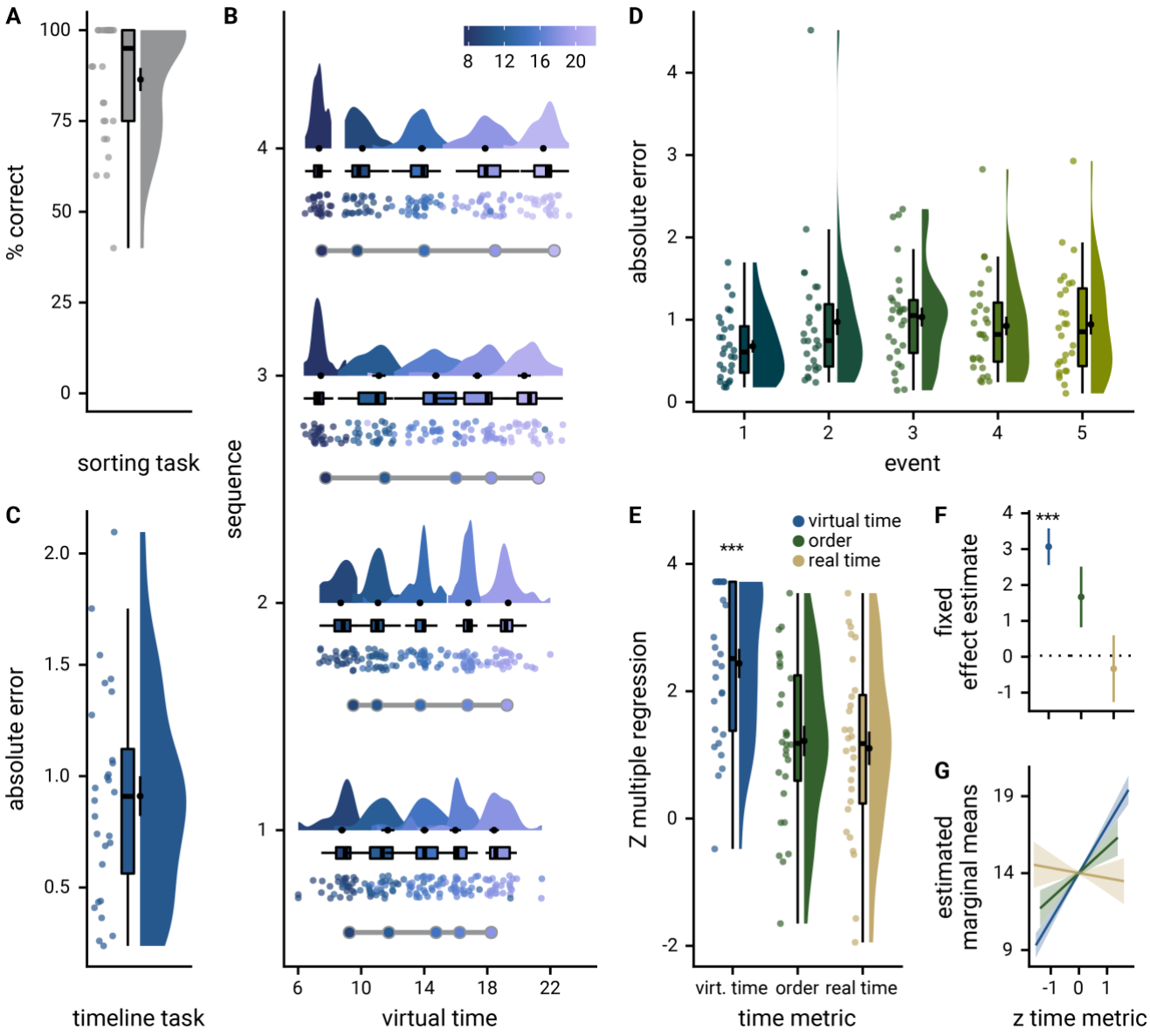


Figure 2. Participants learn the temporal structure of the sequences relative to the virtual clock. A. Plot shows the percentage of correctly sorted event images in the sorting task. **B.** Constructed event times were assessed in the timeline task. Responses are shown separately for each of the five events (color coded according to true virtual time) of each of the four sequences (rows). Colored circles with gray outline at the bottom of each row show true event times. **C, D.** Mean absolute errors in constructed times (in virtual hours) are shown (**C**) averaged across events and sequences and (**D**) averaged separately for the five event positions. **E**. Z-values for the effects of different time metrics from participant-specific multiple regression analyses and permutation tests show that virtual time explained constructed event times with event order and real time in the model as control predictors. **F**. Likewise, parameter estimates and 95% confidence intervals for the fixed effects of the three time metrics from a linear mixed model indicate that virtual time relates to constructed event times beyond the effects of order and real time. **G**. Estimated marginal means (model predictions) illustrate the effects of the three time metrics. **A-E**. Circles are individual participant data; boxplots show median and upper/lower quartile along with whiskers extending to most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bars corresponds to mean±S.E.M.; distributions show probability density function of data points. *** p<0.001

regions of the hippocampus, the mode of representation differs depending on whether events belong to the same sequence or not. In contrast, the anterior-lateral entorhinal cortex uses one shared representational format to map relationships of events from the same and from different sequences.

Results

We asked participants to learn four sequences that consisted of five unique event images each (Figure 1, Supplemental Figure 1). Participants were instructed that each sequence depicted events taking place on a specific day in the life of a family. Their task was to infer the time of each event relative to the temporal reference frame of a virtual clock (Figure 1C). The true virtual times of events were never revealed. Rather, the clock was running hidden from participants. It was uncovered only infrequently between event presentations to briefly show the current virtual time (Supplemental Figure 2, see Methods). Participants had to combine their subjective experience of passing

time with these time cues to construct event times. Importantly, we manipulated the speed of the hidden clock between sequences so that different amounts of virtual time passed in the same real time intervals. With this paradigm, we partially dissociated the virtual time of events from the event order and objectively elapsing time (real time) to test whether actively constructed event times underlie participants' memory for the temporal structure of the sequences.

Successful construction of event times

We assessed memory for the sequences using two behavioral tests administered at the end of the experimental session. First, participants sorted all event images according to sequence membership (Figure 1D). The high performance in this task (Figure 2A; 86.43%±16.82% mean±standard deviation of correct sorts) demonstrates accurate memory for which events belonged to the same sequence. Second, to probe constructed event times, we asked participants to position the events of a sequence on a timeline (Figure 1E). Remembered times were highly accurate (Figure 2B-D; 0.91±0.47 mean±standard deviation of average absolute errors in virtual hours). To test whether these constructed event times were driven by the virtual time of events, we regressed remembered times on virtual times with event order and real time as control predictors of no interest. We did so in a summary statistics approach based on multiple regression for each participant, combined with permutation tests, and using a linear mixed effects model (see Methods). The effect of virtual time on constructed event times was significant when controlling for variance accounted for by event order and real time (Figure 2E-F; summary statistics: t_{27}=10.62, p<0.001; mixed model: $\chi^2(1)$=115.95, p<0.001, Supplemental Table 1). Together, these findings demonstrate that participants formed precise memories of the different sequences and accurately constructed event times.

Hippocampal representations of within-sequence relations reflect constructed event times

Before and after learning the event sequences, participants viewed the event images in random order while undergoing fMRI (Figure 1AB). We quantified changes in the similarity of multi-voxel patterns between pairs of events from before to after learning (Figure 3, see Methods). Using two approaches to model-based representational similarity analysis, we tested whether changes in pattern similarity could be explained by the temporal relationships between pairs of events. Temporal distances between events were measured in virtual time, real elapsing time in seconds and as differences in sequence order position (Figure 1C). In the summary statistics approach, we compared the fit of linear models predicting pattern similarity changes from temporal distances to shuffle distributions for each participant and assessed the resulting Z-values on the group level using permutation-based tests. Second, we fit linear mixed effects models to quantify whether sequence relationships explained pattern similarity changes. Rather than performing inferential statistics on one summary statistic per participant, mixed models estimate fixed effects and their interactions using all data points. We used temporal distance measures as fixed effects while capturing within-participant dependencies with random intercepts and random slopes (see Methods). Based on our previous work (Bellmund et al., 2019; Deuker et al., 2016), we focused our analyses on the anterior hippocampus and the anterior-lateral entorhinal cortex (see Methods).

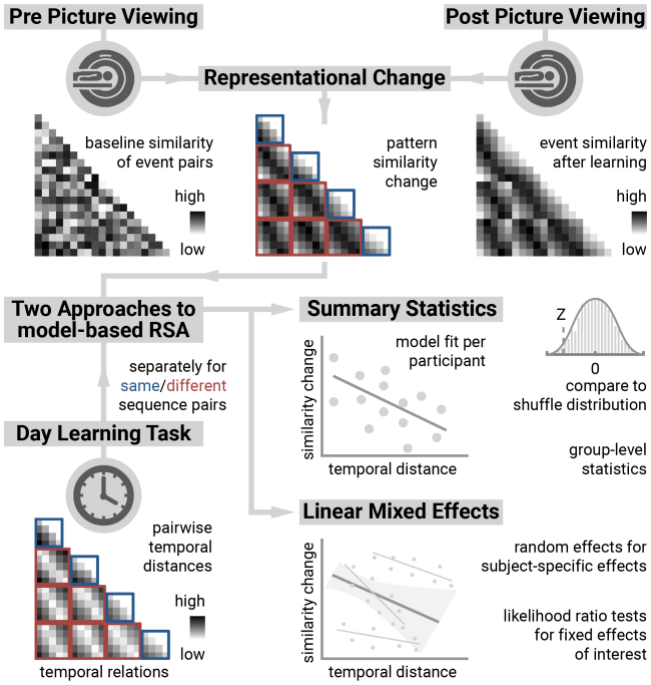


Figure 3. Representational Similarity Analysis Logic. We quantified the representational similarity of all event pairs before and after learning. Representational change was defined by subtracting pre-learning from post-learning pattern similarity (top row). Using two approaches to model-based representational similarity analysis (RSA, see Methods), we analyzed whether pattern similarity changes reflected the temporal structure of the sequences (bottom left). In the summary statistics approach (middle right), we regressed pattern similarity change on temporal distances between events using participant-specific linear models that were compared to null distributions obtained from shuffling similarity change against temporal distances. The resulting Z-values were used for permutation-based group-level statistics. In the mixed model approach (bottom right), we estimated the influence of temporal distances on pattern similarity change using fixed effects, with random effects accounting for within-subject dependencies. The statistical significance of fixed effects was assessed using likelihood ratio tests against reduced models excluding the fixed effect of interest.

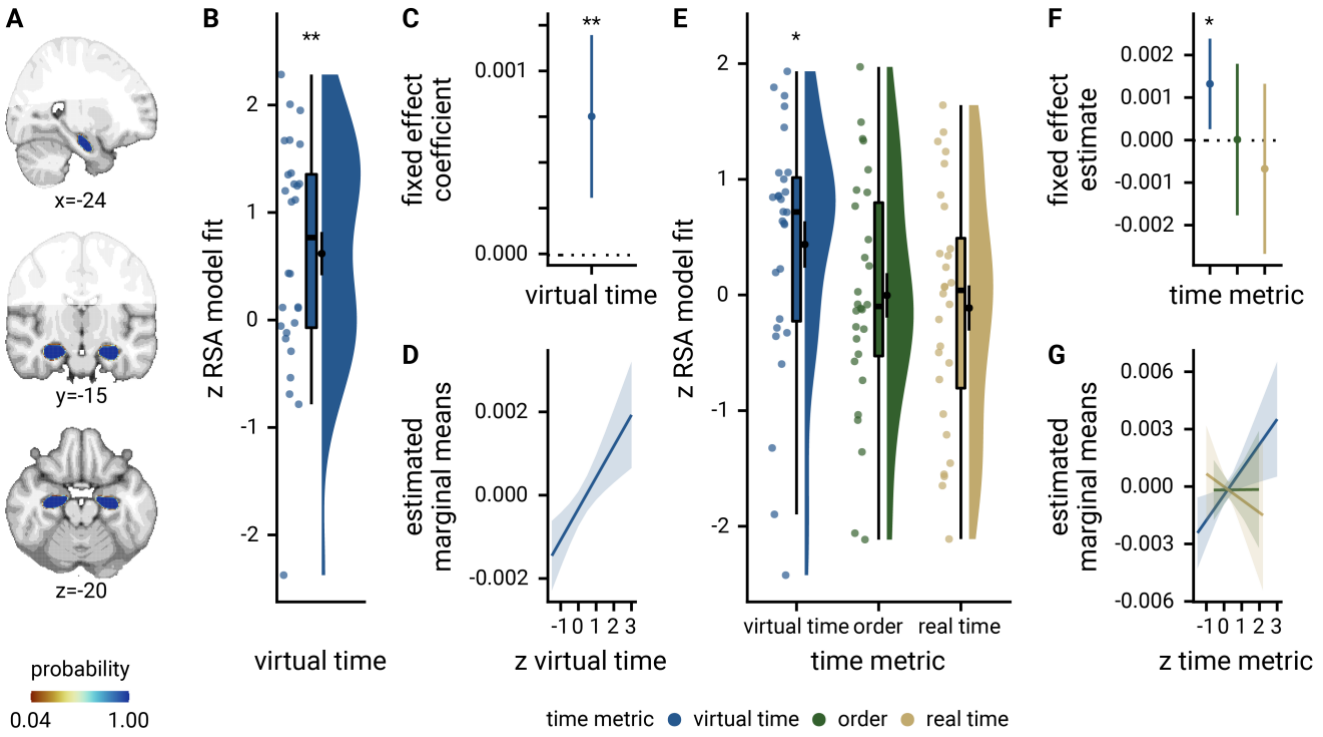


Figure 4. Sequence representations in anterior hippocampus reflect constructed event times. A. The anterior hippocampus region of interest is displayed on the MNI template with voxels outside the field of view shown in lighter shades of gray. Color code denotes probability of a voxel to be included in the mask based on participant-specific ROIs (see Methods). **B**. The Z-values based on permutation tests of participant-specific linear models assessing the effect of virtual time on pattern similarity change for event pairs from the same sequence were significantly positive. **C**. Dot plot shows the parameter estimate and 95% confidence interval for the fixed effect of virtual time on hippocampal pattern similarity change for same-sequence events from a linear mixed effects model. **D**. Estimated marginal means illustrate the positive relationship between virtual temporal distances and pattern similarity change. **E**. Z-values show the relationship of the different time metrics to representational change based on participant-specific multiple regression analyses. Virtual time predicts pattern similarity change with event order and real time in the model as control predictors of no interest. **F**,**G**. Parameter estimates with 95% confidence intervals (**F**) and estimated marginal means (**G**) show the fixed effects of the three time metrics from the corresponding mixed model. **B**,**E**. Circles are individual participant data; boxplots show median and upper/lower quartile along with whiskers extending to most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bars corresponds to mean±S.E.M.; distributions show probability density function of data points. ** p<0.01; * p<0.05

We first tested whether pattern similarity changes in the anterior hippocampus (Figure 4A) could be explained by the virtual temporal distances between event pairs from the same sequence. We observed a positive relationship between similarity changes and temporal distances in both the summary statistics (Figure 4B; t_{27}=3.07, p=0.006; α=0.025, corrected for separate tests of events of the same and different sequences) and the mixed model approach (Figure 4CD; $\chi^2(1)$=9.87, p=0.002, Supplemental Table 2). Thus, pattern similarity was higher for event pairs separated by longer temporal distances than for pairs separated by shorter intervals. A possible explanation for the positive correlation of pattern similarity and temporal distance could be that, in contrast to our previous work (Deuker et al., 2016), participants had to learn multiple sequences. To do so, they might strongly associate the first and last event of each sequence, which mark the transitions between the sequences. Importantly, the effect of virtual temporal distances on hippocampal pattern similarity changes remained significant when competing for variance with a control predictor accounting for these comparisons (Supplemental Figure 3A-C; summary statistics: t_{27}=2.25, p=0.034; mixed model: $\chi^2(1)$=5.36, p=0.021, Supplemental Table 3). Thus, hippocampal event representations changed through learning to reflect temporal distances.

Next, to assess whether this effect was indeed driven by the constructed event times, we included the two additional time metrics as control predictors in the model. Virtual temporal distances significantly predicted hippocampal pattern similarity changes even when controlling for the effects of event order and real time in seconds (Figure 4E-F; summary statistics: t_{27}=2.18, p=0.040, mixed model: $\chi^2(1)$=5.92, p=0.015, Supplemental Table 4). We further observed that the residuals of linear models, in which hippocampal representational change was predicted from order and real time, were related to virtual temporal distances (Supplemental Figure 3D; t_{27}=2.23, p=0.034), demonstrating that virtual time accounts for variance that the other time metrics fail to explain. Together, these data show that hippocampal representations of events from the same sequence changed to reflect actively constructed event times.

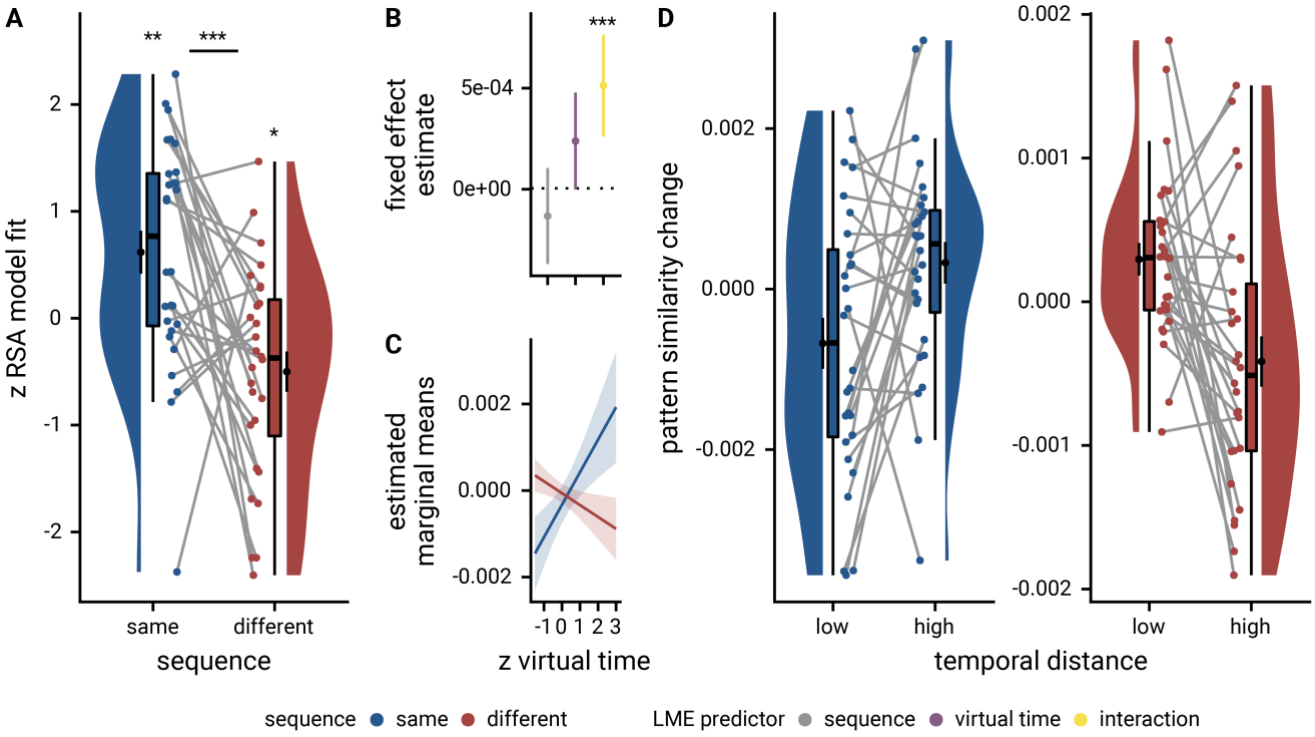


Figure 5. The anterior hippocampus generalizes temporal relations across sequences. A. Z-values show results of participant-specific linear models quantifying the effect of virtual time for event pairs from the same sequence (blue, as in Figure 4B) and from different sequences (red). Temporal distance is negatively related to hippocampal representational change for event pairs from different sequences. See Supplemental Figure 4 for mixed model analysis of across-sequence comparisons. The effect of virtual time differs for comparisons within the same sequence or between two different sequences. **B**. In the corresponding mixed effect model, a significant interaction between virtual time and sequence membership was observed. Dot plot shows fixed effect estimates with 95% confidence intervals. **C**. Estimated marginal means illustrate the diverging effects of virtual time as a function of sequence membership of event pairs. **D**. To visualize the interaction effect in **A-C**, raw pattern similarity change was averaged for events separated by low and high temporal distances, respectively. *** p≤0.001; ** p<0.01; * p<0.05

The hippocampus generalizes temporal relations across sequences

We next tested whether similarity changes of hippocampal representations of events from different sequences mirrored generalized temporal distances. When comparing pairs of events belonging to different sequences, we observed a significant negative effect of virtual temporal distances on pattern similarity change (Figure 5A, Supplemental Figure 4; summary statistics t_{27}=-2.65, p=0.013; mixed model: $\chi^2(1)$=6.01, p=0.014, Supplemental Table 5; α=0.025, corrected for separate tests of events of the same and different sequences). This indicates that the hippocampus generalized temporal relations across sequences, resulting in systematic changes of representations of events from different sequences. Events occurring at similar times relative to the virtual clock, but in different sequences, were represented more similarly than those taking place at more different virtual times.

How temporal relations were reflected in representational change depended on whether events were from the same sequence or not (Figure 5A-C, summary statistics: paired t-test t_{27}=3.71, p=0.001, mixed model: interaction of sequence membership with virtual time $\chi^2(1)$=14.37, p<0.001, Supplemental Table 6). The positive correlation of temporal distance and pattern similarity change implies relatively higher similarity for events that are far apart than those that are close in time, whereas events close together in time became more similar relative to those separated by larger temporal distances when comparing across sequences (Figure 5D). Thus, the way relational knowledge was represented depended on whether events belonged to the same sequence or not.

Sequence representations differ between hippocampus and entorhinal cortex

In our second region of interest, the anterior-lateral entorhinal cortex (Figure 6A), the effect of virtual time on representational change did not differ statistically between event pairs from the same or from different sequences (summary statistics: paired t-test t_{27}=0.07, p=0.942). We thus collapsed across comparisons from the same and different sequences and observed a significant effect of virtual temporal distances on entorhinal pattern similarity change (Figure 6B-D; summary statistics: t_{27}=-2.31, p=0.029; mixed model: $\chi^2(1)$=4.39, p=0.036, Supplemental Table 7; see Supplemental Figure 5 for separate analyses of events from the same and from different sequences). In line with our previous work (Bellmund et al., 2019), events close together in time became more similar than those separated by longer temporal intervals (Figure 6E). We further corroborated that the temporal structure of the sequences was represented differently between the

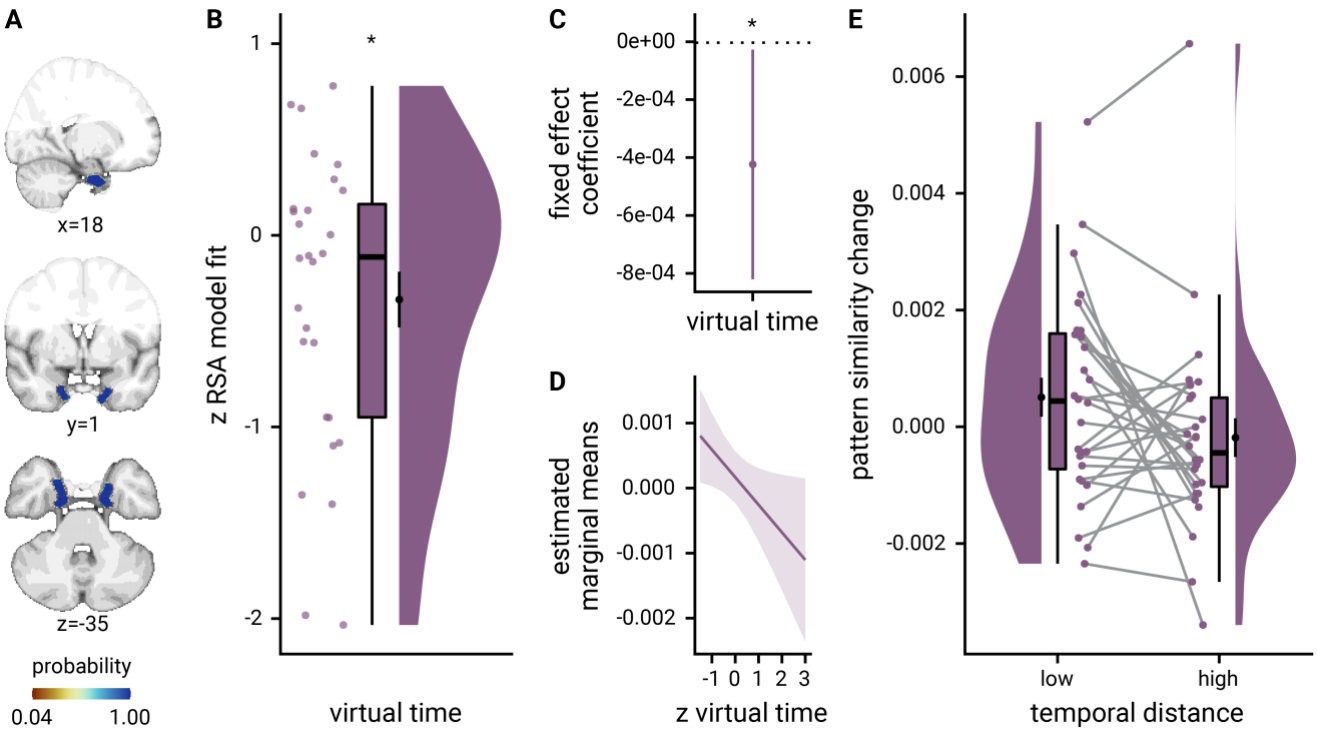


Figure 6. The anterior-lateral entorhinal cortex uses a shared representational format for relations of events from the same and different sequences. **A.** The anterior-lateral entorhinal cortex region of interest is displayed on the MNI template with voxels outside the field of view shown in lighter shades of gray. Color code denotes probability of a voxel to be included based on participant-specific masks (see Methods). **B.** Z-values for participant-specific RSA model fits show a negative relationship between pattern similarity change and virtual temporal distances when collapsing across all event pairs. **C,D**. Parameter estimate with 95% confidence intervals (**C**) and estimated marginal means (**D**) show the fixed effect of virtual time from the corresponding linear mixed effect model. **E.** To illustrate the effect in **B-D**, raw pattern similarity change in the anterior-lateral entorhinal cortex was averaged for events separated by low and high temporal distances. * p<0.05

anterior-lateral entorhinal cortex and the anterior hippocampus (summary statistics: interaction between region and sequence membership in permutation-based repeated-measures ANOVA $F_{1,27}$=7.76, p=0.010, main effect of region $F_{1,27}$=3.10, p=0.086, main effect of sequence $F_{1,27}$=7.41, p=0.012; mixed model: three-way interaction between virtual time, sequence membership and region of interest $\chi^2(1)$=6.31, p=0.012, Supplemental Table 8; see Supplemental Figure 6 for a comparison of the signal-to-noise ratio in these regions). Whereas the hippocampus employed two distinct representational formats for temporal relations depending on whether events belonged to the same sequence or not, we observed consistent negative correlations between representational change and temporal distances when collapsing across all event pairs, but no statistically significant difference between representations of temporal relations from the same or different sequences in the entorhinal cortex.

Anatomical overlap between representations of within-sequence relations and across-sequence generalization

We next asked whether representations of same-sequence relations are distinct from or overlap with the across-sequence generalization of temporal relations. For this purpose and to complement our region-of-interest analyses described above, we performed a searchlight analysis that revealed significant effects of virtual temporal distances on representations of events from the same sequence in the bilateral anterior hippocampus (Figure 7A; peak voxel MNI x=-24, y=-13, z=-20; t=4.53, p_{svc}=0.006, Supplemental Table 9). We used the same-sequence searchlight peak cluster to define a region of interest to test for the independent across-sequence generalization effect (see Methods). Indeed, virtual temporal distances explained pattern similarity change for events from different sequences in these voxels (Figure 7B-D; summary statistics t_{27}=-2.19, p=0.036; mixed model: $\chi^2(1)$=4.13, p=0.042, Supplemental Table 10), demonstrating an overlap between representations of within-sequence relations and their generalization across sequences. Further, we conducted a searchlight analysis looking for negative correlations of temporal distances and pattern similarity change for events from different sequences. We detected clusters in anterior hippocampus that overlapped with the same-sequence searchlight effect (Figure 7EF), though this searchlight generalization effect did not survive corrections for multiple comparisons (peak voxel MNI x=-26, y=-19, z=-15, t=-3.96, p_{svc}=0.071, Supplemental Table 11). Lastly, we directly searched for brain areas in which pattern similarity change differentially scaled with temporal distances depending on whether events were from the same or different sequences. The two largest clusters in our field of view were located in the left and right

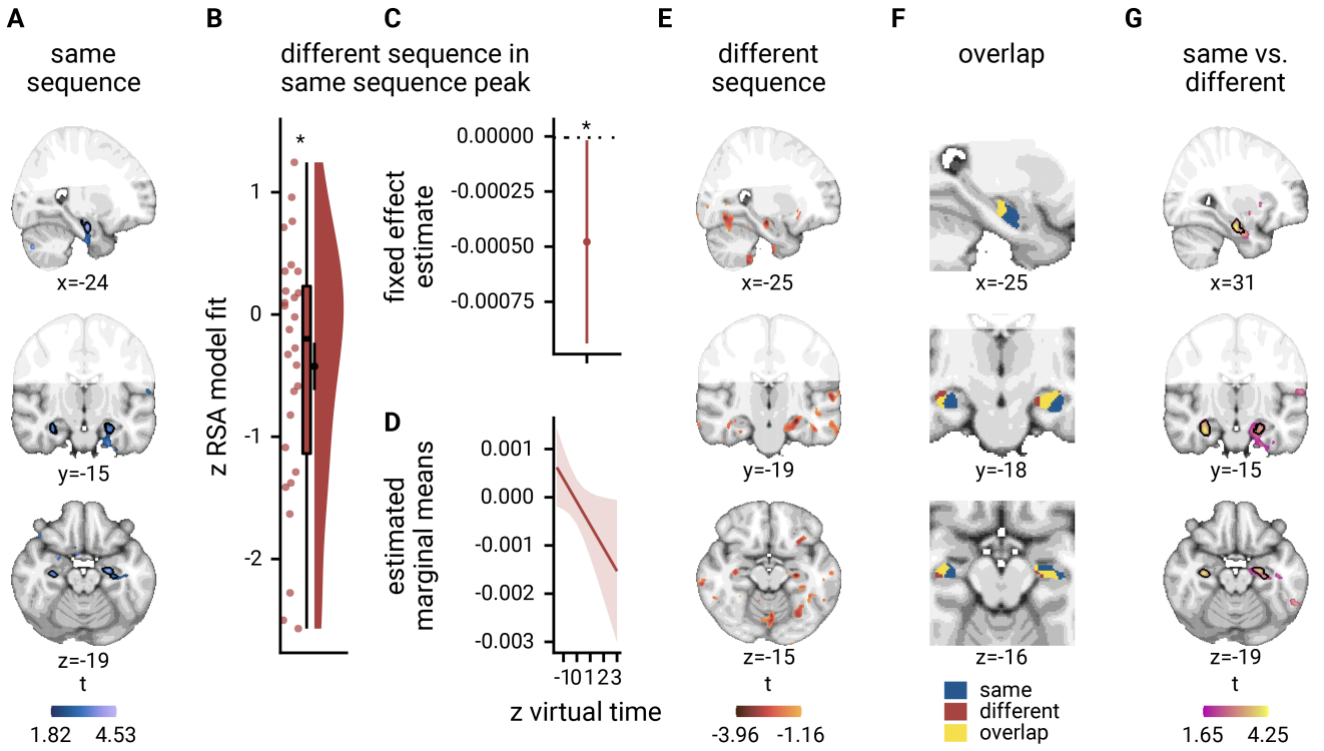


Figure 7. Overlapping representations of within- and across-sequence relations. A. Searchlight analysis results show a positive relationship between representational change and virtual temporal distances for event pairs from the same sequence in the bilateral anterior hippocampus. Statistical image is thresholded at $p_{uncorrected}$<0.01; voxels within black outline are significant after correction for multiple comparisons using small volume correction. **B-D**. In the peak cluster from the independent within-sequence searchlight analysis (**A**), representational change was negatively related to virtual temporal distances between events from different sequences. **B**. Circles show individual participant Z-values from summary statistics approach; boxplot shows median and upper/lower quartile along with whiskers extending to most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bars corresponds to mean±S.E.M.; distribution shows probability density function of data points. **C.** Dot plot shows fixed effect estimate with 95% confidence interval from linear mixed model. **D**. The estimated marginal means from the linear mixed effect model illustrate the negative relationship between virtual temporal distances and representational change. **E**. Searchlight analysis results show negative relationship between representational change and temporal distances for different-sequence event pairs. Statistical image is thresholded at $p_{uncorrected}$ <0.05. **F**. Within the anterior hippocampus, the effects for events from the same sequence and from two different sequences overlap. Visualization is based on statistical images thresholded at $p_{uncorrected}$ <0.05 within small volume correction mask. **G**. Searchlight analysis results show a bilateral interaction effect in the anterior hippocampus that is defined by a differential relationship of virtual temporal distances and representational change for events from the same and different sequences. Statistical image is thresholded at $p_{uncorrected}$<0.01; voxels within black outline are significant after correction for multiple comparisons using small volume correction. **A**, **E**, **G**. Results are shown on the MNI template with voxels outside the field of view displayed in lighter shades of gray. See Supplemental Figure 7 for additional exploratory results. * p<0.05

anterior hippocampus (peak voxel MNI x=31, y=-16, z=-21; t=4.25, p_{svc}=0.007, Supplemental Table 12). Taken together, these findings highlight that hippocampal representations carry information about the specific sequence in which events occur, and that these temporal relations are generalized across similar sequences.

Generalized knowledge about other sequences biases event time construction

Having established generalized hippocampal event representations, we explored whether knowledge about the general structure of event times in other sequences influenced the construction of individual event times. For each event, we quantified when it took place relative to the average virtual time of the events at the same sequence position in the other three sequences (Figure 8A; see Methods). We reasoned that the construction of a specific event time could be biased by knowledge about the general pattern of event times at that sequence position. Indeed, we observed positive relationships between the relative time of other events and signed errors in constructed event times as assessed in the timeline task (Figure 8B-E, Supplemental Figure 8A; summary statistics: t_{27}=5.32, p<0.001; mixed model: $\chi^2(1)$=17.90, p<0.001, Supplemental Table 13). This demonstrates that general knowledge about the sequences biased the construction of event times. When events at the same sequence position, but in different sequences, took place late relative to an event, the constructed virtual time for that event tended to be overestimated, and vice versa when the other events occurred relatively early. In an independent group of participants (Montijn et al., 2021), we replicated this generalization bias (Figure 8F-H, Supplemental Figure 8B; summary

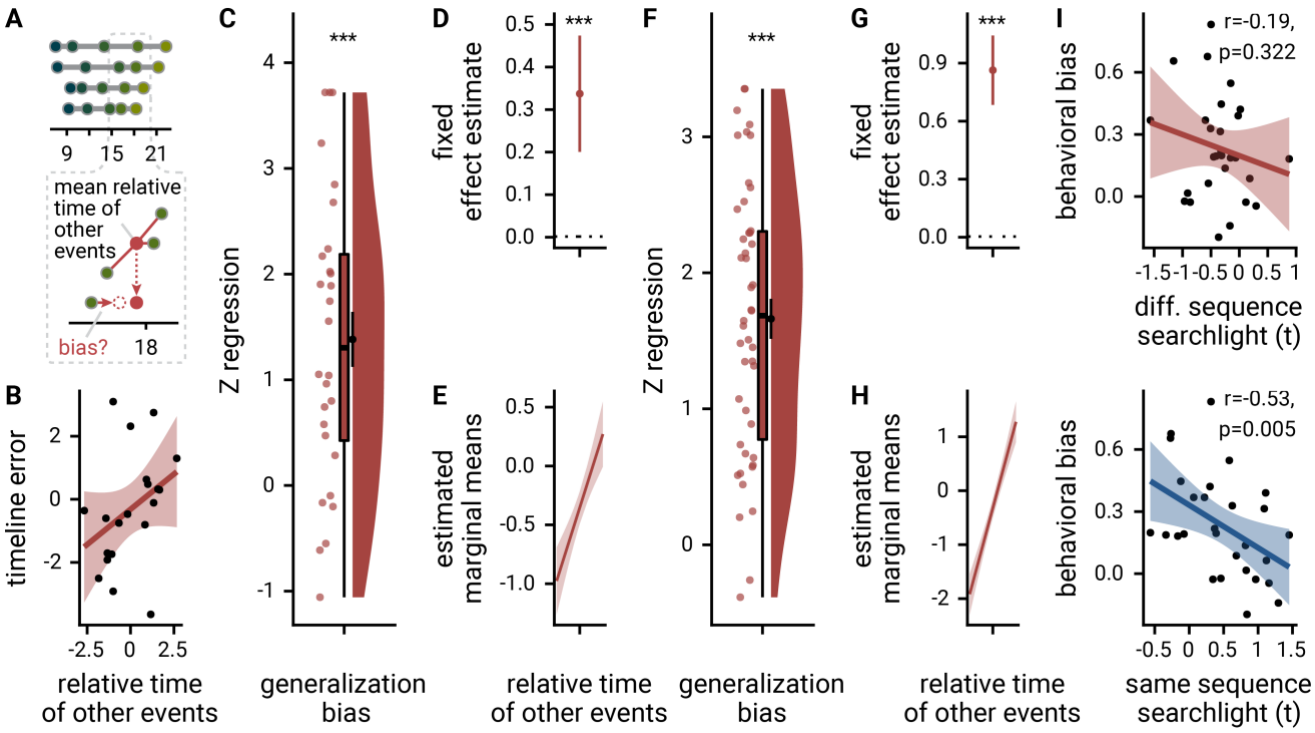


Figure 8. Structural knowledge biases construction of event times. A. The generalization bias quantifies the influence of structural knowledge on the construction of individual event times. For each event, the mean time of events at the same sequence position in the other sequences was calculated to test whether event times were biased towards the relative time of other events. **B.** The scatterplot illustrates the generalization bias for an example participant. Each circle corresponds to one event and the regression line highlights the relationship between the relative time of other events and the errors in constructed event times. The example participant was chosen to have a median-strength generalization bias. See Supplemental Figure 8 for the entire sample. C-E. The relative time of events from other sequences predicted signed event time construction errors as measured in the timeline task. **C.** Circles show individual participant Z-values from participant-specific linear models (**B**); boxplot shows median and upper/lower quartile along with whiskers extending to most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bars corresponds to mean±S.E.M.; distribution shows probability density function of data points. **D.** Dot plot shows fixed effect estimate with 95% confidence interval from linear mixed model. **E**. The estimated marginal means from the linear mixed effect model illustrate the positive relationship between the time of other events and constructed event times. When other events took place late relative to a specific event, the time of that event was estimated to be later than when other events were relatively early. **F-H**. The generalization bias in event time construction through structural knowledge was replicated in an independent sample (n=46) based on Montijn et al. (2021). Data shown as in **C-E**. **I.** The behavioral generalization bias (regression coefficients from summary statistics approach) did not correlate significantly with the across-sequence generalization effect in the anterior hippocampus (searchlight peak voxel t-values). **J.** We observed a significant negative correlation between the same-sequence searchlight effect (peak voxel t-values) and the behavioral generalization bias (regression coefficients from summary statistics approach), suggesting that participants with strong hippocampal representations of the temporal relations between events from the same sequence were less biased by structural knowledge in their construction of event times. Statistics in **I** and **J** are based on Spearman correlation.

statistics: t_{45}=11.30, p<0.001 mixed model: $\chi^2(1)$= 53.74, p<0.001, Supplemental Table 14), confirming the influence of generalized knowledge about the sequences on event time construction. One possibility is that structural knowledge about the sequences biases the construction of specific event times, in particular when uncertainty about the virtual time of events is high. Indeed, we observed a significant negative correlation between how strongly pattern similarity changes in the anterior hippocampus reflected temporal relations between events from the same sequence in the searchlight analysis (Figure 8IJ, Spearman r=-0.53, p=0.005; α=0.025 corrected for two comparisons; correlation with across-sequence effect: Spearman r=-0.19, p=0.322), suggesting that the construction of event times was less biased by time patterns generalized across sequences in those participants with precise representations of within-sequence temporal relations.

Discussion

Our findings show that hippocampal event representations change through learning to reflect temporal relations based on mnemonically constructed event times. Converging region of interest and searchlight analyses demonstrate that, on the one hand, the hippocampus forms highly specific representations of temporal relations of the events in a sequence that mirror actively constructed event times beyond the effects of order and real time. On the other hand, temporal relations are generalized across sequences using a different representational format. In contrast, the similarity of event representations in the

anterior-lateral entorhinal cortex scaled with temporal distances for events irrespective of sequence membership. The behavioral data demonstrate that the construction of specific event times is biased by structural knowledge abstracted from different sequences.

The finding of two alternative representational formats for temporal relations of events from the same or different sequences possibly reconciles memory specificity and generalization in the hippocampus. The hippocampus has long been implicated in memory for specific episodes. It has been suggested to differentiate similar episodes (Chanales et al., 2017; Favila et al., 2016; Lohnas et al., 2018; Milivojevic et al., 2015; Schlichting et al., 2015; Zeithamova et al., 2018), to form associations between related elements (Eichenbaum and Cohen, 1988; Cohen and Eichenbaum, 1993), and to bind events to spatiotemporal contexts (Diana et al., 2007). Our data show that the similarity of event representations in the anterior hippocampus scales with temporal distances between events and thus carries information about the specific sequence in which these events took place. Conversely, there is evidence that the hippocampus contributes to flexible cognition via the recombination of different experiences and statistical learning (Morton et al., 2020; Preston et al., 2004; Schapiro et al., 2012, 2017; Shohamy and Wagner, 2008; Zeithamova and Bowman, 2020; Zeithamova and Preston, 2010; Doeller et al., 2005, 2006). It enables inference and generalization by combining elements across episodes. In our paradigm, we observed that temporal relations were generalized across sequences, resulting in systematic changes of representations of events from different sequences. This finding is in line with the generalization of learned event dynamics to situations with different surface features, but similar relational structure, in a computational model of event memory (Franklin et al., 2020). Our findings further demonstrate that specific and generalized representations overlap anatomically in the anterior hippocampus. The way temporal relations shaped hippocampal multi-voxel pattern similarity differed between pairs of events from the same and different sequences. This might enable the anterior hippocampus to combine fine-grained knowledge about the temporal structure of specific sequences with the generalization across sequences, in line with its contributions to relational binding, memory specificity, and inferential reasoning.

Learning the temporal relations of events in a sequence entails acquiring knowledge about the sequence structure. In navigation, cognitive maps of spatial relations enable flexible behavior (Tolman, 1948). Foundational neuroscientific discoveries uncovered spatially tuned cells in the hippocampal-entorhinal region as potential neural substrates for cognitive maps of space (Moser et al., 2017; O'Keefe and Nadel, 1978). Consistent with its involvement in various cognitive domains, the hippocampal-entorhinal region was suggested to form cognitive maps of relational knowledge more generally (Eichenbaum and Cohen, 1988; O'Keefe and Nadel, 1978). The characteristic activity patterns of spatially tuned cells also encode non-spatial dimensions, suggesting the organization of relational knowledge in cognitive spaces (Bellmund et al., 2018). In the temporal domain, neurons in the rodent hippocampus fire at specific time points during a delay, resulting in characteristic sequences of activity (Kraus et al., 2013; MacDonald et al., 2011; Pastalkova et al., 2008). These neurons are often referred to as time cells and have been suggested to support temporal memory (Eichenbaum, 2014). Recent work characterized similar activity patterns also in the human hippocampus and entorhinal cortex during the encoding and retrieval of word lists (Umbach et al., 2020). In concert with the link between the similarity of event representations and temporal relations between these events (Bellmund et al., 2019; Deuker et al., 2016; Nielson et al., 2015), these findings raise the possibility that the hippocampal-entorhinal region forms cognitive maps reflecting the temporal relational structure of events in a sequence. Knowledge of relational structures enables inference and generalization (Behrens et al., 2018; Park et al., 2020; Whittington et al., 2020). Importantly, our findings demonstrate that fine-grained relational knowledge is stored in generalized event representations.

Structural knowledge influences mnemonic construction. In two independent samples, we show that general time patterns, abstracted from other sequences, bias the construction of specific event times. When events at the same sequence position, but in other sequences, took place relatively late to the time of an event, the time of that event was remembered to be later than when the other events occurred relatively early. This generalization bias shows that knowledge about events at structurally similar positions contributes to constructive memory for specific events. It is in line with biases resulting from the exploitation of environmental statistics when reconstructing stimulus sizes from memory (Hemmer and Steyvers, 2009; Hemmer et al., 2015), when estimating brief time intervals (Jazayeri and Shadlen, 2010; Polti et al., 2021), or when discriminating the order of previously presented stimuli (Orlov et al., 2000). Likewise, prior knowledge can distort memories for short narratives (Bower et al., 1979), spatial associations (Tompary and Thompson-Schill, 2021) and temporal positions (Frisoni et al., 2021). Consistent with the suggested role of grid cells in the representation of spatial structure, distortions in mnemonic reconstructions of spatial relations induced through boundary geometry follow predictions from models of grid-cell functioning (Bellmund et al., 2020b). Further, recombining information across episodes for associative inference can induce false memories for contextual details (Carpenter and

Schacter, 2017), illustrating that generalization impacts memory for specific associations. In line with the greater reproduction of episodic details by participants whose recall follows the temporal structure of an experience more closely (Diamond and Levine, 2020), these findings highlight that structural knowledge and mnemonic construction are intertwined. More broadly, abstract semantic or schematic knowledge may provide a scaffold for the recall of episodic details (Greenberg and Verfaellie, 2010; Irish and Piguet, 2013; Schacter et al., 2017; Addis, 2020). However, our findings show that structural knowledge not only facilitates but can also bias constructive memory.

The hippocampus supports constructive memory and generalization in concert with a distributed network of brain regions. In addition to medial temporal lobe structures, the mental simulation of past and future episodic scenarios recruits a core network including medial prefrontal and retrosplenial cortex as well as lateral parietal and temporal areas (Benoit and Schacter, 2015; Schacter et al., 2007). Notably, this network overlaps with areas supporting the recombination of elements and generalization. For example, both the construction of novel experiences based on the combination of multiple elements (Barron et al., 2013) and memory integration across episodes (Schlichting et al., 2015) are supported by the medial prefrontal cortex and the hippocampus. In sequence processing, representational similarity is increased for items occupying the same position in different sequences in parahippocampal, retrosplenial and medial prefrontal cortices as well as in the angular gyrus (Hsieh et al., 2014; Hsieh and Ranganath, 2015). Likewise, sequence positions can be decoded from magnetoencephalographic responses elicited by visual stimuli presented in scrambled order (Liu et al., 2019). In line with the suggestion that the posterior parietal cortex supports generalization by projecting stimuli onto a low-dimensional manifold (Summerfield et al., 2019), neural magnitude representations that generalize across task contexts have been observed using EEG (Luyckx et al., 2019; Sheahan et al., 2021). While we did not observe effects outside the hippocampal-entorhinal region that survived corrections for multiple comparisons, we note that, based on our prior hypotheses, we opted for high-resolution coverage of the medial temporal lobe at the cost of reducing the field of view of our MR images. As the events in our task can be conceived of as being arranged along one or multiple, parallel mental number lines, future research could test how the parietal cortex encodes event relations to explore commonalities with and differences to the generalization of event times observed in the hippocampus.

Our paradigm allows a highly-controlled read-out of representational change relative to a pre-learning baseline scan. Events are shown in the same random order before and after learning, ruling out that prior associations or the temporal auto-correlation of the blood-oxygen-level-dependent signal drives our effects. Future studies could extend the paradigm to investigate how hierarchically nested sequences are represented, for example by introducing higher-order relations between sequences – akin to different days being grouped in weeks. The precise temporal dynamics of the generalized hippocampal event representation pose another intriguing question. Based on the report that the temporal organization of memory reactivation relative to the hippocampal theta phase reflects semantic relations between items (Estefan et al., 2021), a speculative hypothesis is that a theta phase code could also underlie memory for temporal relations of events from the same and different sequences.

In conclusion, our findings show that the similarity of event representations in the hippocampus reflects relations between events that go back to mnemonically constructed event times, highlighting the impact of active mnemonic construction on sequence memory beyond the effects of event order and real elapsing time. Temporal relations are generalized to events from different sequences, in line with hippocampal contributions to both memory specificity and generalization across episodes. General time patterns abstracted from other sequences systematically influence the construction of specific event times, demonstrating that constructive memory for specific events builds on structural knowledge.

Methods

Participants

31 participants were recruited for this experiment. Participants gave written informed consent prior to participation. All proceedings were approved by the local ethics committee (CMO Regio Arnhem-Nijmegen). One participant aborted the experiment due to feeling claustrophobic when entering the MR scanner. Two participants were excluded from further analysis due to bad memory performance and technical difficulties during data acquisition. Thus, the sample consisted of 28 participants (21 female, age: mean±standard deviation 23.04±3.21 years, range 18-31 years).

Procedure

Overview

The experiment consisted of four parts (Figure 1A) and lasted approximately 2.5 hours in total. The first three parts were performed inside the MR scanner and comprised a learning task lasting around 50 minutes that was completed in between two blocks of a picture viewing task of around 25 minutes each. The tasks inside the scanner were presented on a rear-projection screen with a resolution of 800x600 pixels and implemented using Presentation (version 16.2, Neurobehavioral Systems). Subsequently, outside of the scanner, participants performed two short memory tasks in front of a computer screen, implemented with custom Matlab code. The tasks are described in more detail below. Data analysis was carried out using FSL (version 5.0.4) (Smith et al., 2004) and R (version 3.6.1) (R Core Team, 2020).

Stimuli

The stimuli (Supplemental Figure 1) used throughout the experiment were created within the life-simulation computer game The Sims 3 (Electronic Arts) by taking screenshots. Each image featured a scene in the life of an affluent family. The main character, the family father, was visible in all scenes. In addition, the mother, son, daughter and family dog appeared in some of the images. All of the depicted events took place within the same family home, but showed activities in a number of different rooms. In an effort to design stimuli with minimal to no indication of day time, the house had constant artificial lighting, but no windows or clocks. The 21 pictures used in this study were selected from an initial set of 35 pictures based on an independent sample rating them as the most ambiguous with regard to the time of day they could take place. One image served as a target image for the picture viewing tasks (see below), while the other 20 event images were randomly assigned to different times and days for every participant.

Picture Viewing Tasks

In the picture viewing tasks (Figure 1B), participants viewed a stream of the event images. Their task was to look at the images attentively and to respond via button press whenever a target picture, which showed the father feeding the family's dog, was presented (pre-learning: 95.71%±7.90% mean±standard deviation of percentage of hits; 881.34ms±131.43ms mean±standard deviation of average reaction times; post-learning: 95.71%±6.90% mean±standard deviation of percentage of hits; 841.40ms±162.16ms mean±standard deviation of average reaction times). The task consisted of 10 mini-blocks. In each mini-block, the target image and the 20 images, which would later make up the virtual days (see Day Learning Task), were shown in random order. Mini-blocks were separated by breaks of 15 s. Stimulus presentations lasted 2.5 s and were time-locked to fMRI volume acquisition onsets. Scene stimuli within a mini-block were separated by 2 or 3 repetition times (TR), randomly assigned so that both stimulus onset asynchronies occurred equally often.

For each participant, we generated a random stimulus order with the constraint that no scene was consistently presented at early or late positions across mini-blocks. Specifically, we compared sequence positions across mini-blocks between the images using a one-way ANOVA. We discarded randomizations where this ANOVA was statistically significant to exclude biases in presentation order. Crucially, the same, participant-specific random order of stimuli and inter-stimulus intervals was used in both the pre-learning and the post-learning picture viewing task. Thus, any systematic differences in the representational similarity of event pairs between the two picture viewing tasks do not go back to differences in the timing of stimulus presentations or the temporal auto-correlation of the BOLD-signal. Rather, we interpret such changes to be a consequence of the learning task.

Day Learning Task

In this task, 20 of the 21 scenes, which were shown in the picture viewing tasks, were presented repeatedly. This time, however, they were grouped into multiple sequences introduced as "virtual days" to participants. There were four different sequences, each comprising 5 events. Events from the same sequence were always shown in a specific order and with a specific time delay between them. Scenes were on screen for 1.5 s. At the end of each sequence, an image of a moon was shown for 5 s, then the next sequence began. Every sequence was presented 7 times. There were 7 mini-blocks in this task. Within each of these, every sequence was presented once. At the end of a mini-block, a 30-s break followed, then the next block started. The order in which the sequences were presented differed randomly across the 7 mini-blocks.

We instructed participants that the scenes depicted events from the life of a family and that the sequences of event images corresponded to different days in the family's life. Participants were asked to memorize which events made up the different sequences (Figure

1C). We further instructed them to learn when during the respective sequence each event occurred. Specifically, we asked participants to learn event times relative to a virtual clock. This clock was running hidden from participants and event images were shown whenever the hidden clock reached the specific event time (Figure 1C, Supplemental Figure 2AB). The task was devised such that participants had to rely on their subjective experience and mnemonic processes to construct the times of events.

To give participants an indication of virtual time, the hidden clock was made visible 6 times for every presentation of a sequence: once before the first event, once in between successive events, and once after the last event. Thus, participants had to focus on their experience of passing time between these time cues to infer the event times. Importantly, the exposure of the hidden clock occurred at random times for each sequence presentation (Supplemental Figure 2CD), with the constraint that it could not be revealed closer than 2 s to a preceding or subsequent event. Thus, participants saw different time cues in each repetition of a sequence. For example, while a specific event always happened at the same virtual time, e.g. 2:07 p.m., the virtual clock could be exposed at any time before the event, e.g. corresponding to 1:32 p.m. in the first repetition of the sequence, and corresponding to 1:17 p.m. in the second repetition. Because true event times were never revealed, participants could not exclusively rely on associative learning to solve the task. Time cues were visible for 1.5 s, but displayed only the time at the start of exposure, i.e. the displayed time did not change within the duration of its presentation.

In short, participants had to combine their experience-based estimates of passing time with the time cues provided by the exposures of the otherwise hidden clock to infer the time at which each event in each sequence took place. Crucially, we varied the speed of the hidden clock between sequences in an effort to partly dissociate real time (in seconds) from virtual time (in virtual hours). Thus, for two sequences more virtual time passed in a comparable amount of real elapsing time (Figure 1C, Supplemental Figure 2). This manipulation allowed us to determine in later analysis whether participants successfully constructed the temporal structure relative to the virtual clock or whether real elapsing time or the order of events determined their memories of the sequences.

Sorting Task

The day sorting task (Figure 1D) was performed in front of a computer screen. The 20 event images from the day learning task were presented on the screen in a miniature version. They were arranged in a circle around a central area displaying 4 rectangles. Participants were instructed to drag and drop all events of the same sequence into the same rectangle with a computer mouse. Participants freely chose which rectangle corresponded to which sequence as the sequences were not identifiable by any label and were presented in differing orders across mini-blocks during learning.

Timeline Task

In this task, participants saw a timeline ranging from 6 a.m. to midnight together with miniature versions of the five event images belonging to one sequence (Figure 1E). Participants were instructed to drag and drop the event images next to the timeline so that scene positions reflected the event times they had inferred in the day learning task. To facilitate precise alignment to the timeline, event images were shown with an outward pointing triangle on their left side, on which participants were instructed to base their responses.

MRI Acquisition

MRI data were recorded with a 3T Siemens Skyra scanner (Siemens, Erlangen, Germany). A high-resolution 2D EPI sequence was used for functional scanning (TR=2270 ms, TE=24 ms, 40 slices, distance factor 13%, flip angle 85°, field of view (FOV) 210x210x68 mm, voxel size 1.5 mm isotropic). The field of view (FOV) was aligned to fully cover the medial temporal lobe, parts of ventral frontal cortex and (if possible) calcarine sulcus. Functional images for the two picture viewing tasks and the learning task were acquired in three runs. In addition to these partial-volume acquisitions, 10 scans of a functional whole-brain sequence were also acquired to improve registration during preprocessing. The sequence settings were identical to the functional sequence above, but instead of 40 slices, 120 slices were acquired, leading to a longer TR (6804.1ms). A structural scan was acquired for each participant (TR = 2300 ms; TE = 315 ms; flip angle = 8°; in-plane resolution = 256x256 mm; number of slices = 224, voxel resolution = 0.8x0.8x0.8 mm). Lastly, a gradient field map was acquired (for n = 21 participants only due to time constraints), with a gradient echo sequence (TR = 1020 ms; TE1 = 10 ms; TE2 = 12.46 ms; flip angle = 90°; volume resolution = 3.5x3.5x2 mm; FOV = 224x224 mm).

ROI Definition

Following our previous work on sequence representations in the hippocampus and entorhinal cortex (Bellmund et al., 2019; Deuker et al., 2016), we focused our analysis on the anterior hippocampus and the anterior-lateral entorhinal cortex. Region of interest (ROI) masks were based on participant-specific FreeSurfer segmentations (version 6.0.0-2), which yielded masks for the entire hippocampus and entorhinal cortex. These were co-registered to participants' functional space. We defined anterior hippocampus using the Harvard-Oxford atlas mask (thresholded at 50% probability), selecting all voxels anterior to MNI y=-21 based on Poppenk et al. (2013). The resulting anterior hippocampus mask was also co-

registered to participants' functional space and intersected with the participant-specific hippocampal mask from FreeSurfer. The mask for the anterior-lateral entorhinal cortex was based on Navarro Schröder et al. (2015). It was co-registered to participants' functional space and intersected with the entorhinal cortex mask from FreeSurfer.

Data Analysis

Behavioral Data Analysis

Sorting Task

For analysis of the sorting task, we took the grouping of event images as provided by the participants and assigned them to the four sequences to ensure maximal overlap between actual and sorted sequence memberships. While the assignment of groupings to sequences is unambiguous when performance is, as in our sample, high, this procedure is potentially liberal at lower performance levels. We then calculated the percentage of correctly sorted event images for each participant, see the raincloud plot (Allen et al., 2019) in Figure 2A.

Timeline Task

We analyzed how well participants constructed the event times based on the day learning task. We quantified absolute errors across all events (Figure 2C) as well as separately for the five sequence positions (Figure 2D). Further, using two approaches we tested whether virtual time drove participants' responses rather than the sequence order or objectively elapsing time. For the summary statistics approach, we ran a multiple regression analysis for each participant with virtual time, sequence position (order), and real time since the first event of a day as predictors of responses in the timeline task. To test whether virtual time indeed explained participants' responses even when competing for variance with order and real time, included in the model as control predictors of no interest, we compared the participant-specific t-values of the resulting regression coefficients against null distributions obtained from shuffling the remembered times against the predictors 10,000 times. We converted the resulting p-values to Z-values and tested these against zero using a permutation-based t-test (10,000 random sign-flips, Figure 2E).

Second, we addressed this question using linear mixed effects modeling. Here, we included the three z-scored time metrics as fixed effects. Starting from a maximal random effect structure (Barr et al., 2013), we simplified the random effects structure to avoid convergence failures and singular fits. The final model included random intercepts and random slopes for virtual time for participants. The model results are visualized by dot plots showing the fixed effect parameters with their 95% confidence intervals (Figure 2F) and marginal effects (Figure 2G) estimated using the ggeffects package (Lüdecke, 2018). To assess the statistical significance of virtual time above and beyond the effects of order and real time, we compared this full model to a nested model without the fixed effect of virtual time, but including order and real time, using a likelihood ratio test. Supplemental Table 1 provides an overview of the final model and the model comparison.

To explore whether structural knowledge about general time patterns biases the construction of event times, we assessed errors in remembered event times. Specifically, when constructing the time of one specific event, participants could be biased in their response by the times of the events from other sequences at that sequence position. For each event, we quantified the average time of events in the other sequences at the same sequence position (Figure 8A). For example, for the fourth event of the first sequence, we calculated the average time of the fourth events of sequences two, three and four. We then asked whether the deviation between the average time of other events and an event's true virtual time was systematically related to signed errors in constructed event times. A positive relationship between the relative time of other events and time construction errors indicates that, when other events at the same sequence position are relatively late, participants are biased to construct a later time for a given event than when the other events took place relatively early. In the summary statistics approach, we ran a linear regression for each participant (Figure 8B, Supplemental Figure 8) and tested the resulting coefficients for statistical significance using the permutation-based procedures described above (Figure 8C). The regression coefficients from this approach were used to test for a relationship between the behavioral generalization bias and the hippocampal searchlight effects (see below). Further, we analyzed these data using the linear mixed model approach (Figure 8DE, Supplemental Table 13).

To replicate the results from this exploratory analysis, we conducted the same analysis in an independent group of participants. These participants (n=46) constituted the control groups of a behavioral experiment testing the effect of stress induction on temporal memory (Montijn et al., 2021). They underwent the same learning task as described above with the only difference being the duration of this learning phase (4 rather than 7 mini-blocks of training). The timeline task was administered on the day after learning. The procedures are described in detail in Montijn et al. (2021). The data from this independent sample are shown in Figure 8F-H and Supplemental Figure 8B.

MRI Preprocessing

Preprocessing was performed using FSL FEAT (version 6.00). Functional scans from the picture viewing tasks and the whole-brain functional scan were submitted to motion correction and high-pass filtering using FSL FEAT. For the two picture viewing tasks, data from each mini-block was preprocessed independently. For those participants with a field map scan, distortion

correction was applied to the functional data sets. No spatial smoothing was performed. Functional images from the two picture viewing tasks were then registered to the preprocessed mean image of the whole-brain functional scan. The whole-brain functional images were registered to the individual structural scans. The structural scans were in turn normalized to the MNI template (1-mm resolution). Gray matter segmentation was done on the structural images, and the results were mapped back to the space of the whole-brain functional scan for later use in the analysis.

Representational Similarity Analysis

Representational similarity analysis (RSA) (Kriegeskorte et al., 2008) was first implemented separately for the pre- and post-learning picture viewing task. It was carried out in ROIs co-registered to the whole-brain functional image and in searchlight analyses (see below). For the ROI analyses, preprocessed data were intersected with the participant-specific anterior hippocampus and anterolateral entorhinal cortex ROI masks as well as a brain mask obtained during preprocessing (only voxels within the brain mask in all mini-blocks were analyzed) and the gray matter mask. For each voxel within the ROI mask, motion parameters from FSL MCFLIRT were used as predictors in a general linear model (GLM) with the voxel time series as the dependent variable. The residuals of this GLM (i.e. data that could not be explained by motion) were taken to the next analysis step. As the presentation of images in the picture viewing tasks was locked to the onset of a new volume (see above), the second volume after image onset was selected for every trial, effectively covering the time between 2270 and 4540 ms after stimulus onset. Only data for the 20 event images that were shown in the learning task were analyzed; data for the target stimulus were discarded. The similarity between the multi-voxel activity pattern for every event image in every mini-block with the pattern of every other event in every other mini-block was quantified using Pearson correlation coefficients. Thus, comparisons of scenes from the same mini-block were excluded. Next, we calculated mean, Fisher z-transformed correlation coefficients for every pair of events, yielding separate matrices of pattern similarity estimates for the pre- and the post-learning picture viewing tasks (Figure 3).

In order to assess changes in representational similarity between the two picture viewing tasks, we quantified pattern similarity changes as the difference of the respective correlation coefficients for every pair of events between the post-learning picture viewing task and its pre-learning baseline equivalent (Figure 3). Then, we analyzed how these difference values related to temporal relations between events, which we quantified using the absolute distances in virtual time ("virtual time") between events (Figure 1C, bottom right). We further tested whether the effect of virtual time on anterior hippocampal pattern similarity change persisted when including the absolute difference between sequence positions ("order") and the interval in seconds between events ("real time") as control predictors of no interest in the model. We separately tested the effect of virtual time for event pairs from the same or different sequences and used a Bonferroni-corrected α-level of 0.025 for these tests. Time metrics were z-scored within each participant prior to analysis. To implement these tests, we employed two approaches to model-based RSA that are described in detail below.

Summary Statistics Approach

In the summary statistics approach, we used the different time metrics as predictors for pattern similarity change. We set up a GLM with the given variable from the day learning task as a predictor and the pairwise representational change values as the criterion for every participant. The t-values of the resulting model coefficients were then compared to a null distribution obtained from shuffling the dependent variable of the linear model (i.e. pattern similarity change) 10,000 times. This approach to permutation-testing of regression coefficients controls Type I errors even under situations of collinear regressors (Anderson and Legendre, 1999). Resulting p-values for each coefficient were transformed to a Z-score. The Z-scores were then used for group-level inferential statistics.

Group-level statistics were carried out using permutation-based procedures. For t-tests, we compared the observed t-values against a surrogate distribution obtained from 10,000 random sign-flips to non-parametrically test against 0 or to assess within-participant differences between conditions. Permutation-based repeated measures ANOVAs were carried out using the permuco package (Frossard and Renaud, 2019).

Linear Mixed Effects

Second, we employed linear mixed models to assess how learned sequence relationships were reflected in pattern similarity change using the lme4 package (Bates et al., 2015). Mixed models have the advantage of estimating fixed effects and their interactions using all data, rather than performing inferential statistics on just one value per participant. We used the different time metrics as the fixed effects of interest. Factorial predictors (region of interest: anterior hippocampus and anterior-lateral entorhinal cortex; sequence: same vs. different) were deviation-coded. Within-subject dependencies were captured using random effects. Following the recommendation by Barr et al. (2013), we always first attempted to fit a model with a maximal random effects structure including random intercepts and random slopes for participants. If these models did not converge or resulted in singular fits, we reduced the random effects structure. We always kept random slopes for the fixed effect of interest in the model to avoid anti-conservativity when testing fixed effects or

their interactions (Barr, 2013; Barr et al., 2013). The mixed effects models were fitted using maximum likelihood estimation.

We assessed the statistical significance of fixed effects of interest using likelihood ratio tests. Specifically, the model including the fixed effect of interest was compared against a nested, reduced model excluding this effect, but with the same random effects structure. Throughout the manuscript we report the results of these model comparisons (χ^2-tests with one degree of freedom) and refer to supplemental tables for summaries of the final mixed model parameters. We visualize fixed effect estimates with their 95% confidence intervals as dot plots and further illustrate effects using estimated marginal means (Lüdecke, 2018).

Searchlight Analysis

We further probed how temporal distances between events shaped representational change using searchlight analyses. Using the procedures described above, we calculated pattern similarity change values for search spheres with a radius of 3 voxels around the center voxel. Search spheres were centered on all brain voxels within our field of view. Within a given search sphere, only gray matter voxels were analyzed. Search spheres not containing more than 25 gray matter voxels were discarded. For each search sphere, we implemented linear models to quantify the relationship between representational change and the learned temporal structure. Specifically, we assessed the relationship of pattern similarity change and absolute virtual temporal distances, separately for event pairs from the same sequences and from pairs from different sequences. In a third model, we included all event pairs and tested for an interaction effect of sequence membership (same or different) predictor and virtual temporal distances. The t-values of the respective regressors of interest were stored at the center voxel of a given search sphere.

The resulting t-maps were registered to MNI space for group level statistics and spatially smoothed (FWHM 3mm). Group level statistics were carried out using random sign flipping implemented with FSL Randomise and threshold-free cluster enhancement. We corrected for multiple comparisons using a small volume correction mask including our a priori regions of interest, the anterior hippocampus and the anterior-lateral entorhinal cortex. Further, we used a liberal threshold of $p_{uncorrected}<0.001$ to explore the data for additional effects within our field of view. Exploratory searchlight results are shown in Supplemental Figure 7 and clusters with a minimum extent of 30 voxels are listed in Supplemental Tables 9, 11 and 12.

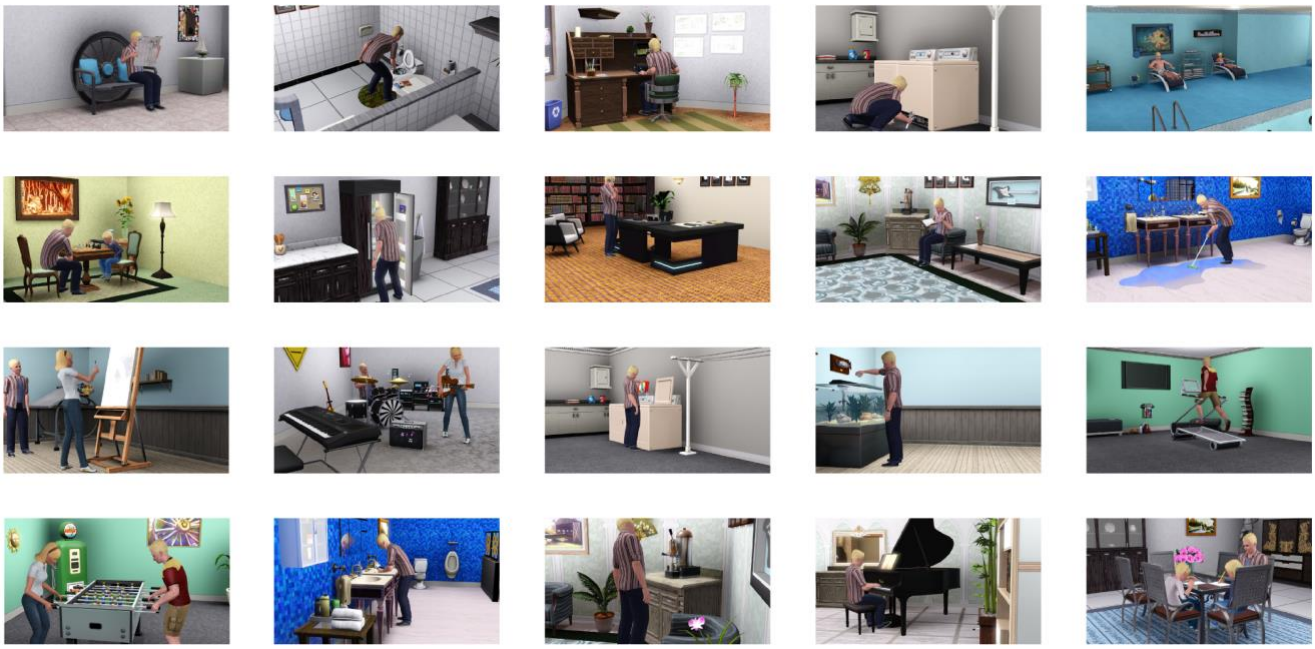
To test whether within- and across-sequence representations overlap, we defined an ROI based on the within-sequence searchlight analysis. Specifically, voxels belonging to the cluster around the peak voxel, thresholded at $p<0.01$ uncorrected within our small volume correction mask, were included. The analysis of representational change was then carried out as described for the other ROIs above.

Relationship to behavior

We used the regression coefficients quantifying the strength of the behavioral generalization bias to test for an across-subject relationship with the RSA searchlight effects. For each participant, we extracted the t-value of the across-sequence and the within-sequence searchlight effects from the peak voxel in our a priori regions of interest. We used Spearman correlations to test for a relationship of the RSA searchlight effects and the behavioral generalization bias ($\alpha=0.025$, corrected for two comparisons).

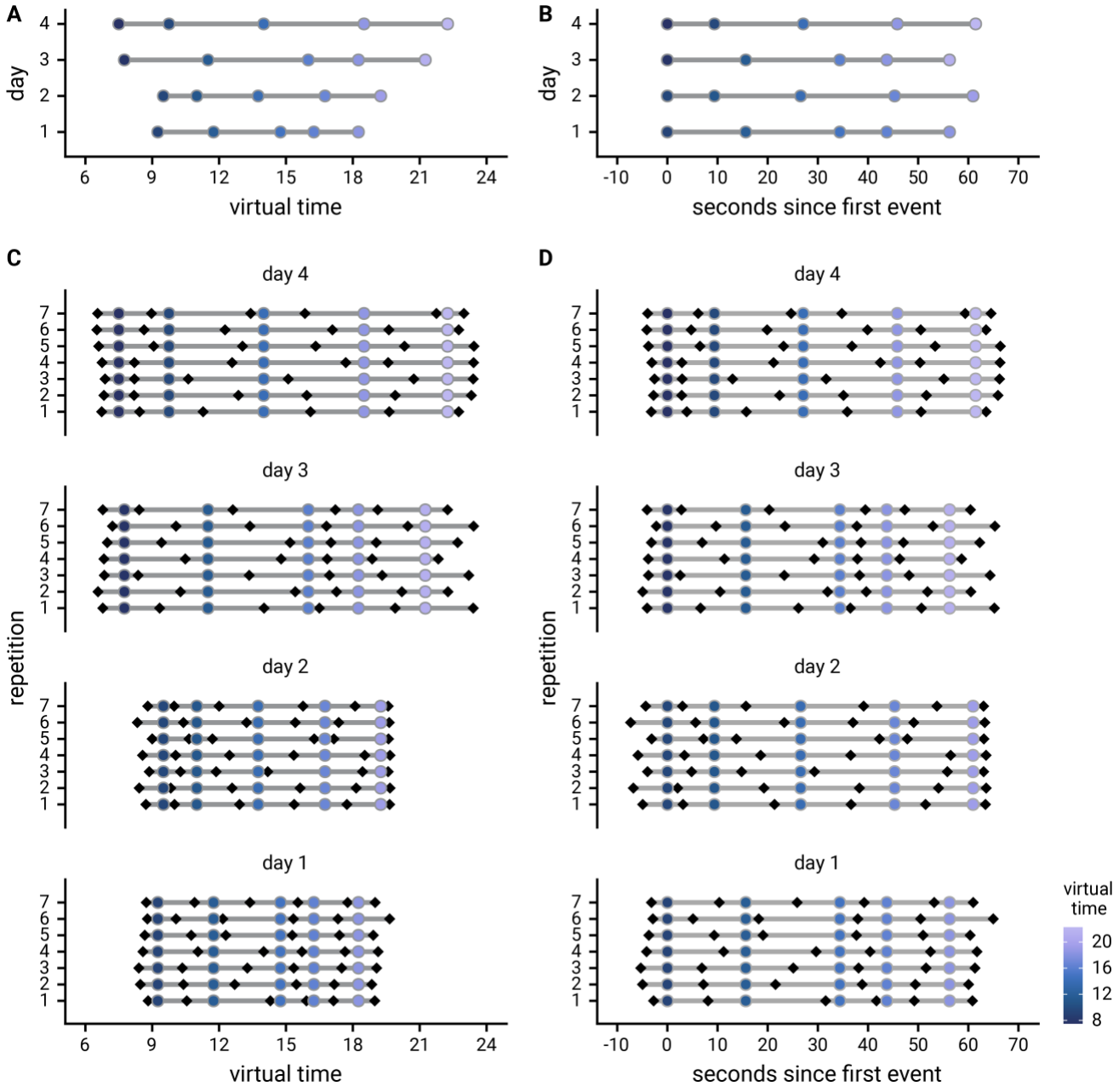
Supplemental Figures

Supplemental Figure 1



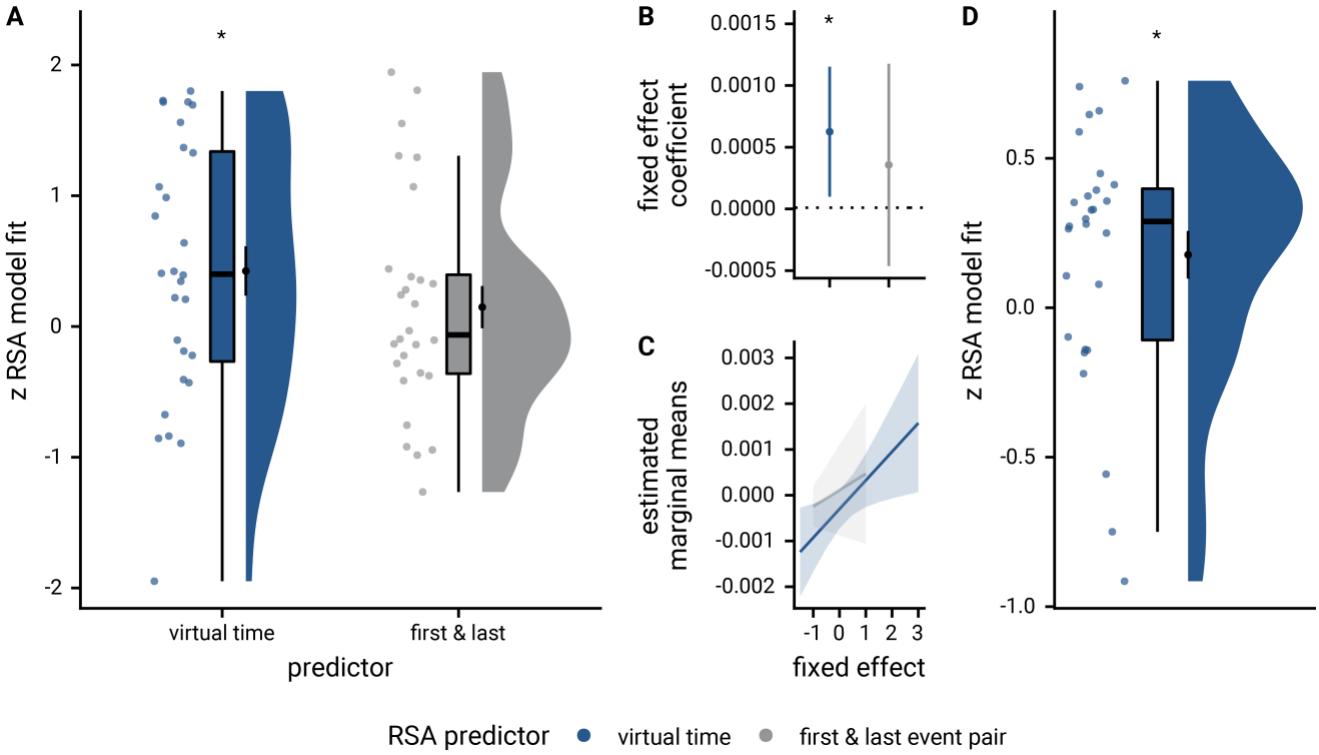
Supplemental Figure 1. Overview of the event images used as stimuli. All scenes were devoid of windows to exclude diurnal cues, such as shadows or light color, and were selected so they would be plausible at any time of day. For each participant, event images were randomly allocated to sequences and event times. Event images were created using the life-simulation computer game The Sims 3 (Electronic Arts). The Sims 3 and screenshots of it are licensed property of Electronic Arts, Inc.

Supplemental Figure 2



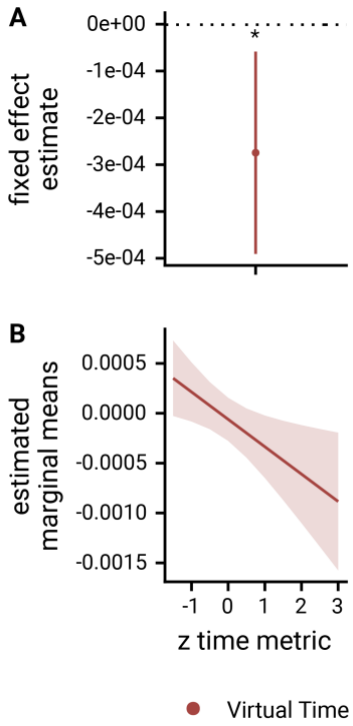
Supplemental Figure 2. Design of the day learning task. **A**. Each of the four virtual days consisted of a sequence of five events. Event sequences are shown in virtual time, i.e. relative to the hidden clock. Less virtual time passes within the bottom two sequences because clock speed was manipulated between sequences. **B**. Event sequences shown in real time relative to the first event. A comparable amount of real time (in seconds) elapses during each event sequence despite different amounts of virtual time passing. **C**, **D**. Sequences in virtual and real time as shown in (**A**) and (**B**), respectively, but separately for each of the seven repetitions of each sequence during the learning task. Black diamonds indicate the time cues shown to one randomly chosen example participant during the task. Time cues varied across repetitions and differed across participants.

Supplemental Figure 3



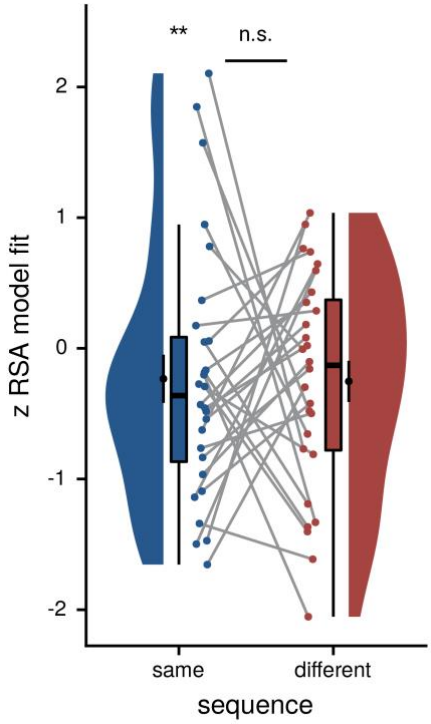
Supplemental Figure 3. The relationship of virtual time and hippocampal pattern similarity change is not driven by the first and last event of a sequence. **A.** Z-values from the summary statistics approach show a significant positive effect of virtual time on pattern similarity change in the anterior hippcampus when competing for variance with a control predictor of no interest accounting for variance explained by whether pairs of events were made up from the first and last event of a sequence or not. **B, C.** Fixed effect estimate with 95% confidence intervals (**B**) and estimated marginal means (**C**) visualize the results of the corresponding mixed model. **D.** We implemented participant-specific regression analyses with order and real time distances as predictors of hippocampal pattern similarity change. The plot shows a significant effect of virtual temporal distances when tested on the residuals of these linear models. Thus, variance that cannot be explained by the other time metrics can be accounted for by virtual temporal distances. This analysis was conducted only using the summary statistics approach because the residuals of a mixed model are more difficult to interpret than those of participant-specific regression analyses using ordinary least squares. **A,D.** Circles show individual participant Z-values from the summary statistics approach; boxplot shows median and upper/lower quartile along with whiskers extending to most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bars corresponds to mean±S.E.M.; distribution shows probability density function of data points. * p<0.05

Supplemental Figure 4



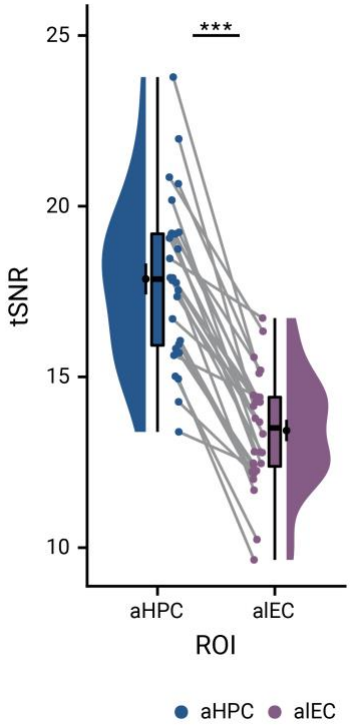
Supplemental Figure 4. Virtual time predicts hippocampal pattern similarity change for events from different sequences. **A, B.** Fixed effect estimate with 95% confidence interval (**A**) and estimated marginal means (**B**) for the effect of virtual time on pattern similarity for events from different sequences are shown to illustrate the across-sequences generalization effect as observed in the linear mixed model analysis. This mixed model complements the summary statistics results shown in Figure 5A. * p<0.05

Supplemental Figure 5



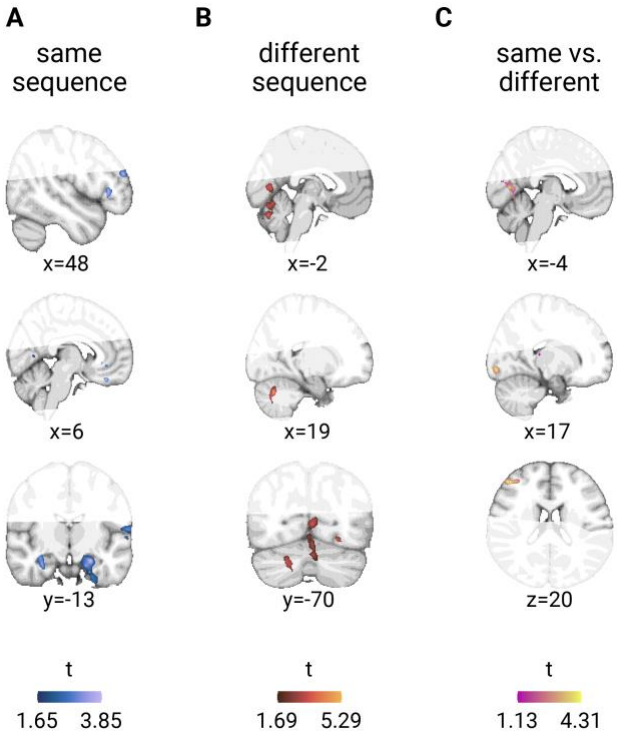
Supplemental Figure 5. Relationship of pattern similarity change and temporal distances between events from the same and different sequences in the anterior-lateral entorhinal cortex. There was no statistically significant difference between correlations of virtual temporal distances and representational change in the anterior-lateral entorhinal cortex depending on whether event pairs were from the same or different sequences. Entorhinal representational change was negatively related to temporal distances between events from the same sequence (summary statistics: t_{24}=-3.54, p=0.002; α=0.025, corrected for separate tests of events of the same and different sequences; three outliers excluded based on the boxplot criterion). The relationship between entorhinal pattern similarity change for events from different sequences was not statistically different from zero (summary statistics: t_{27}=-1.60, p=0.122; α=0.025, corrected for separate tests of events of the same and different sequences). Circles show participant-specific Z-values from summary statistics approach; boxplot shows median and upper/lower quartile along with whiskers extending to most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bars corresponds to mean±S.E.M.; distribution shows probability density function of data points. ** p<0.01 after outlier exclusion

Supplemental Figure 6



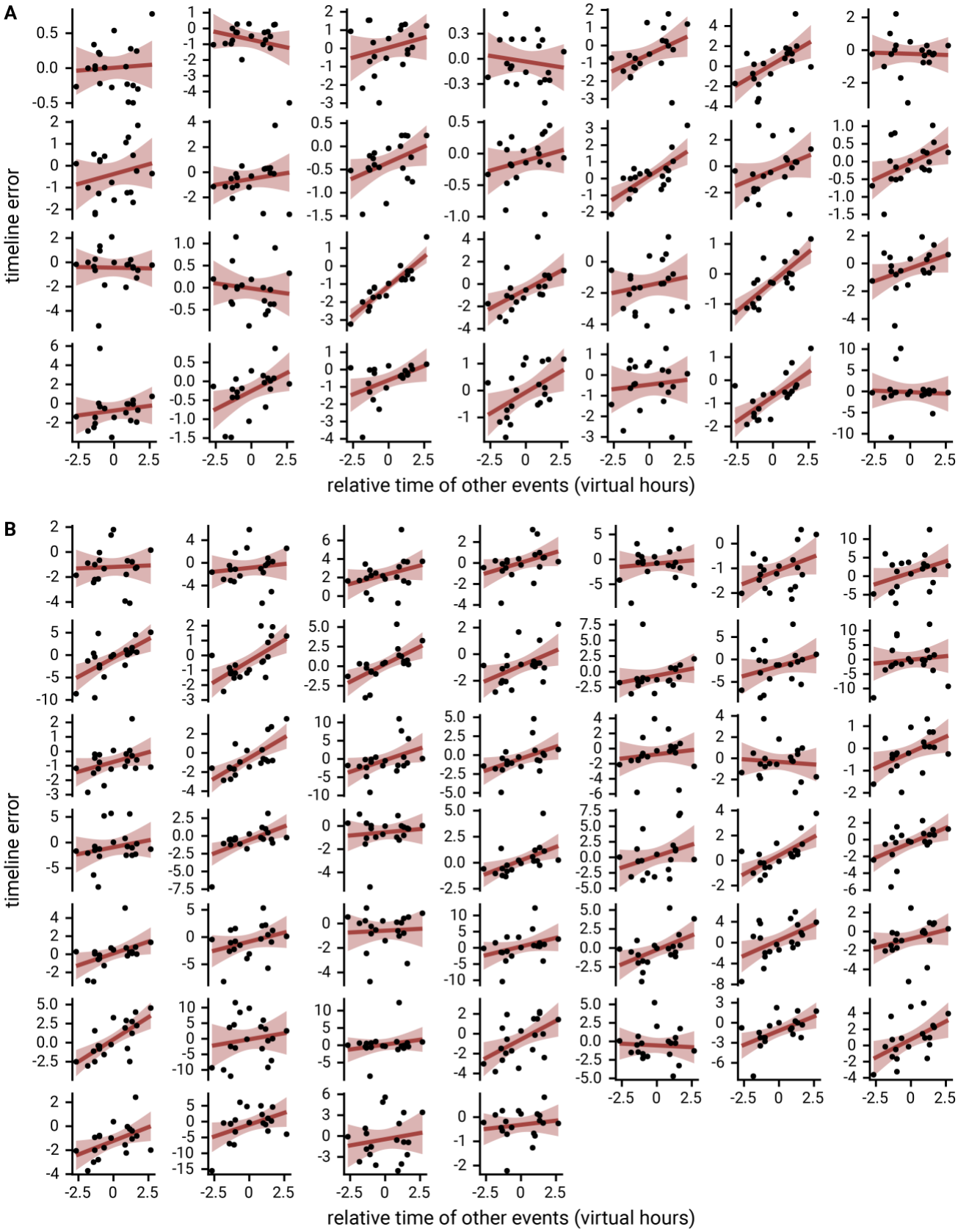
Supplemental Figure 6. Temporal signal-to-noise ratio in the anterior hippocampus and the anterior-lateral entorhinal cortex. A. The temporal signal-to-noise ratio was quantified as the mean unsmoothed signal over time divided by its standard deviation. It was calculated for each voxel and then averaged across voxels in a region of interest. The temporal signal-to-noise ratio was higher in the anterior hippocampus (aHPC) than in the anterior-lateral entorhinal cortex (alEC, summary statistics: t_{27}=12.43, p<0.001). Circles show individual participant values; boxplot shows median and upper/lower quartile along with whiskers extending to most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bars corresponds to mean±S.E.M.; distribution shows probability density function of data points. *** p<0.001

Supplemental Figure 7



Supplemental Figure 7. Exploratory searchlight results. A. For same-sequence event pairs, clusters of voxels in which pattern similarity change correlated positively with temporal distances were detected in the frontal pole, frontal medial cortex and left entorhinal cortex (see Supplemental Table 9). **B.** Pattern similarity change correlated negatively with temporal distances between events from different sequences in the cerebellum and lingual gyrus (see Supplemental Table 11). **C.** The interaction effect, defined as correlations of temporal distances and pattern similarity change depending on whether pairs of events belonged to the same sequence or not, was observed in the occipital pole, lingual gyrus, frontal pole, temporal fusiform cortex and the intracalcerine sulcus (see Supplemental Table 12). **A-C.** Statistical images are thresholded at $p<0.01$ uncorrected for display purposes. No clusters outside the hippocampal-entorhinal region survived corrections for multiple comparisons.

Supplemental Figure 8



Supplemental Figure 8. Generalization bias in individual participants. A, B. Each panel shows the data from one participant. Each circle corresponds to one event. The x-axis indicates the average relative time of the events occupying the same sequence position in other sequences. The y-axis shows the signed error of constructed event times as measured in the timeline task. The regression line and its confidence interval are overlaid in red. Positive slopes of the regression line indicate that constructed event times are biased by the average time of events in the other sequences. **A** shows data from the main sample; **B** from the replication sample.

Supplemental Tables

Supplemental Table 1

Mixed Model: Virtual time explains constructed times with order and real time in the model

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	14.010019	0.069962	200.25	13.868056	14.151981
virtual time	3.069324	0.259967	11.81	2.558874	3.579774
order	1.667630	0.430230	3.88	0.822785	2.512476
real time	-0.332261	0.473306	-0.70	-1.261696	0.597173

random effects		
group	**term**	**estimate**
participant	intercept	0.221991
participant	virtual time (SD)	0.232089
participant	correlation random intercepts and random slopes	0.165592
residual	SD	1.324919

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	7	2053.90	-1019.95			
full model	8	1939.95	-961.98	115.95	1	4.88e-27

model: memory_time~virtual_time_z+order_z+real_time_z+(1+virtual_time_z|sub_id);
SE: standard error, CI: confidence interval, SD: standard deviation, npar: number of parameters, LL: log likelihood, df: degrees of freedom, corr.: correlation

Supplemental Table 2

Mixed Model: Virtual time explains representational change for same-sequence events in the anterior hippocampus

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	-0.000326	0.000211	-1.54	-0.000740	0.000088
virtual time	0.000751	0.000220	3.42	0.000307	0.001196

random effects		
group	**term**	**estimate**
participant	intercept (SD)	0.000001
participant	virtual time	0.000257
residual	SD	0.006917

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	4	-7943.56	3975.78			
full model	5	-7951.43	3980.72	9.87	1	0.002

model: ps_change~vir_time_diff+((1|sub_id)+(0+vir_time_diff|sub_id));
SE: standard error, CI: confidence interval, SD: standard deviation, npar: number of parameters, LL: log likelihood, df: degrees of freedom, corr.: correlation

Supplemental Table 3

Mixed Model: Virtual time explains representational change for same-sequence events in the anterior hippocampus when controlling for the effect of first-last event pairs

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	-0.000015	0.000421	-0.04	-0.000841	0.000810
virtual time	0.000626	0.000264	2.37	0.000099	0.001152
first-last pair	0.000357	0.000418	0.85	-0.000462	0.001176

random effects		
group	**term**	**estimate**
participant	intercept (SD)	0.000001
participant	virtual time (SD)	0.000258
residual	SD	0.006914

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	5	-7946.81	3978.40			
full model	6	-7950.16	3981.08	5.36	1	0.021

model: ps_change~vir_time_diff+first_last+((1|sub_id)+(0+vir_time_diff|sub_id));
SE: standard error, CI: confidence interval, SD: standard deviation, npar: number of parameters, LL: log likelihood, df: degrees of freedom, corr.: correlation

Supplemental Table 4

Mixed Model: Virtual time explains representational change for same-sequence events in the anterior hippocampus when including order and real time in the model

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	-0.000281	0.000219	-1.28	-0.000711	0.000149
virtual time	0.001321	0.000541	2.44	0.000258	0.002383
order	0.000012	0.000908	0.01	-0.001768	0.001793
real time	-0.000676	0.001019	-0.66	-0.002675	0.001323

random effects		
group	**term**	**estimate**
participant	virtual time (SD)	0.000260
residual	SD	0.006913

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	5	-7946.84	3978.42			
full model	6	-7950.76	3981.38	5.92	1	0.015

model: ps_change~vir_time_diff+order_diff+real_time_diff+(0+vir_time_diff|sub_id);
SE: standard error, CI: confidence interval, SD: standard deviation, npar: number of parameters, LL: log likelihood, df: degrees of freedom, corr.: correlation

Supplemental Table 5

Mixed Model: Virtual time explains representational change for different-sequence events in the anterior hippocampus

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	-0.000061	0.000110	-0.55	-0.000276	0.000155
virtual time	-0.000275	0.000110	-2.51	-0.000491	-0.000058

random effects		
group	**term**	**estimate**
participant	virtual time (SD)	0.000000
residual	SD	0.007107

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	3	-29621.39	14813.69			
full model	4	-29625.40	14816.70	6.01	1	0.014

model: ps_change~vir_time_diff+(0+vir_time_diff|sub_id);
SE: standard error, CI: confidence interval, SD: standard deviation, npar: number of parameters, LL: log likelihood, df: degrees of freedom, corr.: correlation

Supplemental Table 6

Mixed Model: The effect of virtual time differs between same-sequence and different-sequence events in the anterior hippocampus

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	-0.000193	0.000121	-1.60	-0.000430	0.000044
virtual time	0.000238	0.000122	1.95	-0.000001	0.000478
day	-0.000133	0.000121	-1.10	-0.000370	0.000104
interaction virtual time and day	0.000513	0.000127	4.05	0.000261	0.000765

random effects		
group	**term**	**estimate**
participant	interaction virtual time and day (SD)	0.000176
residual	SD	0.007066

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	5	-37569.38	18789.69			
full model	6	-37581.75	18796.87	14.37	1	1.50e-04

model: ps_change~vir_time_diff*same_day_dv+(0+vir_time_diff:same_day_dv|sub_id);
SE: standard error, CI: confidence interval, SD: standard deviation, npar: number of parameters, LL: log likelihood, df: degrees of freedom, corr.: correlation

Supplemental Table 7

Mixed Model: Virtual time explains representational change in the anterior-lateral entorhinal cortex (all events)

fixed effects						
term		**estimate**	**SE**	**t-value**	**95% CI**	
intercept		0.000167	0.000202	0.83	-0.000229	0.000563
virtual time		-0.000424	0.000202	-2.09	-0.000820	-0.000027
random effects						
group	**term**					**estimate**
participant	virtual time (SD)					0.000000
residual	SD					0.014734
model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	3	-29767.39	14886.69			
full model	4	-29769.77	14888.89	4.39	1	0.036

model: ps_change~vir_time_diff+(0+vir_time_diff|sub_id);
SE: standard error, CI: confidence interval, SD: standard deviation, npar: number of parameters, LL: log likelihood, df: degrees of freedom, corr.: correlation

Supplemental Table 8

Mixed Model: The effect of virtual time differentially depends on sequence membership in the anterior hippocampus and the anterior-lateral entorhinal cortex

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	-0.000193	0.000219	-0.89	-0.000622	0.000235
virtual time	0.000238	0.000200	1.19	-0.000153	0.000630
day	-0.000133	0.000197	-0.67	-0.000520	0.000254
ROI	0.000455	0.000279	1.63	-0.000093	0.001002
virtual time * day	0.000513	0.000202	2.54	0.000117	0.000909
virtual time * ROI	-0.000810	0.000282	-2.87	-0.001363	-0.000257
day * ROI	0.000261	0.000279	0.94	-0.000286	0.000808
virtual time * day * ROI	-0.000745	0.000294	-2.54	-0.001321	-0.000169

random effects		
group	**term**	**estimate**
participant	intercept (SD)	0.000496
participant	corr. intercept, virtual time:day:ROI1	-1.000000
participant	corr. intercept, virtual time:day:ROI-1	-0.151340
participant	virtual time:day:ROI1 (SD)	0.000170
participant	corr. virtual time:day:ROI1, virtual time:day:ROI-1	0.151340
participant	virtual time:day:ROI-1 (SD)	0.000421
residual	SD	0.011540

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	14	-64699.87	32363.94			
full model	15	-64704.19	32367.09	6.31	1	0.012

model: ps_change~vir_time_diff*same_day_dv*roi_dv+(1+vir_time_diff:same_day_dv:roi_dv|sub_id);
SE: standard error, CI: confidence interval, SD: standard deviation, npar: number of parameters, LL: log likelihood, df: degrees of freedom, corr.: correlation

Supplemental Table 9

Searchlight Analysis: Virtual time explains representational change for same-sequence events

Searchlight results in a priori regions of interest, p-values corrected using small volume correction									
Atlas Label	**Voxel Extent**	**x**	**y**	**z**	**COG x**	**COG y**	**COG z**	**t**	**p**
left hippocampus	193	-24	-13	-20	-23.3	-13.1	-19.8	4.53	0.006
right hippocampus	96	31	-16	-20	30.1	-16.7	-19.8	3.56	0.035
left hippocampus	76	-27	-20	-15	-27.9	-19.5	-16.6	3.47	0.029
Exploratory searchlight results, p-values uncorrected									
Atlas Label	**Voxel Extent**	**x**	**y**	**z**	**COG x**	**COG y**	**COG z**	**t**	**p**
frontal pole	399	50	44	16	48.3	41.6	19.2	3.96	0.0002
frontal pole	173	53	41	-7	51.1	42.9	-4.45	4.56	0.0002
left entorhinal cortex	119	-18	-16	-32	-21.2	-14.6	-31.2	3.45	0.0004
inferior frontal gyrus	91	40	27	2	44.2	28	3.59	4.29	0.0002
lingual gyrus	86	-17	-58	-15	-15.7	-56.9	-9.64	3.82	0.0002
frontal medial cortex	49	7	35	-23	6.29	36.7	-24.1	4.28	0.0004

x, y, z refer to MNI coordinates of minimum p-value in cluster, t denotes the most extreme t-value, COG: center of gravity

Supplemental Table 10

Mixed Model: Virtual time explains representational change for different-sequence events in the peak cluster of the same-sequence searchlight analysis

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	-0.000097	0.000234	-0.41	-0.000557	0.000362
virtual time	-0.000478	0.000234	-2.04	-0.000939	-0.000018

random effects		
group	**term**	**estimate**
participant	virtual time (SD)	0.000000
residual	SD	0.015162

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	3	-23257.87	11631.93			
full model	4	-23260.00	11634.00	4.13	1	0.042

model: ps_change~vir_time_diff+(0+vir_time_diff|sub_id);
SE: standard error, CI: confidence interval, SD: standard deviation, npar: number of parameters, LL: log likelihood, df: degrees of freedom, corr.: correlation

Supplemental Table 11

Searchlight Analysis: Virtual time explains representational change for different-sequence events

Exploratory searchlight results, p-values uncorrected									
Atlas Label	**Voxel Extent**	**x**	**y**	**z**	**COG x**	**COG y**	**COG z**	**t**	**p**
cerebellum	314	19	-68	-34	19.1	-66.3	-29.6	-5.37	0.0002
cerebellum	104	-1	-68	-14	-1.86	-69.1	-14.3	-3.44	0.0002
lingual gyrus	100	-1	-70	4	-2.68	-70.5	4.56	-3.73	0.0002

x, y, z refer to MNI coordinates of minimum p-value in cluster, t denotes the most extreme t-value, COG: center of gravity

Supplemental Table 12

Searchlight Analysis: Interaction of virtual time and sequence membership

Searchlight results in a priori regions of interest, p-values corrected using small volume correction									
Atlas Label	**Voxel Extent**	**x**	**y**	**z**	**COG x**	**COG y**	**COG z**	**t**	**p**
left hippocampus	359	-26	-20	-15	-23.4	-15.5	-18.6	4.15	0.014
right hippocampus	335	31	-16	-21	30.7	-15.1	-20.1	4.25	0.007
Exploratory searchlight results, p-values uncorrected									
Atlas Label	**Voxel Extent**	**x**	**y**	**z**	**COG x**	**COG y**	**COG z**	**t**	**p**
occipital pole	103	17	-91	-8	17.7	-90.6	-6.62	4.08	0.0002
lingual gyrus	102	-5	-73	5	-3.59	-70.4	5.01	3.72	0.0002
frontal pole	96	43	43	18	45.4	43.4	19.7	4.31	0.0006
frontal pole	45	35	43	17	37	43.2	18.5	3.81	0.0006
temporal fusiform cortex	40	-25	-10	-45	-25.3	-10.3	-42.9	3.14	0.0004
intracalcarine sulcus	33	-4	-77	11	-2.85	-75.8	11.5	3.56	0.0002

x, y, z refer to MNI coordinates of minimum p-value in cluster, t denotes the most extreme t-value, COG: center of gravity

Supplemental Table 13
Mixed Model: Behavioral generalization bias

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	-0.352481	0.069962	-5.04	-0.494444	-0.210518
relative time other events	0.337262	0.067360	5.01	0.200579	0.473945

random effects		
group	**term**	**estimate**
participant	intercept	0.220016
participant	relative time other events (SD)	-0.114173
participant	correlation random intercepts and random slopes	0.183681
residual	SD	1.331485

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	5	1958.57	-974.29			
full model	6	1942.67	-965.34	17.90	1	2.32e-05

model: timeline_error~rel_time_other_events_z+(1+rel_time_other_events_z|sub_id);
SE: standard error, CI: confidence interval, SD: standard deviation, npar: number of parameters, LL: log likelihood, df: degrees of freedom, corr.: correlation

Supplemental Table 14
Mixed Model: Behavioral generalization bias (replication)

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	-0.320564	0.089155	-3.60	-0.495488	-0.145640
relative time other events	0.863631	0.091472	9.44	0.684152	1.043110

random effects		
group	**term**	**estimate**
participant	relative time other events (SD)	0.000000
residual	SD	2.704218

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	3	4501.04	-2247.52			
full model	4	4449.30	-2220.65	53.74	1	2.29e-13

model: timeline_error~rel_time_other_events_z+(0+rel_time_other_events_z|sub_id);
SE: standard error, CI: confidence interval, SD: standard deviation, npar: number of parameters, LL: log likelihood, df: degrees of freedom, corr.: correlation

References

Addis, D.R. (2020). Mental Time Travel? A Neurocognitive Model of Event Simulation. Rev. Philos. Psychol. *11*, 233–259.

Allen, M., Poggiali, D., Whitaker, K., Marshall, T.R., and Kievit, R.A. (2019). Raincloud plots: a multi-platform tool for robust data visualization. Wellcome Open Res. *4*, 63.

Anderson, M.J., and Legendre, P. (1999). An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. J. Stat. Comput. Simul. *62*, 271–303.

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

Baram, A.B., Muller, T.H., Nili, H., Garvert, M.M., and Behrens, T.E.J. (2020). Entorhinal and ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning problems. Neuron.

Barr, D.J. (2013). Random effects structure for testing interactions in linear mixed-effects models. Front. Psychol. *4*, 328.

Barr, D.J., Levy, R., Scheepers, C., and Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. J. Mem. Lang. *68*, 255–278.

Barron, H.C., Dolan, R.J., and Behrens, T.E.J. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. Nat. Neurosci. *16*, 1492–1498.

Bartlett, F.C. (1932). Remembering: a study in experimental and social psychology (Cambridge, UK: Cambridge University Press).

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. *67*, 1–48.

Behrens, T.E.J., Muller, T.H., Whittington, J.C.R., Mark, S., Baram, A.B., Stachenfeld, K.L., and Kurth-Nelson, Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. Neuron *100*, 490–509.

Bellmund, J.L.S., Gärdenfors, P., Moser, E.I., and Doeller, C.F. (2018). Navigating cognition: Spatial codes for human thinking. Science *362*, eaat6766.

Bellmund, J.L.S., Deuker, L., and Doeller, C.F. (2019). Mapping sequence structure in the human lateral entorhinal cortex. ELife *8*, e45333.

Bellmund, J.L.S., Polti, I., and Doeller, C.F. (2020a). Sequence Memory in the Hippocampal–Entorhinal Region. J. Cogn. Neurosci. *32*, 2056–2070.

Bellmund, J.L.S., Cothi, W. de, Ruiter, T.A., Nau, M., Barry, C., and Doeller, C.F. (2020b). Deforming the metric of cognitive maps distorts memory. Nat. Hum. Behav. *4*, 177–188.

Benoit, R.G., and Schacter, D.L. (2015). Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. Neuropsychologia *75*, 450–457.

Ben-Yakov, A., and Dudai, Y. (2011). Constructing Realistic Engrams: Poststimulus Activity of Hippocampus and Dorsal Striatum Predicts Subsequent Episodic Memory. J. Neurosci. *31*, 9032–9042.

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

Bright, I.M., Meister, M.L.R., Cruzado, N.A., Tiganj, Z., Buffalo, E.A., and Howard, M.W. (2020). A temporal record of the past with a spectrum of time constants in the monkey entorhinal cortex. Proc. Natl. Acad. Sci. *117*, 20274–20283.

Bunsey, M., and Eichenbaum, H. (1996). Conservation of hippocampal memory function in rats and humans. Nature *379*, 255–257.

Carpenter, A.C., and Schacter, D.L. (2017). Flexible retrieval: When true inferences produce false memories. J. Exp. Psychol. Learn. Mem. Cogn. *43*, 335–349.

Chanales, A.J.H., Oza, A., Favila, S.E., and Kuhl, B.A. (2017). Overlap among Spatial Memories Triggers Repulsion of Hippocampal Representations. Curr. Biol. *27*, 2307-2317.e5.

Cheng, S., Werning, M., and Suddendorf, T. (2016). Dissociating memory traces and scenario construction in mental time travel. Neurosci. Biobehav. Rev. *60*, 82–89.

Cohen, N.J., and Eichenbaum, H. (1993). Memory, amnesia, and the hippocampal system. (Cambridge, MA, US: The MIT Press).

Deuker, L., Bellmund, J.L.S., Navarro Schröder, T., and Doeller, C.F. (2016). An event map of memory space in the hippocampus. ELife *5*, e16534.

Devitt, A.L., Addis, D.R., and Schacter, D.L. (2017). Episodic and semantic content of memory and imagination: A multilevel analysis. Mem. Cognit. *45*, 1078–1094.

Diamond, N.B., and Levine, B. (2020). Linking Detail to Temporal Structure in Naturalistic-Event Recall. Psychol. Sci. *31*, 1557–1572.

Diana, R.A., Yonelinas, A.P., and Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. Trends Cogn. Sci. *11*, 379–386.

Doeller, C.F., Opitz, B., Krick, C.M., Mecklinger, A., and Reith, W. (2005). Prefrontal-hippocampal dynamics involved in learning regularities across episodes. Cereb. Cortex *15*, 1123–1133.

Doeller, C.F., Opitz, B., Krick, C.M., Mecklinger, A., and Reith, W. (2006). Differential hippocampal and prefrontal-striatal contributions to instance-based and rule-based learning. NeuroImage *31*, 1802–1816.

DuBrow, S., and Davachi, L. (2014). Temporal Memory Is Shaped by Encoding Stability and Intervening Item Reactivation. J. Neurosci. *34*, 13998–14005.

Dusek, J.A., and Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. Proc. Natl. Acad. Sci. *94*, 7109–7114.

Ebbinghaus, H. (1885). Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie (Leipzig: Duncker & Humblot).

Eichenbaum, H. (2014). Time cells in the hippocampus:

a new dimension for mapping memories. Nat. Rev. Neurosci. *15*, 732–744.

Eichenbaum, H., and Cohen, N.J. (1988). Representation in the hippocampus: what do hippocampal neurons code? Trends Neurosci. *11*, 244–248.

Estefan, D.P., Zucca, R., Arsiwalla, X., Principe, A., Zhang, H., Rocamora, R., Axmacher, N., and Verschure, P.F.M.J. (2021). Volitional learning promotes theta phase coding in the human hippocampus. Proc. Natl. Acad. Sci. *118*, e2021238118.

Evensmoen, H.R., Rimol, L.M., Rise, H.H., Hansen, T.I., Nili, H., Winkler, A.M., and Håberg, A. (2020). Metric and chronological time in human episodic memory. BioRxiv 2020.05.11.084202.

Ezzyat, Y., and Davachi, L. (2014). Similarity Breeds Proximity: Pattern Similarity within and across Contexts Is Related to Later Mnemonic Judgments of Temporal Proximity. Neuron *81*, 1179–1189.

Favila, S.E., Chanales, A.J.H., and Kuhl, B.A. (2016). Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. Nat. Commun. *7*, 11066.

Franklin, N.T., Norman, K.A., Ranganath, C., Zacks, J.M., and Gershman, S.J. (2020). Structured Event Memory: A neuro-symbolic model of event cognition. Psychol. Rev. *127*, 327–361.

Friedman, W.J. (1993). Memory for the time of past events. Psychol. Bull. *113*, 44.

Friedman, W.J. (2004). Time in Autobiographical Memory. Soc. Cogn. *22*, 591–605.

Frisoni, M., Di Ghionno, M., Guidotti, R., Tosoni, A., and Sestieri, C. (2021). Reconstructive nature of temporal memory for movie scenes. Cognition *208*, 104557.

Frossard, J., and Renaud, O. (2019). permuco: Permutation tests for regression, (repeated measures) ANOVA/ANCOVA and comparison of signals.

Greenberg, D.L., and Verfaellie, M. (2010). Interdependence of episodic and semantic memory: Evidence from neuropsychology. J. Int. Neuropsychol. Soc. *16*, 748–753.

Hassabis, D., and Maguire, E.A. (2007). Deconstructing episodic memory with construction. Trends Cogn. Sci. *11*, 299–306.

Heckers, S., Zalesak, M., Weiss, A.P., Ditman, T., and Titone, D. (2004). Hippocampal activation during transitive inference in humans. Hippocampus *14*, 153–162.

Hemmer, P., and Steyvers, M. (2009). Integrating episodic memories and prior knowledge at multiple levels of abstraction. Psychon. Bull. Rev. *16*, 80–87.

Hemmer, P., Tauber, S., and Steyvers, M. (2015). Moving beyond qualitative evaluations of Bayesian models of cognition. Psychon. Bull. Rev. *22*, 614–628.

Howard, M.W., and Kahana, M.J. (2002). A Distributed Representation of Temporal Context. J. Math. Psychol. *46*, 269–299.

Hsieh, L.-T., and Ranganath, C. (2015). Cortical and subcortical contributions to sequence retrieval: Schematic coding of temporal context in the neocortical recollection network. NeuroImage *121*, 78–90.

Hsieh, L.-T., Gruber, M.J., Jenkins, L.J., and Ranganath, C. (2014). Hippocampal Activity Patterns Carry Information about Objects in Temporal Context. Neuron *81*, 1165–1178.

Irish, M., and Piguet, O. (2013). The Pivotal Role of Semantic Memory in Remembering the Past and Imagining the Future. Front. Behav. Neurosci. *7*, 27.

Irish, M., Addis, D.R., Hodges, J.R., and Piguet, O. (2012). Considering the role of semantic memory in episodic future thinking: evidence from semantic dementia. Brain *135*, 2178–2191.

Jazayeri, M., and Shadlen, M.N. (2010). Temporal context calibrates interval timing. Nat. Neurosci. *13*, 1020–1026.

Jenkins, L.J., and Ranganath, C. (2016). Distinct neural mechanisms for remembering when an event occurred. Hippocampus *26*, 554–559.

Jensen, O., and Lisman, J.E. (2005). Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. Trends Neurosci. *28*, 67–72.

Koster, R., Chadwick, M.J., Chen, Y., Berron, D., Banino, A., Düzel, E., Hassabis, D., and Kumaran, D. (2018). Big-Loop Recurrence within the Hippocampal System Supports Integration of Information across Episodes. Neuron *99*, 1342–1354.

Kraus, B.J., Robinson II, R.J., White, J.A., Eichenbaum, H., and Hasselmo, M.E. (2013). Hippocampal "Time Cells": Time versus Path Integration. Neuron *78*, 1090–1101.

Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., and Bandettini, P.A. (2008). Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. Neuron *60*, 1126–1141.

Kumaran, D., and Maguire, E.A. (2006a). An Unexpected Sequence of Events: Mismatch Detection in the Human Hippocampus. PLOS Biol. *4*, e424.

Kumaran, D., and Maguire, E.A. (2006b). The Dynamics of Hippocampal Activation during Encoding of Overlapping Sequences. Neuron *49*, 617–629.

Kumaran, D., and McClelland, J.L. (2012). Generalization Through the Recurrent Interaction of Episodic Memories. Psychol. Rev. *119*, 573–616.

Kyle, C.T., Smuda, D.N., Hassan, A.S., and Ekstrom, A.D. (2015). Roles of human hippocampal subfields in retrieval of spatial and temporal context. Behav. Brain Res. *278*, 549–558.

Lewandowsky, S., and Murdock, B.B. (1989). Memory for serial order. Psychol. Rev. *96*, 25–57.

Liu, Y., Dolan, R.J., Kurth-Nelson, Z., and Behrens, T.E.J. (2019). Human Replay Spontaneously Reorganizes Experience. Cell *178*, 640-652.e14.

Lohnas, L.J., Duncan, K., Doyle, W.K., Thesen, T., Devinsky, O., and Davachi, L. (2018). Time-resolved neural reinstatement and pattern separation during memory decisions in human hippocampus. Proc. Natl. Acad. Sci. *115*, E7418–E7427.

Lositsky, O., Chen, J., Toker, D., Honey, C.J.,

Shvartsman, M., Poppenk, J.L., Hasson, U., and Norman, K.A. (2016). Neural pattern change during encoding of a narrative predicts retrospective duration estimates. ELife *5*, e16070.

Lüdecke, D. (2018). ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. J. Open Source Softw. *3*, 772.

Luyckx, F., Nili, H., Spitzer, B., and Summerfield, C. (2019). Neural structure mapping in human probabilistic reward learning. ELife *8*, e42816.

MacDonald, C.J., Lepage, K.Q., Eden, U.T., and Eichenbaum, H. (2011). Hippocampal "Time Cells" Bridge the Gap in Memory for Discontiguous Events. Neuron *71*, 737–749.

McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995). Why There Are Complementary Learning Systems in the Hippocampus and Neocortex: Insights From the Successes and Failures of Connectionist Models of Learning and Memory. Psychol. Rev. *102*, 419–457.

Milivojevic, B., Vicente-Grabovetsky, A., and Doeller, C.F. (2015). Insight Reconfigures Hippocampal-Prefrontal Memories. Curr. Biol. *25*, 821–830.

Montchal, M.E., Reagh, Z.M., and Yassa, M.A. (2019). Precise temporal memories are supported by the lateral entorhinal cortex in humans. Nat. Neurosci. *22*, 284–288.

Montijn, N.D., Gerritsen, L., and Engelhard, I.M. (2021). The effect of stress on memory for temporal context: an exploratory study. BioRxiv.

Morton, N.W., Schlichting, M.L., and Preston, A.R. (2020). Representations of common event structure in medial temporal lobe and frontoparietal cortex support efficient inference. Proc. Natl. Acad. Sci. *117*, 29338–29345.

Moser, E.I., Moser, M.-B., and McNaughton, B.L. (2017). Spatial representation in the hippocampal formation: a history. Nat. Neurosci. *20*, 1448–1464.

Navarro Schröder, T., Haak, K.V., Zaragoza Jimenez, N.I., Beckmann, C.F., and Doeller, C.F. (2015). Functional topography of the human entorhinal cortex. ELife *4*, e06738.

Nielson, D.M., Smith, T.A., Sreekumar, V., Dennis, S., and Sederberg, P.B. (2015). Human hippocampus represents space and time during retrieval of real-world memories. Proc. Natl. Acad. Sci. *112*, 11078–11083.

O'Keefe, J., and Nadel, L. (1978). The Hippocampus as a Cognitive Map (Oxford: Clarendon Press).

Orlov, T., Yakovlev, V., Hochstein, S., and Zohary, E. (2000). Macaque monkeys categorize images by their ordinal number. Nature *404*, 77–80.

Park, S.A., Miller, D.S., Nili, H., Ranganath, C., and Boorman, E.D. (2020). Map Making: Constructing, Combining, and Inferring on Abstract Cognitive Maps. Neuron *107*, 1226–1238.

Pastalkova, E., Itskov, V., Amarasingham, A., and Buzsáki, G. (2008). Internally Generated Cell Assembly Sequences in the Rat Hippocampus. Science *321*, 1322–1327.

Polti, I., Nau, M., Kaplan, R., van Wassenhove, V., and Doeller, C.F. (2021). Rapid updating of sensorimotor representations in the human hippocampus and striatum. Annu. Meet. Cogn. Neurosci. Soc.

Poppenk, J., Evensmoen, H.R., Moscovitch, M., and Nadel, L. (2013). Long-axis specialization of the human hippocampus. Trends Cogn. Sci. *17*, 230–240.

Preston, A.R., Shrager, Y., Dudukovic, N.M., and Gabrieli, J.D.E. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. Hippocampus *14*, 148–152.

R Core Team (2020). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).

Radvansky, G.A., and Zacks, J.M. (2014). Event cognition (New York, NY, US: Oxford University Press).

Ranganath, C., and Hsieh, L. (2016). The hippocampus: a special place for time. Ann. N. Y. Acad. Sci. *1369*, 93–110.

Schacter, D.L., and Addis, D.R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. Philos. Trans. R. Soc. B Biol. Sci. *362*, 773–786.

Schacter, D.L., and Addis, D.R. (2020). Memory and imagination: Perspectives on constructive episodic simulation. In The Cambridge Handbook of Imagination, A. Abraham, ed. (Cambridge: Cambridge University Press), pp. 111–131.

Schacter, D.L., Addis, D.R., and Buckner, R.L. (2007). Remembering the past to imagine the future: the prospective brain. Nat. Rev. Neurosci. *8*, 657–661.

Schacter, D.L., Benoit, R.G., and Szpunar, K.K. (2017). Episodic Future Thinking: Mechanisms and Functions. Curr. Opin. Behav. Sci. *17*, 41–50.

Schapiro, A.C., Kustner, L.V., and Turk-Browne, N.B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. Curr. Biol. *22*, 1622–1627.

Schapiro, A.C., Turk-Browne, N.B., Botvinick, M.M., and Norman, K.A. (2017). Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. Phil Trans R Soc B *372*, 20160049.

Schlichting, M.L., Mumford, J.A., and Preston, A.R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. Nat. Commun. *6*, 8151.

Sheahan, H., Luyckx, F., Nelli, S., Teupe, C., and Summerfield, C. (2021). Neural state space alignment for magnitude generalization in humans and recurrent networks. Neuron *109*, 1214–1226.

Shohamy, D., and Wagner, A.D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. Neuron *60*, 378–389.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage *23*, S208–S219.

Summerfield, C., Luyckx, F., and Sheahan, H. (2020).

Structure Learning and the Posterior Parietal Cortex. Prog. Neurobiol. *184*, 101717.

Thavabalasingam, S., O'Neil, E.B., and Lee, A.C.H. (2018). Multivoxel pattern similarity suggests the integration of temporal duration in hippocampal event sequence representations. NeuroImage *178*, 136–146.

Thavabalasingam, S., O'Neil, E.B., Tay, J., Nestor, A., and Lee, A.C.H. (2019). Evidence for the incorporation of temporal duration information in human hippocampal long-term memory sequence representations. Proc. Natl. Acad. Sci. *116*, 6407–6414.

Tolman, E.C. (1948). Cognitive maps in rats and men. Psychol. Rev. *55*, 189–208.

Tompary, A., and Thompson-Schill, S.L. (2021). Semantic influences on episodic memory distortions. J. Exp. Psychol. Gen. Advance online publication.

Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J.J., Moser, M.-B., and Moser, E.I. (2018). Integrating time from experience in the lateral entorhinal cortex. Nature *561*, 57–62.

Umbach, G., Kantak, P., Jacobs, J., Kahana, M., Pfeiffer, B.E., Sperling, M., and Lega, B. (2020). Time cells in the human hippocampus and entorhinal cortex support episodic memory. Proc. Natl. Acad. Sci. *117*, 28463–28474.

Whittington, J.C.R., Muller, T.H., Mark, S., Chen, G., Barry, C., Burgess, N., and Behrens, T.E.J. (2020). The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation. Cell *183*, 1249-1263.e23.

Zacks, J.M. (2020). Event Perception and Memory. Annu. Rev. Psychol. *71*, 165–191.

Zeithamova, D., and Bowman, C.R. (2020). Generalization and the hippocampus: More than one story? Neurobiol. Learn. Mem. *175*, 107317.

Zeithamova, D., and Preston, A.R. (2010). Flexible Memories: Differential Roles for Medial Temporal Lobe and Prefrontal Cortex in Cross-Episode Binding. J. Neurosci. *30*, 14676–14684.

Zeithamova, D., Dominick, A.L., and Preston, A.R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. Neuron *75*, 168–179.

Zeithamova, D., Gelman, B.D., Frank, L., and Preston, A.R. (2018). Abstract Representation of Prospective Reward in the Hippocampus. J. Neurosci. *38*, 10093–10101.

Acknowledgements

The authors would like to thank Ignacio Polti for helpful discussions on the behavioral generalization bias, Iris M. Engelhard for making the data of the replication sample available, and Roland Benoit for helpful comments on a previous version of the manuscript. This work was supported by the Netherlands Organisation for Scientific Research (NWO-Vidi 452-12-009; NWO-MaGW 406-14-114), the European Research Council (ERC-CoG GEOCOG 724836) and the Max Planck Society. C.F.D.'s research is further supported by the Kavli Foundation, the Centre of Excellence scheme of the Research Council of Norway—Centre for Neural Computation (223262/F50), The Egil and Pauline Braathen and Fred Kavli Centre for Cortical Microcircuits, the Jebsen Centre for Alzheimer's Disease, and the National Infrastructure scheme of the Research Council of Norway—NORBRAIN (197467/F50). N.D.M. was funded by the Netherlands Organization for Scientific Research (NWO 453-15-005, awarded to Iris M. Engelhard) during the acquisition of the data used for the replication of the behavioral generalization effect.

Data & Code Availability

Analysis code and documentation is available on GitHub (https://jacbel.github.io/virtem_code/ovierview). Data to reproduce the statistical analyses reported in this manuscript will be made openly available upon publication.

Author Contributions

J.L.S.B., L.D, N.D.M. and C.F.D. conceived the experiment and analyses. N.D.M. and L.D. developed the tasks and acquired the data. L.D. preprocessed the MRI data. J.L.S.B. analyzed the data and wrote the manuscript with input from C.F.D. All authors discussed the results and contributed to the paper.