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[1,4,5]

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
5: Institute of Psychology – Wilhelm Wundt, Leipzig University, Leipzig, Germany
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 memories reflect temporal relations derived from mnemonic constructions, event order, or elapsing time, and whether these sequence representations 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 temporal relations based on constructed times. Temporal relations were 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 mnemonic construction and the generalization of relational knowledge combine in the hippocampus, consistent with the simulation of scenarios from episodic details and structural knowledge.

Introduction

Our memories are not veridical records, but constructions of our past[1]. When constructing scenarios of the past or future, we often combine specific episodic details with general, semantic knowledge[2–7]. 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[8,9]. 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[8,9]. Event representations in the hippocampal-entorhinal region carry information about sequence relationships[10,11], 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[12,13], 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[10]. For instance, learning sequences recruits the hippocampus and entorhinal cortex[14,15], and hippocampal activity increases at event boundaries delineating sequences[16,17]. Hippocampal multi-voxel patterns are sensitive to objects shown at learned sequence positions[18], and recent work suggests that the hippocampus incorporates the duration of intervals between elements in sequence representations[19,20]. Further, pattern correlations in the hippocampus and entorhinal cortex relate to memory for temporal relations[21–26].

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[21,27]. After relative to before learning, pattern similarity in the anterior hippocampus and the anterior-lateral entorhinal cortex elicited by event images reflected the sequence relationships between pairs of events. Events closer in time elicited more similar activity patterns relative to events separated by longer intervals, resulting in negative correlations between pattern similarity and temporal distances[21,27]. Within the entorhinal cortex, this effect was specific to the anterior-lateral subregion[27], consistent with the involvement of this area in precise temporal memory recall[28,29]. Negative correlations between pattern similarity and distances are in line with sequence representations akin to cognitive maps of space – positions separated by low distances share similar representations, whereas positions with high distances between them are represented less similarly, i.e. pattern similarity scales with distance.

However, whether event representations in the anterior hippocampus and anterior-lateral entorhinal cortex reflect temporal distances based on 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[30–32]. 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[33,34]. Slowly drifting activity patterns could be incorporated into event memories as temporal tags, providing a potential mechanism for temporal memory[35]. 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.

Mnemonic construction enables prospective cognition[2,5,36]. The hippocampal-entorhinal region integrates and recombines episodic details across experiences for future simulation, inferential reasoning and generalization[5,12,13,37–40]. 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[41–43]. Further, it combines separately learned associations, enabling inferences about shared associations[44–50]. 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[12,51]. Indeed, entorhinal activity patterns reflected structural similarities between choice options in a reinforcement learning task[52]. 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[53]. Work in rodents suggests that hippocampal representations of events in a sequence generalize across comparable experiences in a different environment[54]. Applying abstract structural knowledge enables adaptive behavior through the generalization of relations to novel situations[12,51]. Whether representations of temporal relations of events in a sequence are

constructed such that they generalize across sequences with a similar structure is unclear.

Knowledge about structural regularities and semantic associations closely interacts with episodic construction[4,7,38]. When estimating the size of studied images, participants' reconstructions were systematically distorted towards category averages[55,56]. For relatively small fruits like strawberries, participants tended

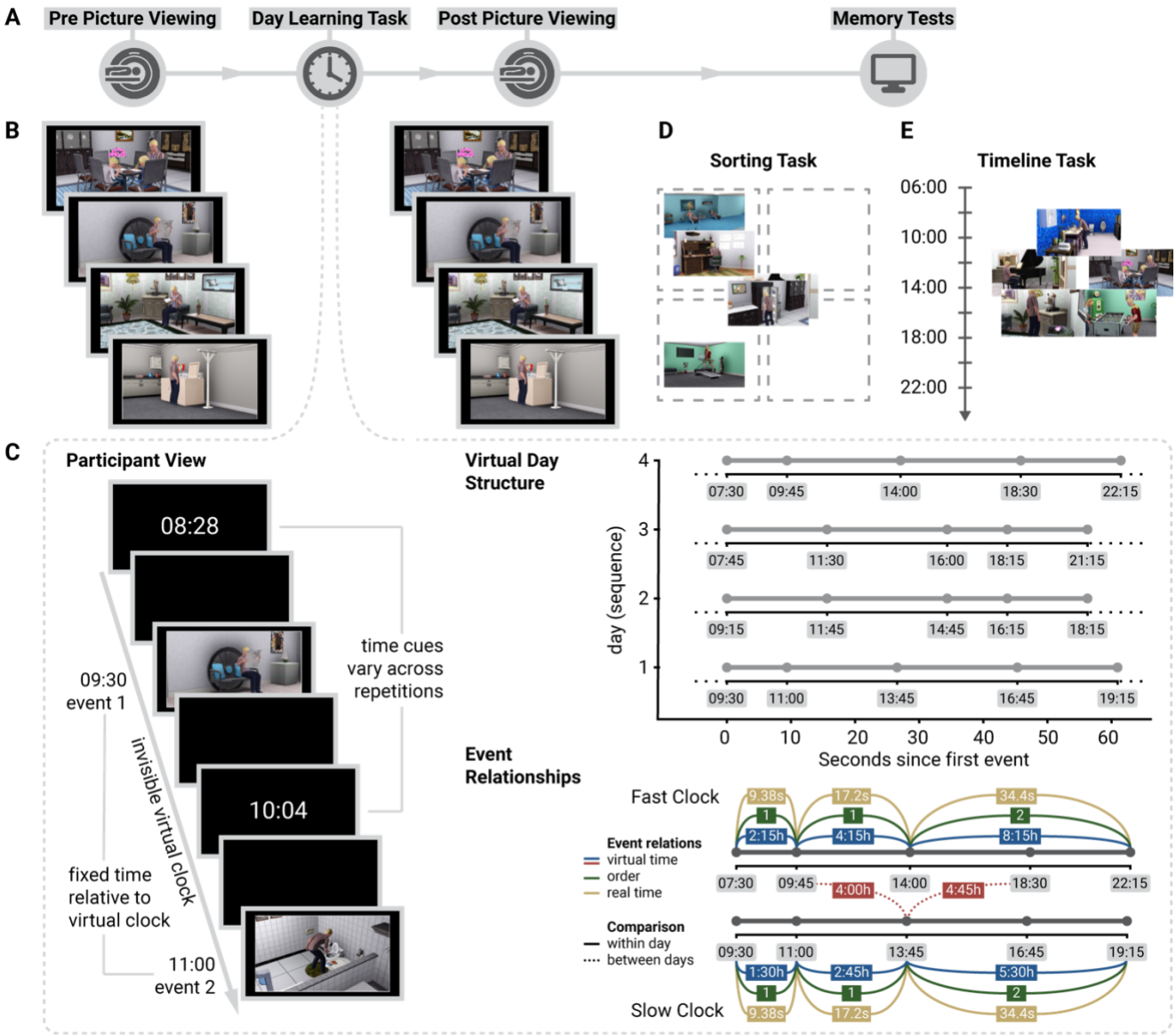


Figure 1. Experimental Design. A. Overview of the experiment. **B**. In the picture viewing tasks before and after learning, participants saw event images presented in the same random order and using identical stimulus timings. **C**. The day learning task took place in between the picture viewing tasks. 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 real 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.

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[55]. Consistent with the notion that learned event structures contribute to event cognition[57–59], external and semantic details are used to furnish past and future scenarios when few episodic details are generated[60,61]. 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[62]. These findings suggest that abstract knowledge about general patterns could systematically distort constructions of specific 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 day you arrived late towards this time.

Here, we combine functional magnetic resonance imaging (fMRI) with a sequence learning task requiring the 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 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 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). 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). Event images with minimal or no indication of time of day (Supplemental Figure 1) were randomly assigned to sequences and sequence positions for each participant. Thus it was impossible to infer specific event times or sequence memberships from the stimuli. 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 experience of objectively elapsing time (real time) with the virtual 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 real time to test whether mnemonically 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. The distribution of sorting errors did not differ from uniformity across sequence positions (χ^2=2.55, p=0.635). 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). The accuracy of constructed virtual times differed between sequences ($F_{3,81}$=5.86, p<0.001), but not as a function of virtual clock speed (t_{27}=-0.82, p=0.423, Supplemental Figure 3AB). We did not observe an across-subject relationship between the number of sorting errors and mean absolute errors in the timeline task (Supplemental Figure 3CD). To test

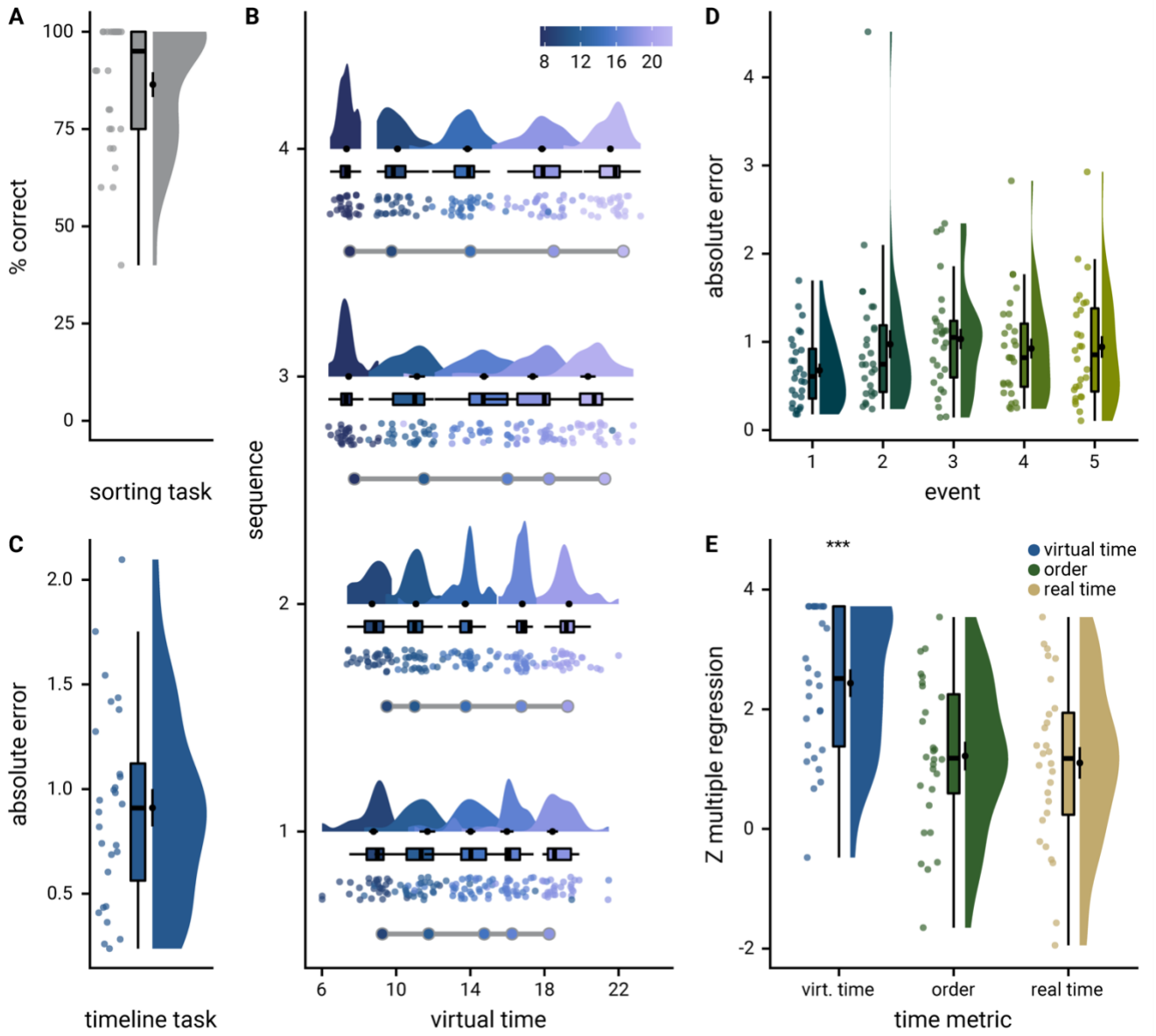


Figure 2. Participants learn the temporal structure of the sequences relative to the virtual clock. A. Plot shows the percentage of correctly sorted event images in the sorting task. **B.** Constructed event times were assessed in the timeline task. Responses are shown separately for the five events (color coded according to true virtual time) of each sequence (rows). Colored circles with gray outline 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. **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

whether the 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; summary statistics: t_{27}=10.62, p<0.001, d=1.95, 95% CI [1.38, 2.70]; mixed model: $\chi^2(1)$=115.95, p<0.001, Supplemental Figure 4AB, Supplemental Table 1). Together, these findings demonstrate that participants formed precise memories of the different sequences and accurately constructed event times.

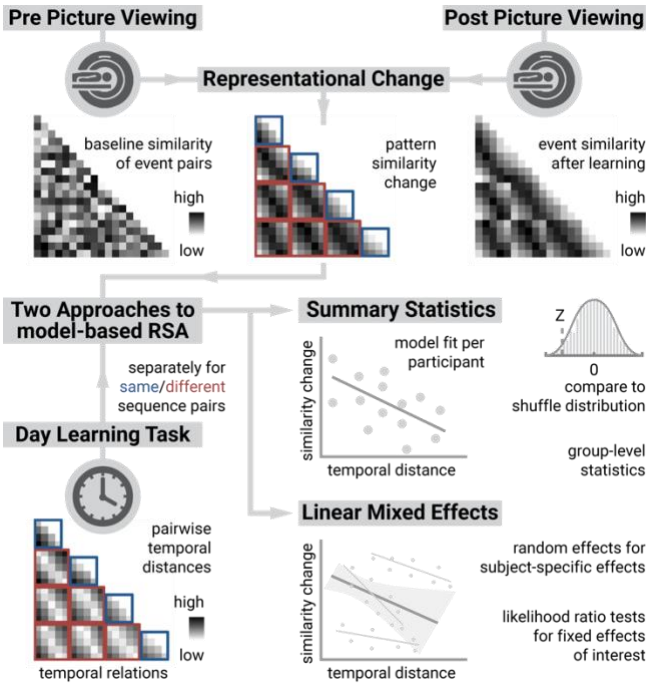


Figure 3. Representational Similarity Analysis Logic. We quantified the representational similarity of all event pairs before and after learning. Representational change was defined by subtracting pre-learning from post-learning pattern similarity (top row). Using two approaches to model-based representational 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.

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). The converging results of these analyses demonstrate that our findings do not depend on the specific statistical methods employed. We centered our analyses on the anterior hippocampus and the anterior-lateral entorhinal cortex (see Methods) based on our previous work implicating these regions in representing sequence relations[21,27].

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. Surprisingly, 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, d=0.56, 95% CI [0.18, 1.00]; 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 Figure 4CD, Supplemental Table 2). This effect was further characterized by higher pattern similarity for event pairs separated by longer temporal distances than for pairs separated by shorter intervals (Figure 4C, t_{27}=2.48, p=0.020, d=0.64, 95% CI [0.08, 0.87]). In contrast to our previous work[21], where we observed negative correlations of pattern similarity and temporal distances, participants learned multiple

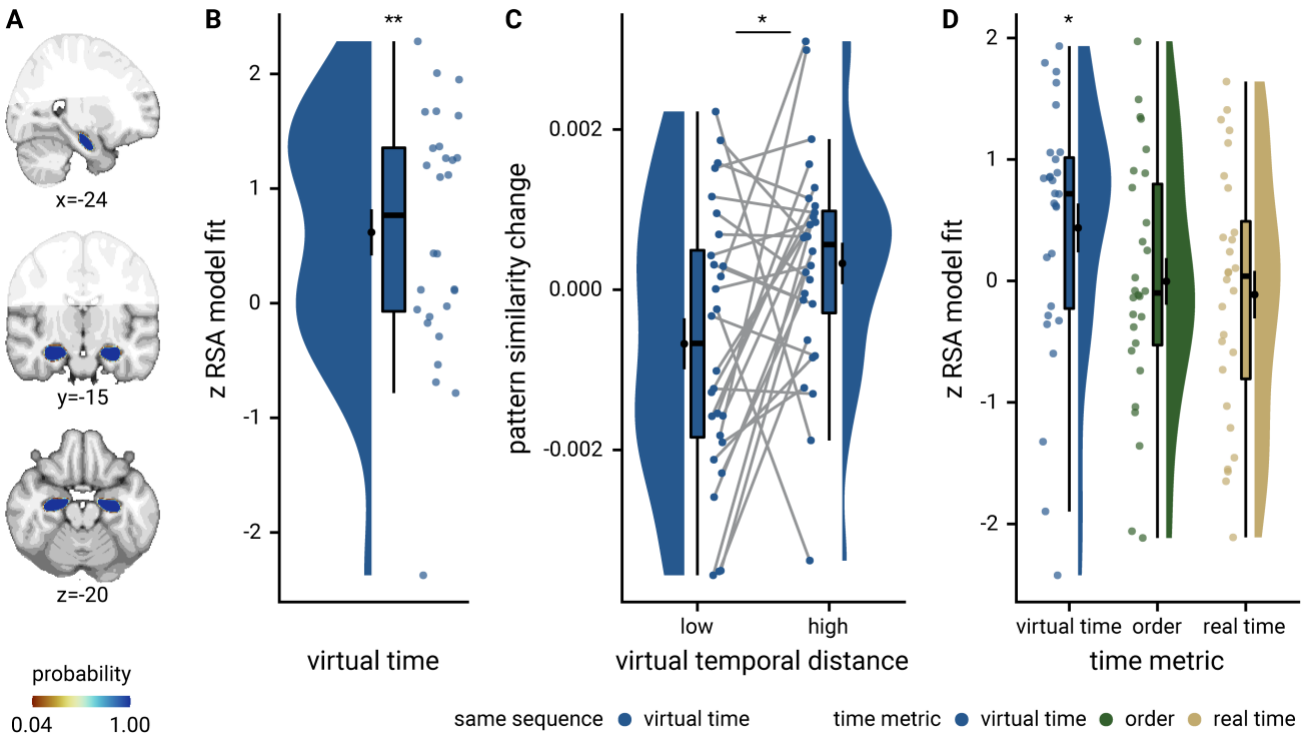


Figure 4. Sequence representations in anterior hippocampus reflect constructed event times. A. The anterior hippocampus region of interest is displayed on the MNI template with voxels outside the field of view shown in lighter shades of gray. Color code denotes probability of a voxel to be included in 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**. To illustrate the effect shown in **B**, average pattern similarity change values are shown for same-sequence event pairs that are separated by low and high temporal distances based on a median split. **D**. 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. **B-D**. 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

sequences in this study. They might have formed strong associations of same-sequence events on top of inferring each event's virtual time, potentially altering how temporal distances affected hippocampal pattern similarity (see Discussion). The effect of virtual temporal distances on pattern similarity changes remained significant when competing for variance with a control predictor accounting for comparisons of the first and last event of each sequence (Supplemental Figure 5A-C; summary statistics: t_{27}=2.25, p=0.034, d=0.41, 95% CI [0.04, 0.82]; mixed model: $\chi^2(1)$=5.36, p=0.021, Supplemental Table 3). Thus, the relationship of hippocampal event representations and temporal distances is not exclusively driven by associations of the events marking the transitions between sequences.

Having established that hippocampal pattern similarity changes relate to temporal distances, we next assessed whether this effect was driven by virtual event times beyond sequence order and real time. We thus included the two additional time metrics as control predictors in the model. Virtual temporal distances significantly predicted pattern similarity changes even when controlling for the effects of event order and real time in seconds (Figure 4D; summary statistics: t_{27}=2.18, p=0.040, d=0.40, 95% CI [0.02, 0.81]; mixed model: $\chi^2(1)$=5.92, p=0.015, Supplemental Figure 4EF, Supplemental Table 4). Further, 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 5D; t_{27}=2.23, p=0.034, d=0.41, 95% CI [0.03, 0.82]),

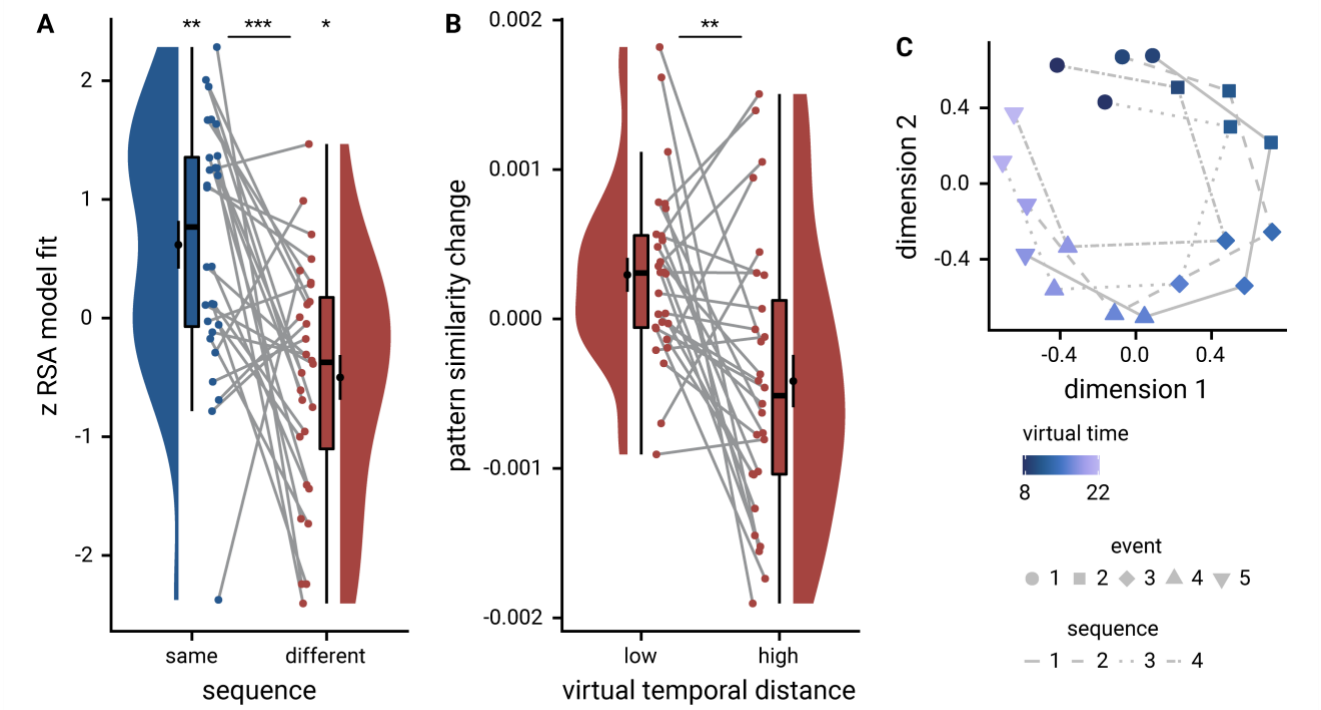


Figure 5. The anterior hippocampus generalizes temporal relations across sequences. A. Z-values show results of participant-specific linear models quantifying the effect of virtual time for event pairs from the same sequence (blue, as in Figure 4B) and from different sequences (red). Temporal distance is negatively related to hippocampal representational change for event pairs from different sequences. See Supplemental Figure 4EF 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**. To illustrate the effect shown in **A**, average pattern similarity change values are shown for across--sequence event pairs that are separated by low and high temporal distances based on a median split. **C**. Multidimensional scaling results show low-dimensional embedding of the event sequences. Shapes indicate event order, color shows virtual times of events. The different lines connect the events belonging to the four sequences for illustration. *** p≤0.001; ** p<0.01; * p<0.05

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 mnemonically constructed event times.

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, summary statistics t_{27}=-2.65, p=0.013, d=-0.49, 95% CI [-0.91, -0.10]; mixed model: $\chi^2(1)$=6.01, p=0.014, Supplemental Figure 4GH, Supplemental Table 5; α=0.025, corrected for separate tests of events of the same and different sequences). This indicates that hippocampal representations of events from different sequences changed systematically to reflect generalized temporal relations. 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 (Figure 5B, t_{27}=-3.26, p=0.002, d=-0.89, 95% CI [-1.03, -0.21]). Virtual time was a significant predictor of hippocampal pattern similarity change for events from different sequences when competing for variance with order and real time (Supplemental Figure 6A-C; summary statistics: t_{26}=-2.62, p=0.015, d=-0.49, 95% CI [-0.92, -0.10], mixed model: $\chi^2(1)$=4.48, p=0.034, Supplemental Table 6; one outlier excluded). The relationship of temporal distances and representational change differed significantly between events from the same or different sequences (Figure 5A, summary statistics: paired t-test t_{27}=3.71, p=0.001, d=1.05, 95% CI [0.29, 1.13]; mixed model: interaction of sequence

membership with virtual time $\chi^2(1)=14.37$, $p<0.001$, Supplemental Figure 4IJ, Supplemental Table 7). Similar interactions of sequence membership with order ($\chi^2(1)=9.98$, $p=0.002$) and real time ($\chi^2(1)=9.27$, $p=0.002$) were observed, but, crucially, the interaction of sequence membership and virtual time remained significant when including interactions of sequence membership with order and real time in the model ($\chi^2(1)=8.57$, $p=0.003$, Supplemental Table 8). Thus, the way knowledge about virtual temporal relations was represented in the hippocampus depended on whether events belonged to the same sequence or not.

To explore how event sequences may be arranged in a low-dimensional representational space to give rise to the effects described above, we generated a distance matrix from the mixed effects model fitted to hippocampal pattern similarity change and subjected it to non-metric multidimensional scaling (see Methods, Supplemental Figure 6D). The resulting configuration in two dimensions (Figure 5C), chosen for intuitive visualization, exhibited a c-shaped pattern for each sequence. Similar representational geometries have previously been described in parietal cortex[63–65]. Events occurring at similar virtual times occupy similar locations, in line with high pattern similarity for events from different sequences that are separated by low temporal distances. Thus, the generalization across sequences results in a comparable configuration for each sequence. While the observed configuration resulted in stress values significantly lower than those obtained in a permutation test (see Methods; $z=-3.5$, $p=0.001$, Supplemental Figure 6E), the high representational distances between temporally close events from the same sequence are not perfectly captured by the c-shaped arrangement (Supplemental Figure 6FG). More than the two dimensions chosen for visualization would likely better capture the complex representational structure of the sequences.

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; summary statistics: $t_{27}=-2.31$, $p=0.029$, $d=-0.42$, 95% CI [-0.84, -0.05]; mixed model: $\chi^2(1)=4.39$, $p=0.036$, Supplemental Figure 4KL, Supplemental Table 9; see Supplemental Figure 7A for separate analyses of events from the same and from different sequences). In line with our previous work[27],

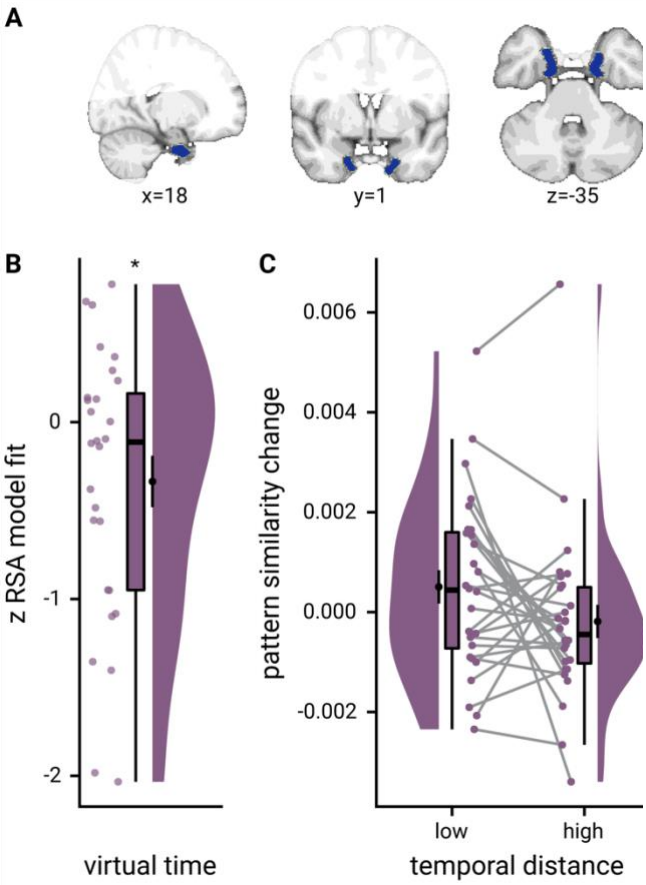


Figure 6. The anterior-lateral entorhinal cortex uses a shared representational format for relations of events from the same and different sequences. A. The anterior-lateral entorhinal cortex region of interest is displayed on the MNI template with voxels outside the field of view shown in lighter shades 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. To illustrate the effect in B, raw pattern similarity change in the anterior-lateral entorhinal cortex was averaged for events separated by low and high temporal distances based on a median split. * $p<0.05$

events close together in time became more similar than those separated by longer temporal intervals (Figure 6C). The relationship of virtual temporal distances and entorhinal pattern similarity change was not statistically significant when competing for variance with distances based on order and real time (Supplemental Figure 7B-D; summary statistics: t_{27}=-0.7, p=0.495, d=-0.13, 95% CI [-0.51, 0.25], mixed model: $\chi^2(1)$=1.18, p=0.278, Supplemental Table 10).

We further corroborated that the temporal structure of the sequences was represented differently between the 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, η^2=0.08, main effect of region $F_{1,27}$=3.10, p=0.086, η^2=0.02, main effect of sequence $F_{1,27}$=7.41, p=0.012, η^2=0.08; mixed model: three-way

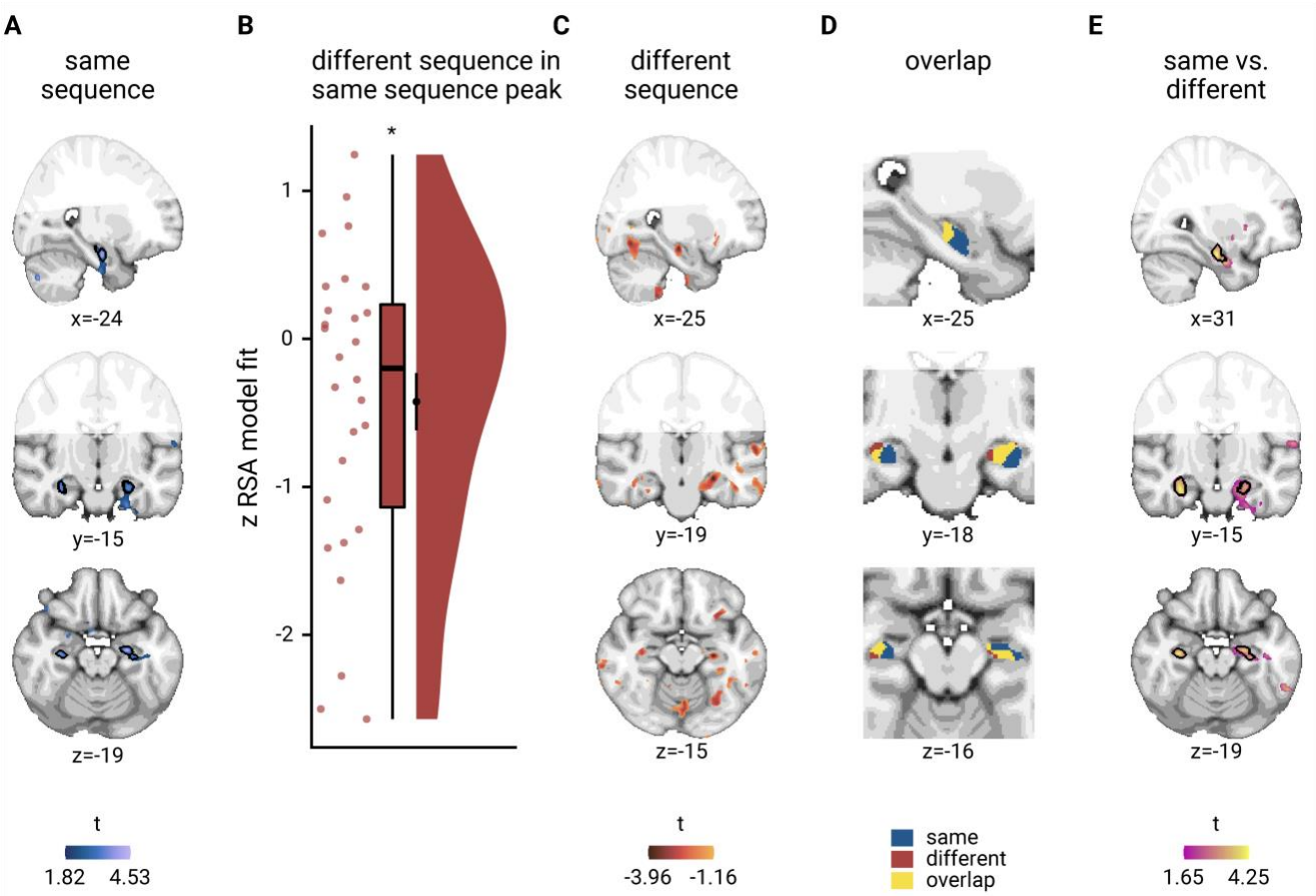


Figure 7. Overlapping representations of within- and across-sequence relations. A. Searchlight analysis results show a positive relationship between representational change and virtual temporal distances for event pairs from the same sequence in the bilateral anterior hippocampus. Statistical image is thresholded at $p_{uncorrected}$<0.01; voxels within black outline are significant after correction for multiple comparisons using small volume correction. **B**. 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. 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**. 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. **D**. 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. **E**. 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**, **C-E**. Results are shown on the MNI template with voxels outside the field of view displayed in lighter shades of gray. See Supplemental Figure 9 for additional exploratory results. * p<0.05

interaction between virtual time, sequence membership and region of interest $\chi^2(1)=6.31$, $p=0.012$, Supplemental Table 11; see Supplemental Figure 8 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 12). 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; summary statistics $t_{27}=-2.19$, $p=0.036$, $d=-0.40$, 95% CI [-0.81, -0.03]; mixed model: $\chi^2(1)=4.13$, $p=0.042$, Supplemental Figure 4MN, Supplemental Table 13), 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 7CD), 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 14). 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 anterior hippocampus (Figure 7E, peak voxel MNI x=31, y=-16, z=-21; t=4.25, $p_{svc}=0.007$, Supplemental Table 15). 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 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 8BC, Supplemental Figure 10A; summary statistics: $t_{27}=5.32$, $p<0.001$, $d=0.98$, 95% CI [0.55, 1.48]; mixed model: $\chi^2(1)=17.90$, $p<0.001$, Supplemental Figure 4OP, Supplemental Table 16). This demonstrates that structural knowledge about the sequences biased the construction of event times. The constructed virtual time of an event tended to be overestimated when the events occupying the same sequence position in the other sequences took place late relative to the event in question, and vice versa when the other events occurred relatively early. In an independent group of participants[66], we replicated this generalization bias (Figure 8D, Supplemental Figure 10B; summary statistics: $t_{45}=11.30$, $p<0.001$, $d=1.64$,

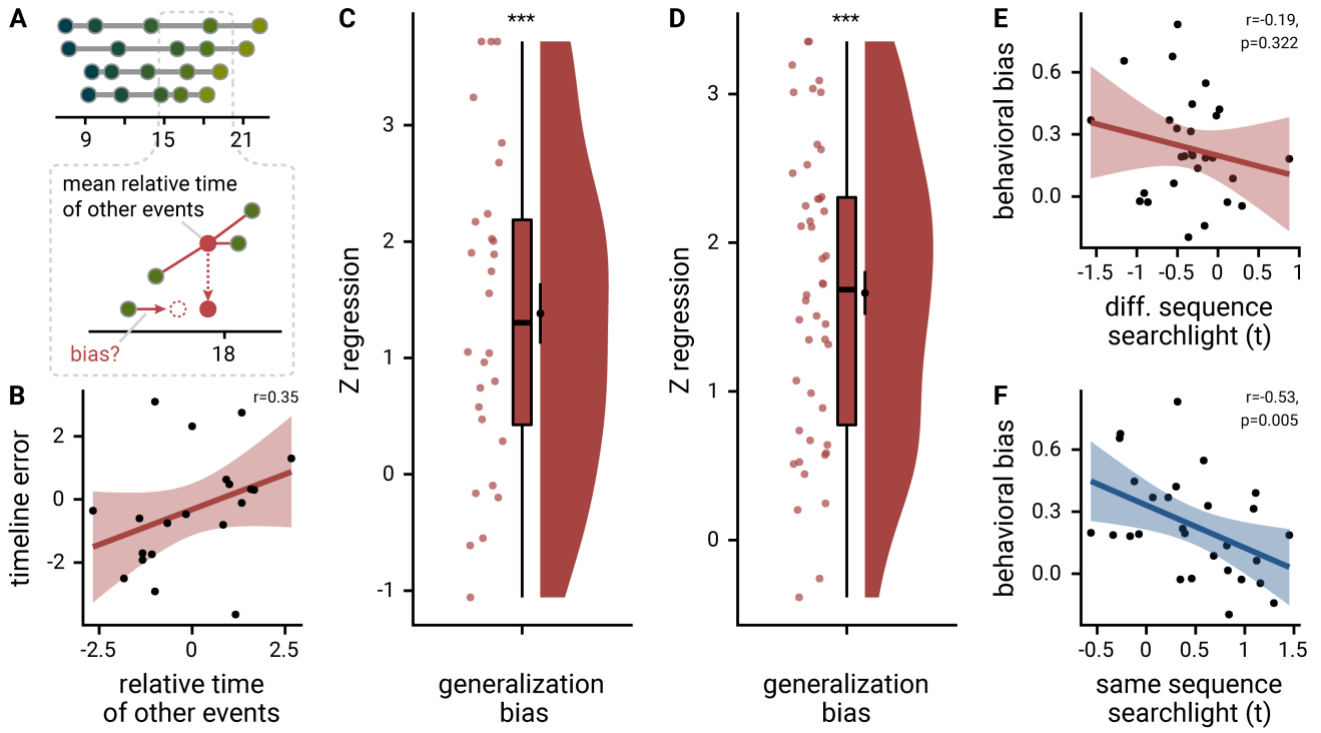


Figure 8. Structural knowledge biases construction of event times. A. The generalization bias quantifies the influence of structural knowledge on the construction of individual event times. For each event, the mean time of events at the same sequence position in the other sequences was calculated to 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 10 for the entire sample. Correlation coefficient is based on Pearson correlation. **C.** The relative time of events from other sequences predicted signed event time construction errors as measured in the timeline task. Positive values indicate that 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. 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**. The generalization bias in event time construction through structural knowledge was replicated in an independent sample (n=46) based on Montijn et al.[66]. Data shown as in **B**. **E.** 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). **F.** 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 **E** and **F** are based on Spearman correlation.

95% CI [1.23, 2.13]; mixed model: $\chi^2(1)= 53.74$, $p<0.001$, Supplemental Figure 4QR, Supplemental Table 17), 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 same-sequence events in the searchlight analysis and the strength of the behavioral generalization bias (Figure 8EF, 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.

We further explored whether participants made systematic errors in the sorting task that might point towards generalization across sequences. Specifically, we searched for swap errors where participants interchanged events between sequences that occupied the same sequence position. Indeed, 57.5%±34.3% (mean±S.D) of sorting errors were swap errors and 12 of the 14 participants who made sorting errors also made swap errors (Supplemental Figure 3EF, mean±S.D of 3.1±2.1 swap errors per participant with sorting errors). The proportion of swap errors in our sample was larger than expected from random sorting errors ($z=5.07$, $p<0.001$, Supplemental Figure 3G), indicating that participants systematically swapped events belonging to the same position between sequences. While we did not observe statistically significant relationships between swap errors and the generalization bias (Supplemental Figure 3HI), the prevalence of these errors is compatible with the view that participants generalized across events occupying the same sequence position.

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 specific representations of temporal relations of the events in a sequence that mirror 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 entorhinal cortex scaled with temporal distances irrespective of sequence membership. The behavioral data demonstrate that the construction of specific event times is biased by structural knowledge abstracted from different sequences.

In our paradigm, participants mentally constructed the times of events relative to a hidden virtual clock. To do so, they needed to combine their experience of passing real time with infrequent cues about the current virtual time. Thus, real time was critical for the successful construction of event times, despite not being cued explicitly. Participants' responses in a memory test and the similarity structure of hippocampal multi-voxel patterns were explained by virtual event times beyond the effects of real time and sequence order, showing that sequence representations reflect mnemonically constructed time. Recent work demonstrated the scaling of time cell representations to different real time intervals in the rodent hippocampus[67]. Temporal scaling of hippocampal representations could potentially underlie our observation that temporal distances in virtual time are related to the similarity of event representations even when accounting for the effects of real time and order. This finding highlights that the anterior hippocampus maps relational knowledge derived from mnemonic constructions.

The hippocampus constructed an integrated representation that generalized temporal relations across sequences. Multi-voxel patterns of events taking place at similar virtual times, but in different sequences, were more similar than those of events occurring at different points in time. Thus, representations of events from different sequences changed systematically to reflect generalized temporal distances. Speculatively, this effect could be related to the observation that, in mice trained to run a number of laps on a maze to obtain rewards, lap-specific firing patterns in the hippocampus generalize across sequences of laps on geometrically distinct mazes[54]. While it is possible that the first and last events of the sequences are particularly important to sequence processing, our data show that virtual time explained representational changes when competing for variance with order and real time also for events from different sequences. This makes it unlikely that the hippocampal generalization effect was driven exclusively by events at the first or last sequence position. The generalization of temporal distances across sequences in the hippocampus is in line with the contribution of constructive mnemonic processes to flexible cognition via the recombination of elements across experiences and statistical learning[13,40,43,46,48,49,68,69]. More generally, it is consistent with the role of the hippocampus in forming cognitive maps of relational structures and in generalizing structural knowledge to novel situations[12,38,51,53,57,70,71].

Structural knowledge and mnemonic construction are intertwined. 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[55,56], when estimating brief time intervals[72,73], or when discriminating the order of previously presented stimuli[74]. Likewise, prior knowledge can distort memories for short narratives[75], spatial associations[76] and temporal positions[62]. 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[77]. Further, recombining information across episodes for associative inference can induce false memories for contextual details[68,78], 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[79], these findings highlight that structural knowledge and mnemonic construction are interwoven. More broadly, abstract semantic or schematic knowledge may provide a scaffold for the recall of episodic details[4,7,38,80,81]. Our findings show that structural knowledge not only facilitates, but also biases constructive memory.

The way temporal relations shaped hippocampal multi-voxel pattern similarity differed between pairs of events from the same and different sequences. We observed positive correlations between temporal distances and hippocampal representational change, which were characterized by relatively decreased pattern similarity for nearby compared to increased pattern similarity for more distant events from the same sequence. One possible explanation for the surprising direction of this effect could be that, compared to our previous work where participants encountered only one sequence[21], participants relied more on associative encoding strategies when learning multiple sequences in the present experiment. Possibly, the need to link events belonging to the same sequence altered how pattern similarity changes relate to temporal distances for these same-sequence events. In line with this interpretation, prior work has shown that the relationship of hippocampal pattern similarity and temporal memory can depend on factors like the use of associative encoding strategies and the presence of event boundaries marking switches between sequences of images from the same category[22,82,24]. Successful recency discrimination was associated with more similar hippocampal representations during encoding when participants were encouraged to use associative strategies to encode the order of image sequences from two alternating visual categories[22]. A different study found more dissimilar hippocampal representations for stimuli whose order was later remembered correctly[24]. Thus, the formation of associations between same-sequence events could explain why correlations of pattern similarity change were, in contrast to our previous work[21], positive. A second possible interpretation of this effect is based on observations that the hippocampus differentiates similar episodes[47,83–86]. Hippocampal differentiation could explain the relative decrease of pattern similarity for temporally close events from the same sequence. However, the generalization across sequences does not directly follow from a differentiation account.

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[39,87]. 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[88] and memory integration across episodes[47] 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[18,89]. Likewise, sequence positions can be decoded from magnetoencephalographic responses elicited by visual stimuli presented in scrambled order[90]. In line with the suggestion that the posterior parietal cortex supports generalization by projecting stimuli onto a low-dimensional manifold[91], neural magnitude representations that generalize across task contexts have been observed using EEG[63,92]. 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[93], 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 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 the abstraction of structural knowledge and the generalization across episodes. General time patterns abstracted from other sequences systematically influence the construction of specific event times, demonstrating that constructions of specific scenarios build 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)[94] and R (version 3.6.1)[95].

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 to participants as "virtual days". 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 experience of passing real time and mnemonic construction to infer the times of events.

Specifically, 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. Participants received no cues about elapsing real time, but had to use their experience of passing real time between virtual time cues to infer the event times relative to the hidden virtual clock. 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). Correlations between the linearly increasing time metrics are inevitably high (Pearson correlation of virtual time with order r=0.969 and virtual time with real time r=0.975). Still this manipulation allowed us to determine using multiple regression whether virtual time explained constructed event times when competing for variance with real elapsed time and event order and whether hippocampal pattern similarity changes related to temporal distances in virtual time beyond ordinal distances and real time distances. Regression models including collinear predictor variables do not result in biased parameter estimates[96,97].

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

Our previous work demonstrates representations reflecting the temporal relations of events from one sequence in the anterior hippocampus[21] and the anterior-lateral entorhinal cortex[27]. More generally, these regions have been implicated in temporal coding and memory (for review, see[10]). Further, the hippocampus has been linked to inferential reasoning and generalization[46,48,49,51,53]. We thus focused our analyses on these regions. 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.[98]. 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.[99]. 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[100] in Figure 2A.

In an exploratory analysis, we searched for systematic errors in the sorting task. Specifically, we looked for swap errors where participants interchanged events occurring at the same position between two or more sequences. We used a χ^2-test to assess whether the number of

swap errors deviated from uniformity across sequence positions. To test whether participants made more swap errors than expected from chance we ran a permutation test where we introduced sorting errors for randomly selected events. For each of 10 000 iterations, we generated a surrogate sample of sorting results with the number of randomly introduced sorting errors matching the number of errors made by the different participants in our sample. We then quantified the proportion of swap errors across this surrogate sample. This resulted in a distribution of the proportion of swap errors that would be expected from random sorting errors. We assessed how many permutations yielded proportions of swap errors larger or equal to the proportion of swap errors observed in the fMRI sample to compute a p-value and further quantified a z-value as the difference between the observed swap error proportion and the mean of the chance distribution divided by the standard deviation of the chance distribution. We tested whether the number of swap errors was related to absolute errors in the timeline task (see below) using Spearman's correlation and a t-test for independent samples.

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), the four sequences (Supplemental Figure 3A) and as a function of virtual clock speed (Supplemental Figure 3B). 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 (two-sided; α=0.05; 10,000 random sign-flips, Figure 2E). As a measure of effect size, we report Cohen's d with Hedges' correction and its 95% confidence interval as computed using the effsize-package[101].

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[102], 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 (Supplemental Figure 4A) and marginal effects (Supplemental Figure 4B) estimated using the ggeffects package[103]. To assess the statistical significance (α=0.05) 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 10A) 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 (Supplemental Figure 4OP, Supplemental Table 16).

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[66]. 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.[66]. The data from this independent sample are shown in Figure 8D and Supplemental Figure 10B.

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)[104] 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. Time metrics were z-scored within each participant prior to analysis. 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. To implement these tests, we employed two approaches to model-based RSA that are described in detail below. We used a summary statistics approach, which uses permutation-based procedures on the subject-level as well as on the group-level, in line with recommendations for the analysis of multi-voxel patterns [105]. We also implemented our statistical analyses using linear mixed effects models, which capture within-subject dependencies using random effects while estimating the fixed of interest on all data points. Mixed effects models are well-suited to test more complex interactions. The fact that the results of the two analysis approaches converge demonstrates that our findings are robust to the specific statistical technique. We used an α-level of 0.05 for both approaches because they are not independent as they are implemented on the same data and test the same hypotheses.

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[106]. 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 (two-sided tests; α=0.05 unless stated otherwise). We report Cohen's d with Hedges' correction and its 95% confidence interval as computed using the effsize-package for R. For paired tests, Cohen's d was calculated using pooled standard deviations and confidence intervals are based on the non-central t-distribution. Permutation-based repeated measures ANOVAs were carried out using the permuco-package[107] and we report generalized η^2 as effect sizes computed using the afex-package[108].

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[109]. 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.[102], 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[102,110]. The mixed effects models were fitted using maximum likelihood estimation.

We assessed the statistical significance of fixed effects of interest using likelihood ratio tests (α=0.05). 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[103].

Multidimensional Scaling

We aimed to explore how hippocampal event representations of the different sequences could be embedded in a low-dimensional representational space to give rise to the positive and negative correlations of pattern similarity change and temporal distances for same-sequence and different-sequence events, respectively. For each pair of events, we generated an expected similarity value (Supplemental Figure 6D) using the fixed effects of the mixed model fitted to hippocampal pattern similarity that captures the interaction between virtual temporal distances and sequence membership (c.f. Figure 5, Supplemental Figure 4IJ, and Supplemental Table 7). Using the predict-method implemented in the lme4-package[109], we generated model-derived similarity values for all event pairs given their temporal distances and sequence membership. We chose this approach over the raw pattern similarity values to obtain less noisy estimates of the pairwise distances. Using the smacof-package[111], the model-predicted similarities were converted to distances and the resulting distance matrix (Supplemental Figure 6D) was subjected to non-metric multidimensional scaling using two dimensions. We chose two dimensions to be able to intuitively visualize the results. Because MDS is sensitive to starting values, we ran multidimensional scaling 1000 times with random initial configurations and visualized the resulting configuration with the lowest stress value. Basing this analysis on the model-derived similarities assumes the same relationship of virtual temporal distances for all event pairs from different sequences, but we would like to note that not all solutions we observed, in particular those with higher stress values, resulted in parallel configurations for the four sequences.

We tested the stress value of the resulting configuration against a surrogate distribution of stress values obtained from permuting the input distances on each of 1000 iterations. Using the mean and standard deviation of the resulting null distribution, we obtained a z-value as a test statistic and report the proportion of stress values in the null distribution that were equal to or smaller than the observed stress value (Supplemental Figure 6E). Additionally, we contrasted the distances between pairs of events in the resulting configuration between distances separated by high or low (median split) input distances using a t-test for independent samples (Supplemental Figure 6F). Using a Spearman correlation, we quantified the relationship of the input distances and the distances in the resulting configuration (Supplemental Figure 6G).

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 9 and clusters with a

minimum extent of 30 voxels are listed in Supplemental Tables 12, 14 and 15.

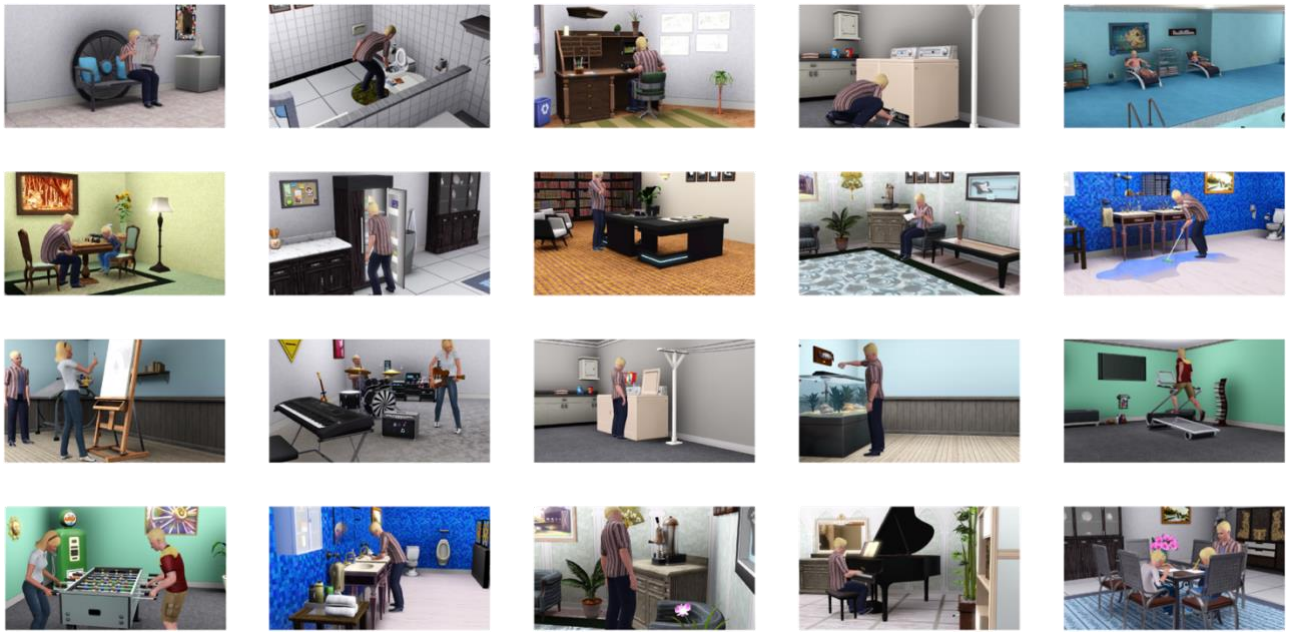
To test whether within- and across-sequence representations overlap, we defined an ROI based on the within-sequence searchlight analysis. Specifically, voxels belonging to the cluster around the peak voxel, thresholded at $p<0.01$ uncorrected within our small volume correction mask, were included. The analysis of representational change was then carried out as described for the other ROIs above. The results observed using a threshold of $p<0.001$ were not statistically different from those obtained with a threshold of $p<0.01$ ($t_{27}=-0.95$, $p=0.338$; test against 0 using the ROI resulting from the $p<0.001$ threshold: $t_{27}=-1.98$, $p=0.056$).

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 chose this approach because the searchlight analyses provide greater spatial precision than anatomically defined region of interest masks. 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).

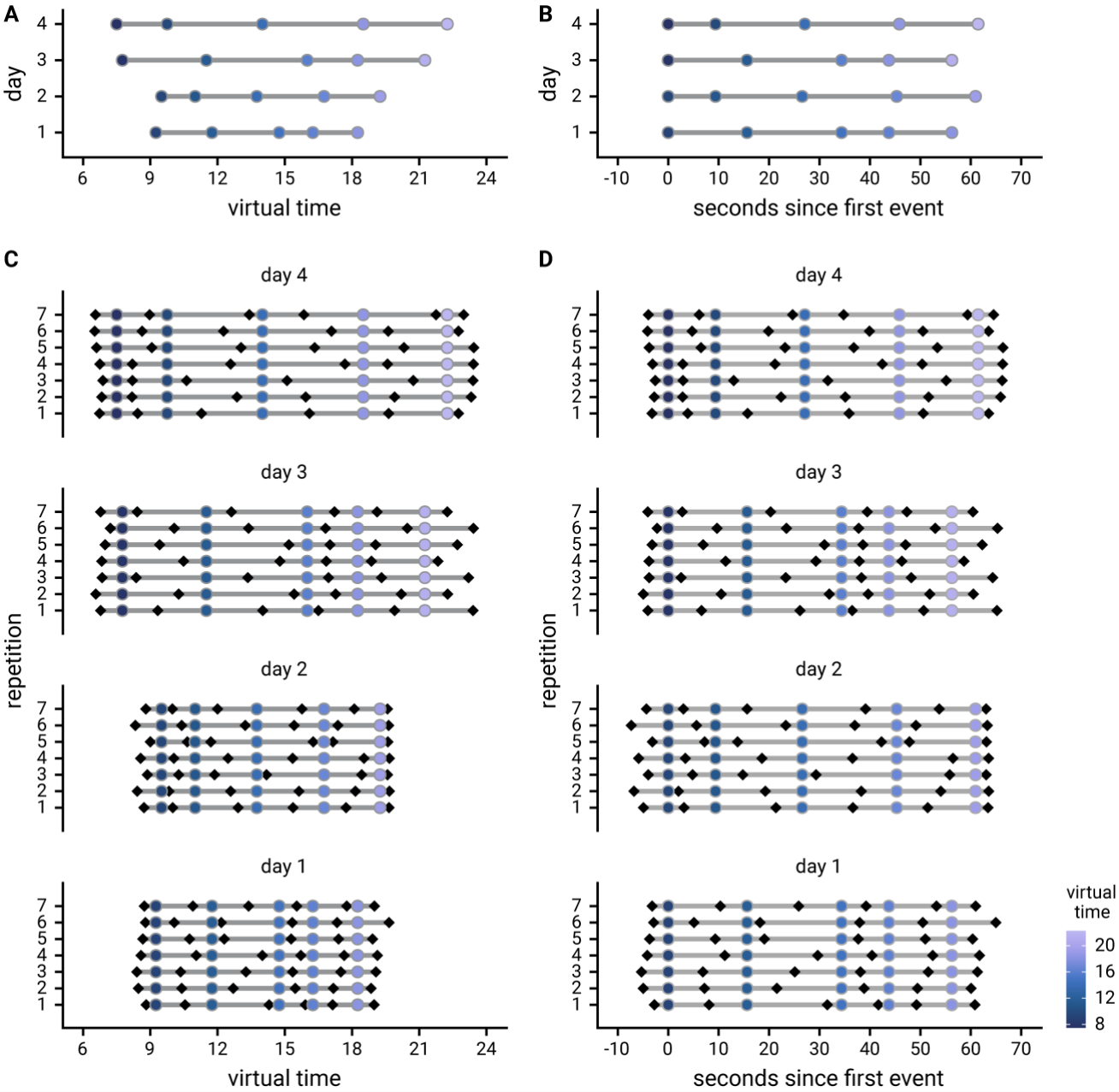
Supplemental Figures

Supplemental Figure 1



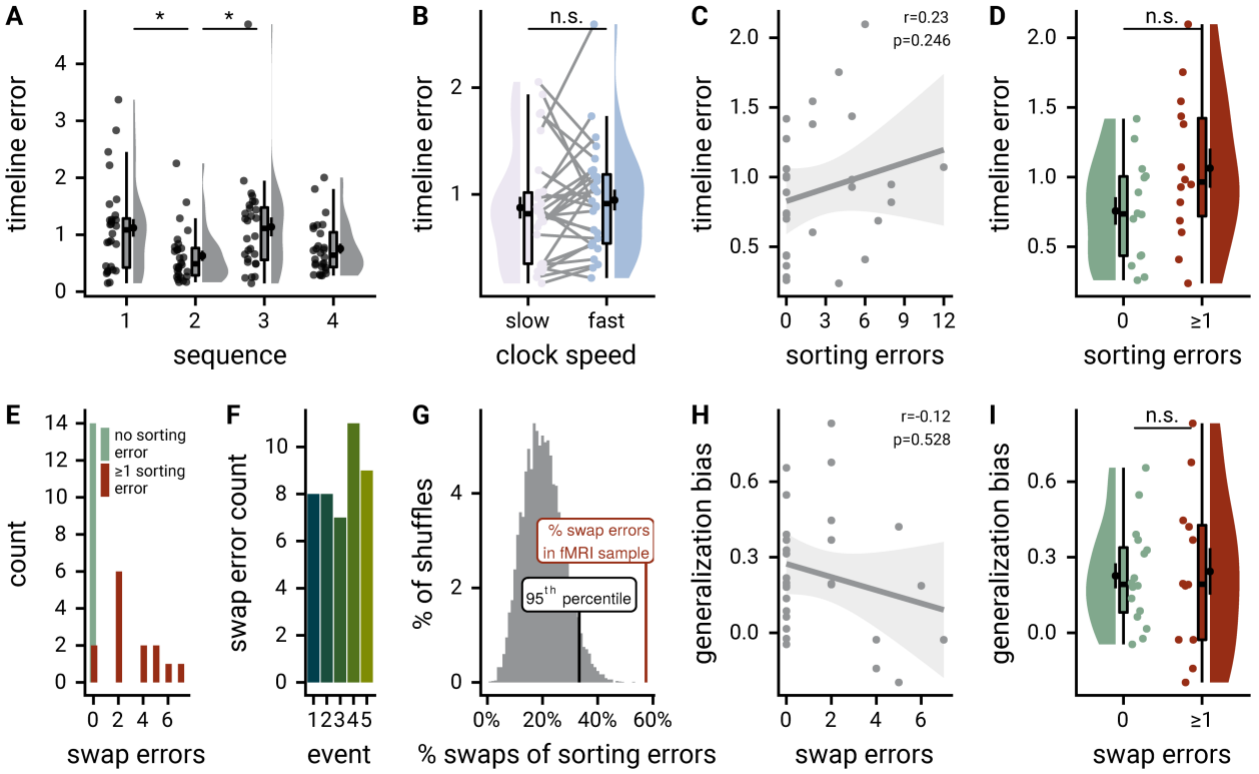
Supplemental Figure 1. Overview of the event images used as stimuli. All scenes were devoid of windows to exclude diurnal cues, such as shadows or light color, and were selected so they would be plausible at any time of day. For each participant, event images were randomly allocated to sequences and 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



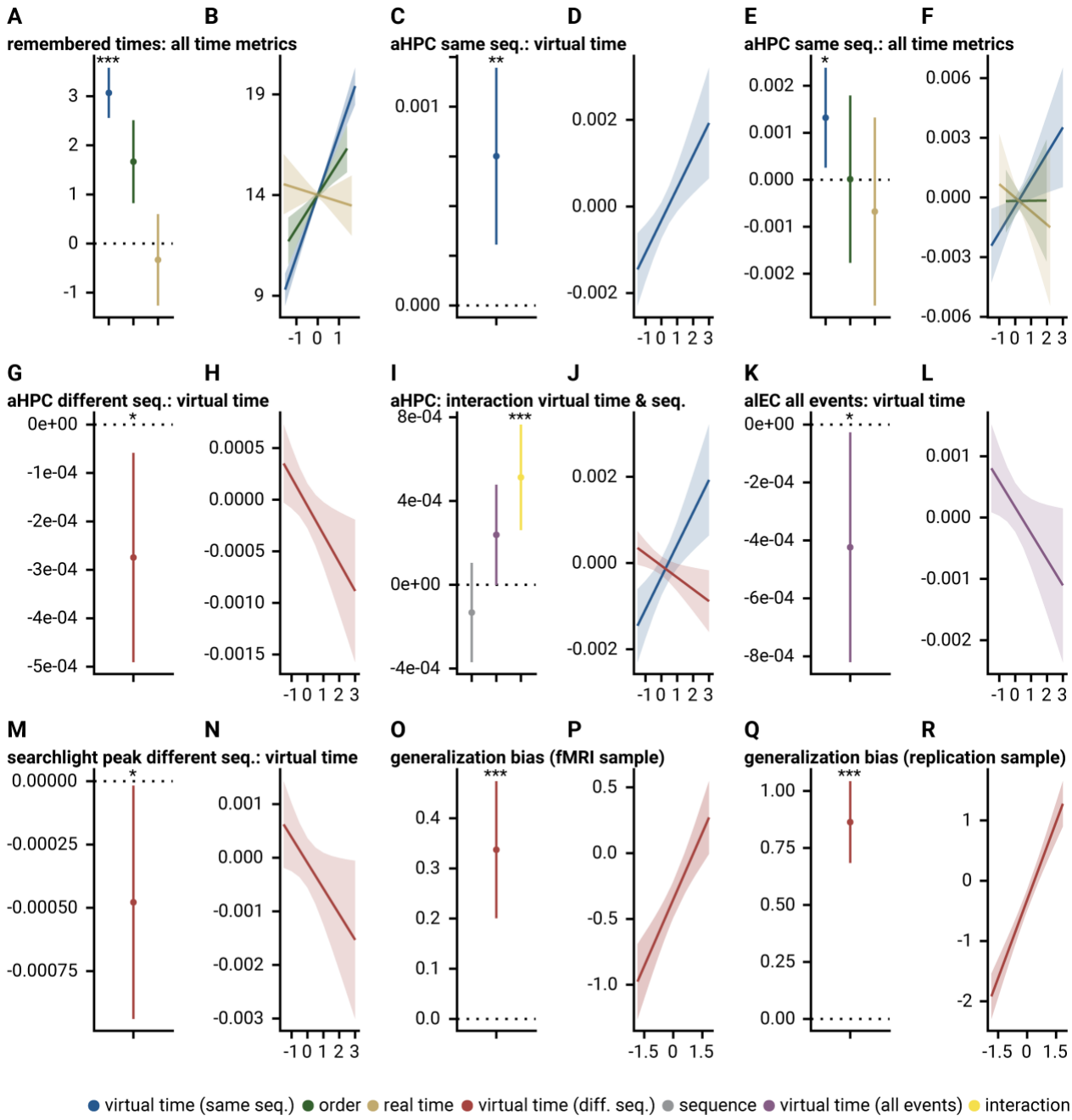
Supplemental Figure 2. Design of the day learning task. **A**. Each of the four virtual days consisted of a sequence of five events. Event sequences are shown in virtual time, i.e. relative to the hidden clock. Less virtual time passes within the bottom two sequences because clock speed was manipulated 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



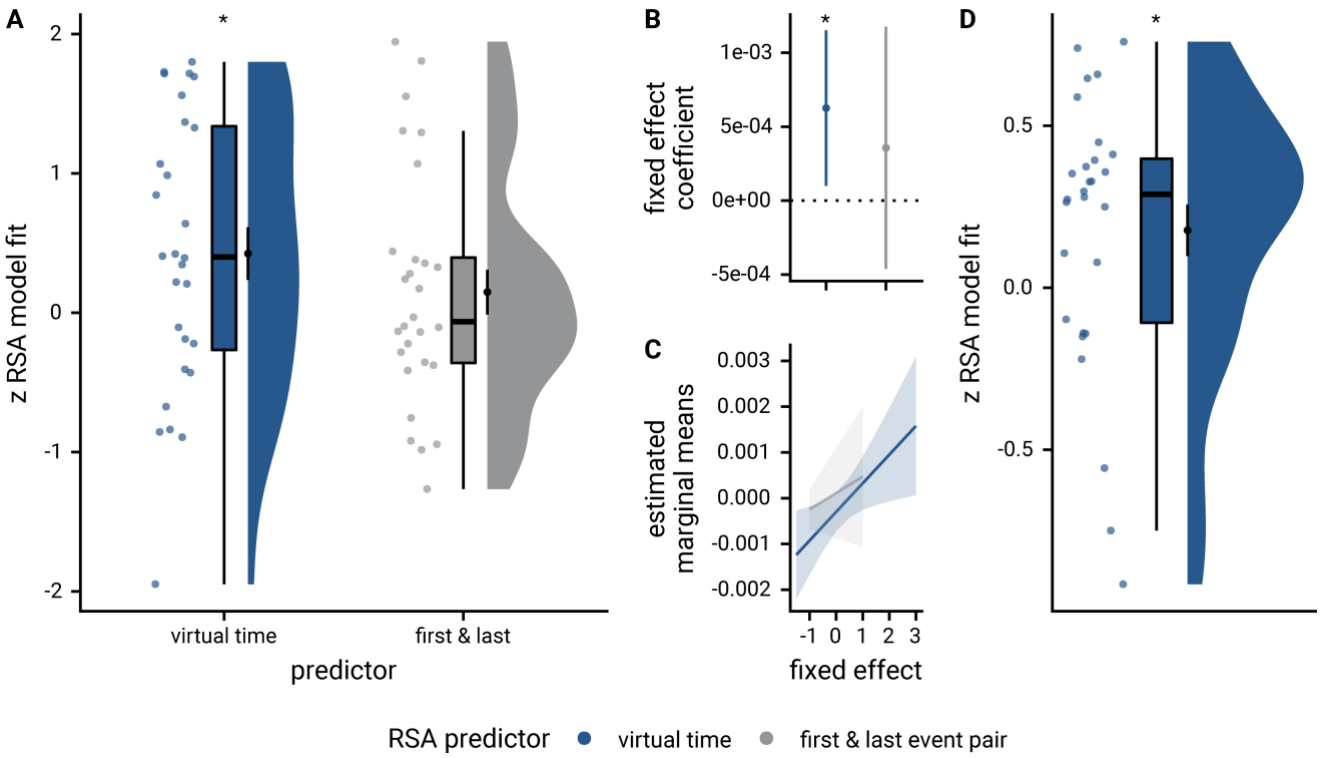
Supplemental Figure 3. Memory performance. A. A permutation-based repeated measures ANOVA revealed a significant effect of sequence on mean absolute errors in the timeline task ($F_{3,81}$=5.86, p<0.001, post hoc contrasts: sequence 1 vs. 2: t_{27}=3.38, p=0.001, sequence 1 vs. 3: t_{27}=-0.12, p=0.912, sequence 1 vs. 4: t_{27}=2.59, p=0.013, sequence 2 vs. 3: t_{27}=-2.92, p=0.001, sequence 2 vs. 4: t_{27}=-1.15, p=0.271, sequence 3 vs. 4: t_{27}=2.15, p=0.023). *p < Bonferroni-adjusted alpha-level of 0.008, corrected for 6 pairwise post hoc comparisons. **B**. Mean absolute timeline errors did not differ statistically between sequences with fast and slow clock speed (t_{27}=-0.82, p=0.423). **C**. The number of errors in the sorting task did not correlate with the mean absolute error in the timeline task across participants (r=0.23, p=0.246). **D**. Mean absolute errors in the timeline task were not statistically different between participants who made one or more errors (red) or no errors in the sorting task (green) in the sorting task (t-test for independent samples, t_{26}=-1.79, p=0.085). **E**. Histogram shows the number of swap errors for participants with (red) and without (green) errors in the sorting task. **F**. The distribution of swap errors over sequence positions did not deviate statistically from uniformity ($\chi^2(1)$=1.07, p=0.899). **G**. Histogram shows the null distribution of the proportion of swap errors expected under random sorting errors. The proportion of swap errors observed in our sample (red line) exceeded the 95th percentile of the null distribution (black line). **H**. The number of swap errors was not significantly correlated with the generalization bias (Spearman r=0.12, p=0.528). **I**. The generalization bias in the timeline task was not significantly different between participants who made one or more swap errors (red) or no swap errors (green) in the sorting task (t_{26}=0.18, p=0.861). **A**, **B**, **D**, **H**. 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. **C**, **H**. Each circle shows data from one participant, grey line and shaded region indicate least squares line and confidence interval.

Supplemental Figure 4



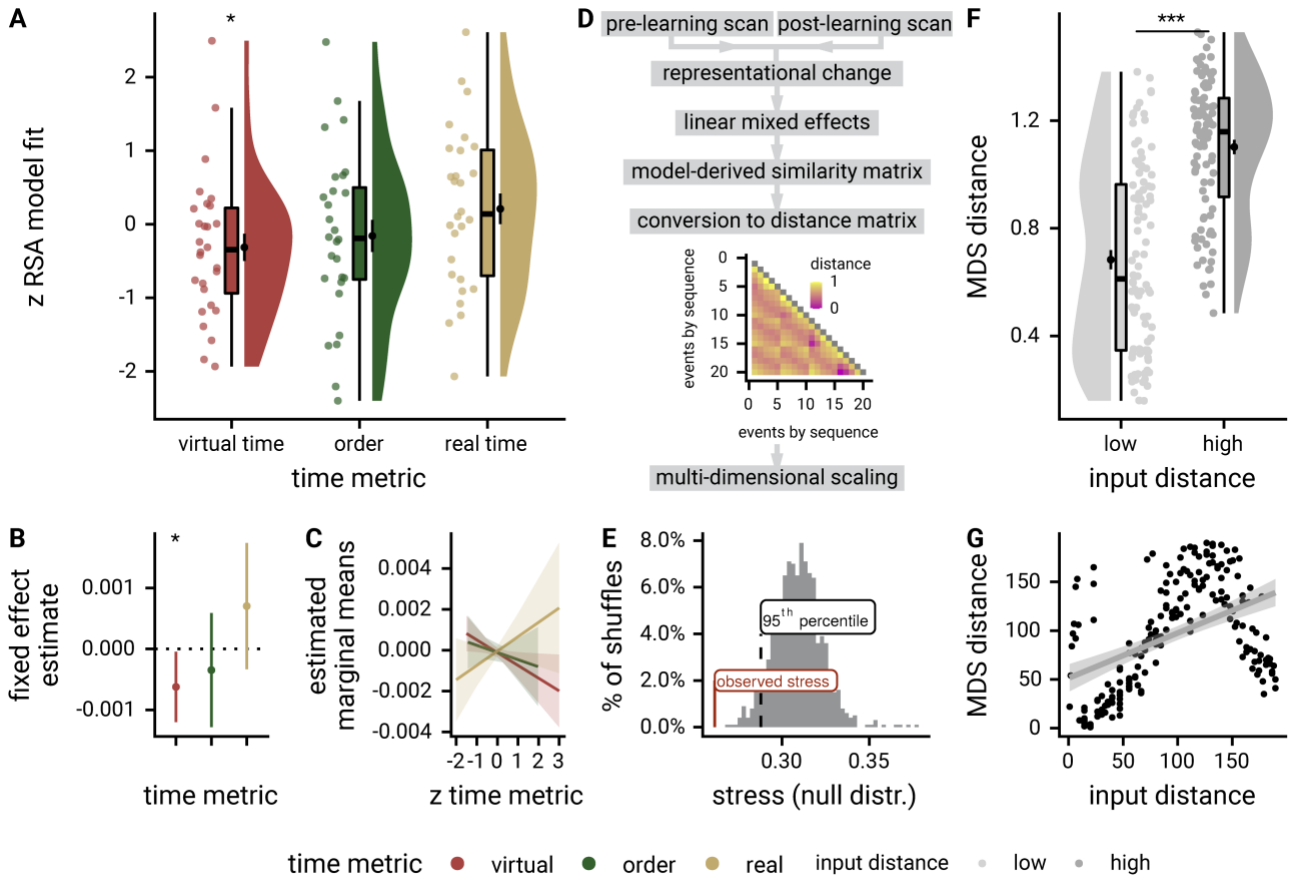
Supplemental Figure 4. Mixed model results. Dot plots show parameter estimates and 95% confidence intervals for fixed effects of mixed model analyses. Line plots show estimated marginal means. **A**, **B**. Remembered times in the time line task are predicted by virtual event times with order and real time in the model (c.f. **Figure 2B**). **C**, **D**. Temporal distances in virtual time explain representational change in the anterior hippocampus (aHPC) for same-sequence events (c.f. **Figure 4B**). **E**, **F**. Temporal distances in virtual time explain representational change in the aHPC for same-sequence events when competing for variance with temporal distances based on order and real time (c.f. **Figure 4D**). **G**, **H**. Temporal distances in virtual time explain representational change in the aHPC for different-sequence events (c.f. **Figure 5A**). **I**, **J**. There was a significant interaction of virtual temporal distances and sequence membership characterized by a differential relationship between temporal distances and aHPC representational change for event pairs from the same sequence or from different sequences (c.f. **Figure 5A**). **K**, **L**. Virtual temporal distances explain representational change in the anterior-lateral entorhinal cortex (alEC) when collapsing across all event pairs (c.f. **Figure 6B**). **M**, **N**. In the aHPC peak cluster of the same-sequence searchlight analysis, virtual temporal distances were siginificantly related to representational change for events from different sequences (c.f. **Figure 7B**). **O**-**R**. The relative time of events from other sequences predicted signed event time construction errors as measured in the timeline task (c.f. **Figure 8CD**) in the main fMRI sample (**O**, **P**) and in the independent replication sample (**Q**, **R**).

Supplemental Figure 5



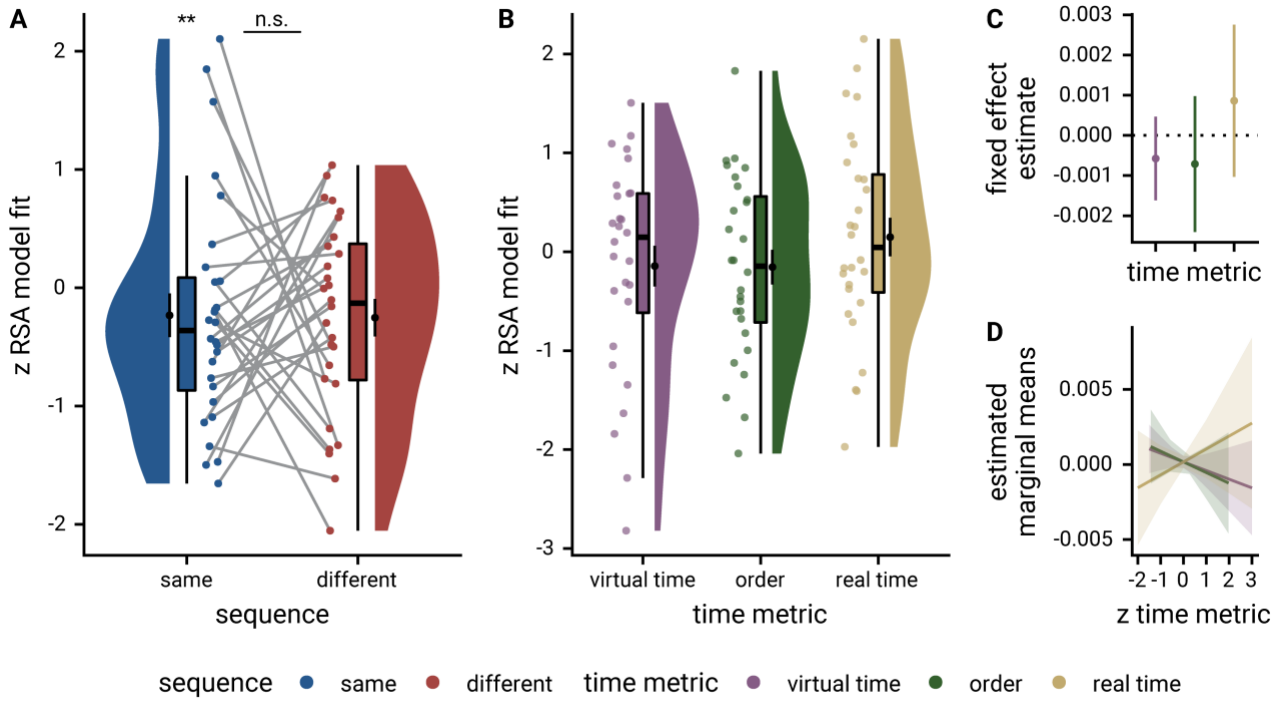
Supplemental Figure 5. The relationship of virtual time and hippocampal pattern similarity change is not driven by the first and last event of a sequence. **A**. Z-values from the summary statistics approach show a significant positive effect of virtual time on pattern similarity change in the anterior 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 6



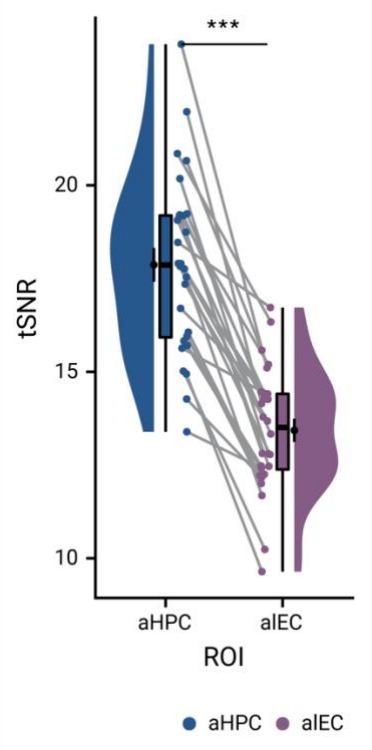
Supplemental Figure 6. Virtual time predicts hippocampal pattern similarity change for events from different sequences. A. Z-values show the relationship of the different time metrics to representational change in the anterior hippocampus based on participant-specific multiple regression analyses for pairs of events from 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. **B**, **C**. Parameter estimates with 95% confidence intervals (**B**) and estimated marginal means (**C**) show the fixed effects of the three time metrics from the corresponding mixed model. * p<0.05 after exclusion of one outlier excluded based on the boxplot criterion. **D**. A linear mixed model capturing the interaction effect of virtual temporal distances and sequence membership (Figure 5, Supplemental Figure 4IJ) was fitted to hippocampal representational change. An event-by-event similarity matrix was derived from the fixed effects of this model. Similarities were converted distances and then used as input for multidimensional scaling (see Methods). **E**. The stress value observed in the MDS analysis (red line) was significantly smaller than the 5th percentile (black dashed line) of a surrogate distribution of stress values obtained from shuffling the dissimilarities before running MDS in each of 1000 iterations. **F**. Pairs of events separated by a large distance in the input distance matrix were separated by a larger Euclidean distance in the resulting MDS configuration (t_{188}=9.35, p<0.001, d=1.35, 95% CI [1.03, 1.67]). *** p <0.001. G. There was a significant Spearman correlation of input distances and MDS configuration distances (r=0.46, p<0.001), but visual inspection reveals a non-linear relationship where very high distances are systematically underestimated in the MDS configuration. This is likely because the data were projected onto only two dimensions for visualization. More dimensions would be needed to improve the fit of the MDS configuration and the input distance matrix. Distances are shown as ranks because non-metric MDS was used (high ranks for high distances).

Supplemental Figure 7



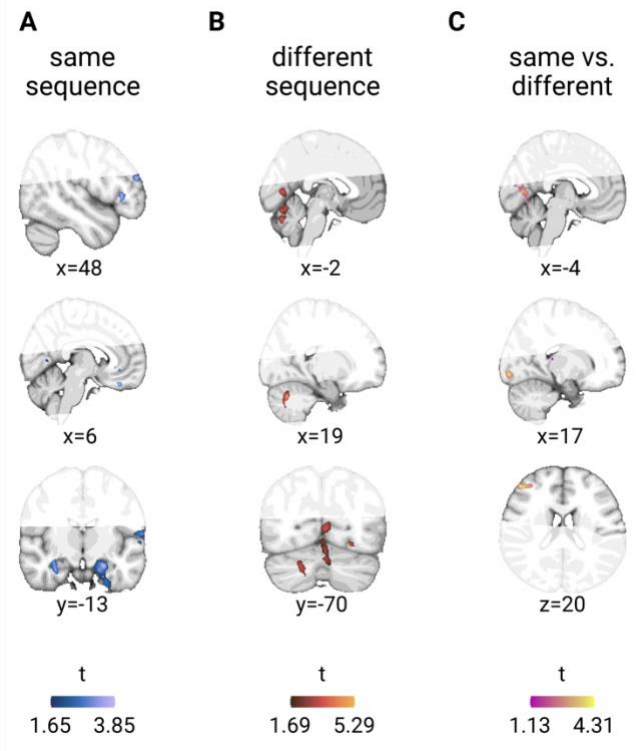
Supplemental Figure 7. Pattern similarity change in the anterior-lateral entorhinal cortex. **A**. Relationship of pattern similarity change and temporal distances between events from the same and different sequences in the anterior-lateral entorhinal cortex. There was no statistically significant difference 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, d=-0.69, 95% CI [-1.17, -0.27]; 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, d=-0.29, 95% CI [-0.69, 0.08]; α=0.025, corrected for separate tests of events of the same and different sequences). ** p<0.01 after outlier exclusion. **B**. Z-values show the relationship of the different time metrics to representational change in the anterior-lateral entorhinal cortex based on participant-specific multiple regression analyses. Analysis includes all pairs of events. **C**, **D**. Parameter estimates with 95% confidence intervals (**C**) and estimated marginal means (**D**) show the fixed effects of the three time metrics from the corresponding mixed model. A,B. 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.

Supplemental Figure 8



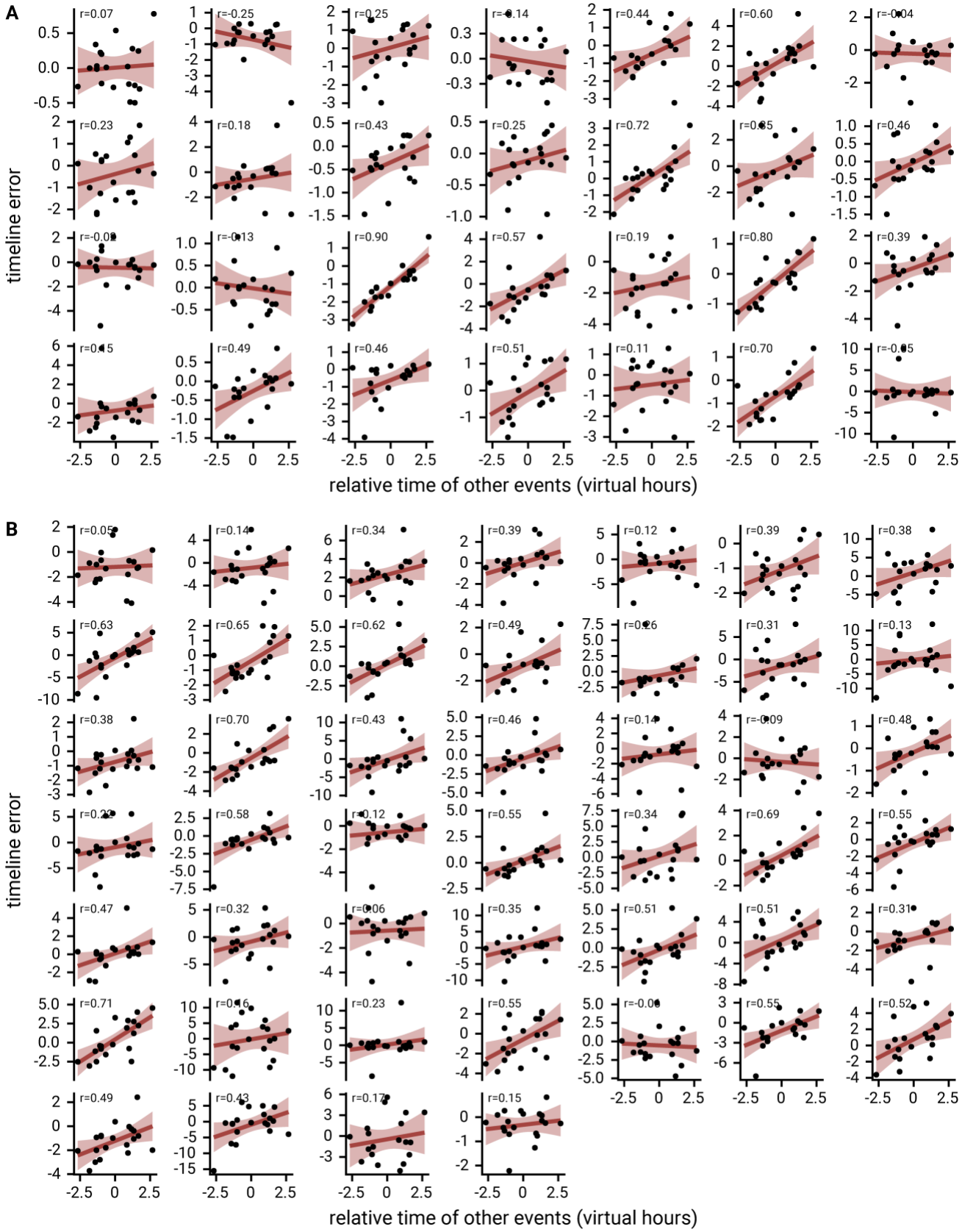
Supplemental Figure 8. Temporal signal-to-noise ratio in the anterior hippocampus and the anterior-lateral entorhinal cortex. The temporal signal-to-noise ratio was quantified as the mean unsmoothed signal over time divided by its standard deviation. It was calculated for each voxel and then averaged 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, d=1.99, 95% CI [1.65, 3.13]). 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 9



Supplemental Figure 9. 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 12). **B.** Pattern similarity change correlated negatively with temporal distances between events from different sequences in the cerebellum and lingual gyrus (see Supplemental Table 14). **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 15). **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 10



Supplemental Figure 10. 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. Correlation coefficients are based on Pearson correlation. **A** shows data from the main sample; **B** from the replication sample.

References

1. Bartlett, F. C. *Remembering: a study in experimental and social psychology*. (Cambridge University Press, 1932).
2. Cheng, S., Werning, M. & Suddendorf, T. Dissociating memory traces and scenario construction in mental time travel. *Neurosci. Biobehav. Rev.* **60**, 82–89 (2016).
3. Hassabis, D. & Maguire, E. A. Deconstructing episodic memory with construction. *Trends Cogn. Sci.* **11**, 299–306 (2007).
4. Irish, M. & Piguet, O. The Pivotal Role of Semantic Memory in Remembering the Past and Imagining the Future. *Front. Behav. Neurosci.* **7**, 27 (2013).
5. Schacter, D. L. & Addis, D. R. The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philos. Trans. R. Soc. B Biol. Sci.* **362**, 773–786 (2007).
6. Schacter, D. L. & Addis, D. R. Memory and imagination: Perspectives on constructive episodic simulation. in *The Cambridge Handbook of Imagination* (ed. Abraham, A.) 111–131 (Cambridge University Press, 2020).
7. Schacter, D. L., Benoit, R. G. & Szpunar, K. K. Episodic Future Thinking: Mechanisms and Functions. *Curr. Opin. Behav. Sci.* **17**, 41–50 (2017).
8. Friedman, W. J. Memory for the time of past events. *Psychol. Bull.* **113**, 44 (1993).
9. Friedman, W. J. Time in Autobiographical Memory. *Soc. Cogn.* **22**, 591–605 (2004).
10. Bellmund, J. L. S., Polti, I. & Doeller, C. F. Sequence Memory in the Hippocampal–Entorhinal Region. *J. Cogn. Neurosci.* **32**, 2056–2070 (2020).
11. Ranganath, C. & Hsieh, L. The hippocampus: a special place for time. *Ann. N. Y. Acad. Sci.* **1369**, 93–110 (2016).
12. Behrens, T. E. J. *et al.* What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. *Neuron* **100**, 490–509 (2018).
13. Zeithamova, D. & Bowman, C. R. Generalization and the hippocampus: More than one story? *Neurobiol. Learn. Mem.* **175**, 107317 (2020).
14. Kumaran, D. & Maguire, E. A. An Unexpected Sequence of Events: Mismatch Detection in the Human Hippocampus. *PLOS Biol.* **4**, e424 (2006).
15. Kumaran, D. & Maguire, E. A. The Dynamics of Hippocampal Activation during Encoding of Overlapping Sequences. *Neuron* **49**, 617–629 (2006).
16. Baldassano, C. *et al.* Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron* **95**, 709-721.e5 (2017).
17. Ben-Yakov, A. & Dudai, Y. Constructing Realistic Engrams: Poststimulus Activity of Hippocampus and Dorsal Striatum Predicts Subsequent Episodic Memory. *J. Neurosci.* **31**, 9032–9042 (2011).
18. Hsieh, L.-T., Gruber, M. J., Jenkins, L. J. & Ranganath, C. Hippocampal Activity Patterns Carry Information about Objects in Temporal Context. *Neuron* **81**, 1165–1178 (2014).
19. Thavabalasingam, S., O'Neil, E. B. & Lee, A. C. H. Multivoxel pattern similarity suggests the integration of temporal duration in hippocampal event sequence representations. *NeuroImage* **178**, 136–146 (2018).
20. Thavabalasingam, S., O'Neil, E. B., Tay, J., Nestor, A. & Lee, A. C. H. Evidence for the incorporation of temporal duration information in human hippocampal long-term memory sequence representations. *Proc. Natl. Acad. Sci.* **116**, 6407–6414 (2019).
21. Deuker, L., Bellmund, J. L. S., Navarro Schröder, T. & Doeller, C. F. An event map of memory space in the hippocampus. *eLife* **5**, e16534 (2016).
22. DuBrow, S. & Davachi, L. Temporal Memory Is Shaped by Encoding Stability and Intervening Item Reactivation. *J. Neurosci.* **34**, 13998–14005 (2014).
23. Ezzyat, Y. & Davachi, L. Similarity Breeds Proximity: Pattern Similarity within and across Contexts Is Related to Later Mnemonic Judgments of Temporal Proximity. *Neuron* **81**, 1179–1189 (2014).
24. Jenkins, L. J. & Ranganath, C. Distinct neural mechanisms for remembering when an event occurred. *Hippocampus* **26**, 554–559 (2016).
25. Kyle, C. T., Smuda, D. N., Hassan, A. S. & Ekstrom, A. D. Roles of human hippocampal subfields in retrieval of spatial and temporal context. *Behav. Brain Res.* **278**, 549–558 (2015).
26. Lositsky, O. *et al.* Neural pattern change during encoding of a narrative predicts retrospective duration estimates. *eLife* **5**, e16070 (2016).
27. Bellmund, J. L. S., Deuker, L. & Doeller, C. F. Mapping sequence structure in the human lateral entorhinal cortex. *eLife* **8**, e45333 (2019).
28. Montchal, M. E., Reagh, Z. M. & Yassa, M. A. Precise temporal memories are supported by the lateral entorhinal cortex in humans. *Nat. Neurosci.* **22**, 284–288 (2019).
29. Evensmoen, H. R. *et al.* Metric and chronological time in human episodic memory. *bioRxiv* 2020.05.11.084202 (2020) doi:10.1101/2020.05.11.084202.
30. Ebbinghaus, H. Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie. (Duncker & Humblot, 1885).
31. Lewandowsky, S. & Murdock, B. B. Memory for serial order. *Psychol. Rev.* **96**, 25–57 (1989).
32. Jensen, O. & Lisman, J. E. Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends Neurosci.* **28**, 67–72 (2005).
33. Bright, I. M. *et al.* A temporal record of the past with a spectrum of time constants in the monkey

entorhinal cortex. *Proc. Natl. Acad. Sci.* **117**, 20274–20283 (2020).
34. Tsao, A. *et al.* Integrating time from experience in the lateral entorhinal cortex. *Nature* **561**, 57–62 (2018).
35. Howard, M. W. & Kahana, M. J. A Distributed Representation of Temporal Context. *J. Math. Psychol.* **46**, 269–299 (2002).
36. Szpunar, K. K., Spreng, R. N. & Schacter, D. L. A taxonomy of prospection: Introducing an organizational framework for future-oriented cognition. *Proc. Natl. Acad. Sci.* **111**, 18414–18421 (2014).
37. Kumaran, D. & McClelland, J. L. Generalization Through the Recurrent Interaction of Episodic Memories. *Psychol. Rev.* **119**, 573–616 (2012).
38. Renoult, L., Irish, M., Moscovitch, M. & Rugg, M. D. From Knowing to Remembering: The Semantic–Episodic Distinction. *Trends Cogn. Sci.* S1364661319302323 (2019) doi:10.1016/j.tics.2019.09.008.
39. Schacter, D. L., Addis, D. R. & Buckner, R. L. Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci.* **8**, 657–661 (2007).
40. Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M. & Norman, K. A. 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 (2017).
41. Bunsey, M. & Eichenbaum, H. Conservation of hippocampal memory function in rats and humans. *Nature* **379**, 255–257 (1996).
42. Heckers, S., Zalesak, M., Weiss, A. P., Ditman, T. & Titone, D. Hippocampal activation during transitive inference in humans. *Hippocampus* **14**, 153–162 (2004).
43. Park, S. A., Miller, D. S., Nili, H., Ranganath, C. & Boorman, E. D. Map Making: Constructing, Combining, and Inferring on Abstract Cognitive Maps. *Neuron* **107**, 1226–1238 (2020).
44. Dusek, J. A. & Eichenbaum, H. The hippocampus and memory for orderly stimulus relations. *Proc. Natl. Acad. Sci.* **94**, 7109–7114 (1997).
45. Koster, R. *et al.* Big-Loop Recurrence within the Hippocampal System Supports Integration of Information across Episodes. *Neuron* **99**, 1342–1354 (2018).
46. Preston, A. R., Shrager, Y., Dudukovic, N. M. & Gabrieli, J. D. E. Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus* **14**, 148–52 (2004).
47. Schlichting, M. L., Mumford, J. A. & Preston, A. R. Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nat. Commun.* **6**, 8151 (2015).
48. Shohamy, D. & Wagner, A. D. Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron* **60**, 378–89 (2008).
49. Zeithamova, D. & Preston, A. R. Flexible Memories: Differential Roles for Medial Temporal Lobe and Prefrontal Cortex in Cross-Episode Binding. *J. Neurosci.* **30**, 14676–14684 (2010).
50. Zeithamova, D., Dominick, A. L. & Preston, A. R. Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron* **75**, 168–79 (2012).
51. Whittington, J. C. R. *et al.* The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation. *Cell* **183**, 1249-1263.e23 (2020).
52. Baram, A. B., Muller, T. H., Nili, H., Garvert, M. M. & Behrens, T. E. J. Entorhinal and ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning problems. *Neuron* (2020) doi:10.1016/j.neuron.2020.11.024.
53. Morton, N. W., Schlichting, M. L. & Preston, A. R. Representations of common event structure in medial temporal lobe and frontoparietal cortex support efficient inference. *Proc. Natl. Acad. Sci.* **117**, 29338–29345 (2020).
54. Sun, C., Yang, W., Martin, J. & Tonegawa, S. Hippocampal neurons represent events as transferable units of experience. *Nat. Neurosci.* 1–13 (2020) doi:10.1038/s41593-020-0614-x.
55. Hemmer, P. & Steyvers, M. Integrating episodic memories and prior knowledge at multiple levels of abstraction. *Psychon. Bull. Rev.* **16**, 80–87 (2009).
56. Hemmer, P., Tauber, S. & Steyvers, M. Moving beyond qualitative evaluations of Bayesian models of cognition. *Psychon. Bull. Rev.* **22**, 614–628 (2015).
57. Franklin, N. T., Norman, K. A., Ranganath, C., Zacks, J. M. & Gershman, S. J. Structured Event Memory: A neuro-symbolic model of event cognition. *Psychol. Rev.* **127**, 327–361 (2020).
58. Radvansky, G. A. & Zacks, J. M. *Event cognition*. (Oxford University Press, 2014).
59. Zacks, J. M. Event Perception and Memory. *Annu. Rev. Psychol.* **71**, 165–191 (2020).
60. Irish, M., Addis, D. R., Hodges, J. R. & Piguet, O. Considering the role of semantic memory in episodic future thinking: evidence from semantic dementia. *Brain* **135**, 2178–2191 (2012).
61. Devitt, A. L., Addis, D. R. & Schacter, D. L. Episodic and semantic content of memory and imagination: A multilevel analysis. *Mem. Cognit.* **45**, 1078–1094 (2017).
62. Frisoni, M., Di Ghionno, M., Guidotti, R., Tosoni, A. & Sestieri, C. Reconstructive nature of temporal memory for movie scenes. *Cognition* **208**, 104557 (2021).
63. Luyckx, F., Nili, H., Spitzer, B. & Summerfield, C. Neural structure mapping in human probabilistic reward learning. *eLife* **8**, e42816 (2019).
64. Okazawa, G., Hatch, C. E., Mancoo, A., Machens, C. K. & Kiani, R. Representational geometry of

perceptual decisions in the monkey parietal cortex. *Cell* **184**, 3748-3761.e18 (2021).
65. Nelli, S., Braun, L., Dumbalska, T., Saxe, A. & Summerfield, C. Neural knowledge assembly in humans and deep networks. *bioRxiv* (2021) doi:https://doi.org/10.1101/2021.10.21.465374.
66. Montijn, N. D., Gerritsen, L. & Engelhard, I. M. The effect of stress on memory for temporal context: an exploratory study. *bioRxiv* (2021).
67. Shimbo, A., Izawa, E.-I. & Fujisawa, S. Scalable representation of time in the hippocampus. *Sci. Adv.* **7**, eabd7013 (2021).
68. Carpenter, A. C., Thakral, P. P., Preston, A. R. & Schacter, D. L. Reinstatement of item-specific contextual details during retrieval supports recombination-related false memories. *NeuroImage* **236**, 118033 (2021).
69. Schapiro, A. C., Kustner, L. V. & Turk-Browne, N. B. Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Curr. Biol.* **22**, 1622–7 (2012).
70. Bellmund, J. L. S., Gärdenfors, P., Moser, E. I. & Doeller, C. F. Navigating cognition: Spatial codes for human thinking. *Science* **362**, eaat6766 (2018).
71. Nielson, D. M., Smith, T. A., Sreekumar, V., Dennis, S. & Sederberg, P. B. Human hippocampus represents space and time during retrieval of real-world memories. *Proc. Natl. Acad. Sci.* **112**, 11078–11083 (2015).
72. Jazayeri, M. & Shadlen, M. N. Temporal context calibrates interval timing. *Nat. Neurosci.* **13**, 1020–1026 (2010).
73. Polti, I., Nau, M., Kaplan, R., Wassenhove, V. van & Doeller, C. F. Rapid encoding of task regularities in the human hippocampus guides sensorimotor timing. *bioRxiv* (2021) doi:https://doi.org/10.1101/2021.08.03.454928.
74. Orlov, T., Yakovlev, V., Hochstein, S. & Zohary, E. Macaque monkeys categorize images by their ordinal number. *Nature* **404**, 77–80 (2000).
75. Bower, G. H., Black, J. B. & Turner, T. J. Scripts in memory for text. *Cognit. Psychol.* **11**, 177–220 (1979).
76. Tompary, A. & Thompson-Schill, S. L. Semantic influences on episodic memory distortions. *J. Exp. Psychol. Gen.* Advance online publication (2021) doi:10.1037/xge0001017.
77. Bellmund, J. L. S. *et al.* Deforming the metric of cognitive maps distorts memory. *Nat. Hum. Behav.* **4**, 177–188 (2020).
78. Carpenter, A. C. & Schacter, D. L. Flexible retrieval: When true inferences produce false memories. *J. Exp. Psychol. Learn. Mem. Cogn.* **43**, 335–349 (2017).
79. Diamond, N. B. & Levine, B. Linking Detail to Temporal Structure in Naturalistic-Event Recall. *Psychol. Sci.* **31**, 1557–1572 (2020).
80. Greenberg, D. L. & Verfaellie, M. Interdependence of episodic and semantic memory: Evidence from neuropsychology. *J. Int. Neuropsychol. Soc.* **16**, 748–753 (2010).
81. Addis, D. R. Mental Time Travel? A Neurocognitive Model of Event Simulation. *Rev. Philos. Psychol.* **11**, 233–259 (2020).
82. DuBrow, S. & Davachi, L. Commentary: Distinct neural mechanisms for remembering when an event occurred. *Front. Psychol.* **8**, (2017).
83. Chanales, A. J. H., Oza, A., Favila, S. E. & Kuhl, B. A. Overlap among Spatial Memories Triggers Repulsion of Hippocampal Representations. *Curr. Biol.* **27**, 2307-2317.e5 (2017).
84. Favila, S. E., Chanales, A. J. H. & Kuhl, B. A. Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. *Nat. Commun.* **7**, 11066 (2016).
85. Lohnas, L. J. *et al.* Time-resolved neural reinstatement and pattern separation during memory decisions in human hippocampus. *Proc. Natl. Acad. Sci.* **115**, E7418–E7427 (2018).
86. Zeithamova, D., Gelman, B. D., Frank, L. & Preston, A. R. Abstract Representation of Prospective Reward in the Hippocampus. *J. Neurosci.* **38**, 10093–10101 (2018).
87. Benoit, R. G. & Schacter, D. L. Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. *Neuropsychologia* **75**, 450–457 (2015).
88. Barron, H. C., Dolan, R. J. & Behrens, T. E. J. Online evaluation of novel choices by simultaneous representation of multiple memories. *Nat. Neurosci.* **16**, 1492–1498 (2013).
89. Hsieh, L.-T. & Ranganath, C. Cortical and subcortical contributions to sequence retrieval: Schematic coding of temporal context in the neocortical recollection network. *NeuroImage* **121**, 78–90 (2015).
90. Liu, Y., Dolan, R. J., Kurth-Nelson, Z. & Behrens, T. E. J. Human Replay Spontaneously Reorganizes Experience. *Cell* **178**, 640-652.e14 (2019).
91. Summerfield, C., Luyckx, F. & Sheahan, H. Structure Learning and the Posterior Parietal Cortex. *Prog. Neurobiol.* **184**, 101717 (2020).
92. Sheahan, H., Luyckx, F., Nelli, S., Teupe, C. & Summerfield, C. Neural state space alignment for magnitude generalization in humans and recurrent networks. *Neuron* **109**, 1214–1226 (2021).
93. Estefan, D. P. *et al.* Volitional learning promotes theta phase coding in the human hippocampus. *Proc. Natl. Acad. Sci.* **118**, e2021238118 (2021).
94. Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* **23**, S208–S219 (2004).
95. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2020).
96. Morrissey, M. B. & Ruxton, G. D. Multiple Regression Is Not Multiple Regressions: The Meaning of Multiple Regression and the Non-Problem of Collinearity. *Philos. Theory Pract. Biol.* **10**, (2018).

97. Vanhove, J. Collinearity isn't a disease that needs curing. *PsyArXiv* (2020) doi:10.31234/osf.io/mv2wx.
98. Poppenk, J., Evensmoen, H. R., Moscovitch, M. & Nadel, L. Long-axis specialization of the human hippocampus. *Trends Cogn. Sci.* **17**, 230–240 (2013).
99. Navarro Schröder, T., Haak, K. V., Zaragoza Jimenez, N. I., Beckmann, C. F. & Doeller, C. F. Functional topography of the human entorhinal cortex. *eLife* **4**, e06738 (2015).
100. Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R. & Kievit, R. A. Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Res.* **4**, 63 (2019).
101. Torchiano, M. *Effsize - A Package For Efficient Effect Size Computation*. (Zenodo, 2016). doi:10.5281/ZENODO.196082.
102. Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* **68**, 255–278 (2013).
103. Lüdecke, D. ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *J. Open Source Softw.* **3**, 772 (2018).
104. Kriegeskorte, N. *et al.* Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. *Neuron* **60**, 1126–1141 (2008).
105. Stelzer, J., Chen, Y. & Turner, R. Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. *NeuroImage* **65**, 69–82 (2013).
106. Anderson, M. J. & Legendre, P. An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. *J. Stat. Comput. Simul.* **62**, 271–303 (1999).
107. Frossard, J. & Renaud, O. permuco: Permutation tests for regression, (repeated measures) ANOVA/ANCOVA and comparison of signals. (2019).
108. Singmann, H. et al. afex: Analysis of Factorial Experiments. (2021).
109. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
110. Barr, D. J. Random effects structure for testing interactions in linear mixed-effects models. *Front. Psychol.* **4**, 328 (2013).
111. Leeuw, J. de & Mair, P. Multidimensional Scaling Using Majorization: SMACOF in R. *J. Stat. Softw.* **31**, 1–30 (2009).

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/). 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, 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.

Competing interests

The authors declare no competing interests.

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: Virtual time explains representational change for different-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	-0.000101	0.000112	-0.90	-0.000319	0.000118
virtual time	-0.000623	0.000294	-2.12	-0.001201	-0.000046
order	-0.000348	0.000478	-0.73	-0.001284	0.000589
real time	0.000702	0.000529	1.33	-0.000334	0.001739

random effects		
group	**term**	**estimate**
participant	virtual time (SD)	0.000000
residual	SD	0.007077

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	5	-28594.64	14302.32			
full model	6	-28597.12	14304.56	4.48	1	0.034

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 7

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 8

Mixed Model: The effect of virtual time differs between same-sequence and different-sequence events in the anterior hippocampus when including interactions with other time metrics

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	-0.000190	0.000122	-1.55	-0.000428	0.000049
virtual time	0.000237	0.000124	1.92	-0.000005	0.000480
day	-0.000130	0.000121	-1.07	-0.000368	0.000108
interaction virtual time and day	0.000769	0.000262	2.93	0.000255	0.001283
interaction order and day	0.000287	0.000418	0.69	-0.000533	0.001106
interaction real time and day	-0.000558	0.000464	-1.20	-0.001468	0.000351

random effects		
group	**term**	**estimate**
participant	interaction virtual time and day (SD)	0.000176
residual	SD	0.007065

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	7	-37572.96	18793.48			
full model	8	-37579.53	18797.77	8.57	1	0.003

model:
ps_change~vir_time_diff*same_day_dv+order_diff:same_day_dv+real_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 9

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 10

Mixed Model: Virtual time does not explain representational change for different-sequence events in the anterior-lateral entorhinal cortex when including order and real time in the model

fixed effects					
term	**estimate**	**SE**	**t-value**	**95% CI**	
intercept	0.000167	0.000202	0.83	-0.000229	0.000563
virtual time	-0.000576	0.000531	-1.09	-0.001617	0.000464
order	-0.000712	0.000862	-0.83	-0.002402	0.000978
real time	0.000862	0.000966	0.89	-0.001031	0.002754

random effects		
group	**term**	**estimate**
participant	virtual time (SD)	0.000000
residual	SD	0.014733

model comparison						
model	**npar**	**AIC**	**LL**	**χ^2**	**df**	**p**
reduced model	5	-29767.41	14888.71			
full model	6	-29766.59	14889.30	1.18	1	0.278

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 11

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 12

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 13

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 14

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 15

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 16

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 17

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