Causal bridges and agentic engagement shape temporal memory for a virtual escape room game

Wangjing Yu[1,2], Benjamin M. Silver[2], Kevin N. Ochsner[2], Lila Davachi[2,3,*]

[1] Department of Psychology, New York University, New York 10003, USA.
[2] Department of Psychology, Columbia University, New York 10027, USA.
[3] Department of Clinical Research, Nathan Kline Institute for Psychiatric Research, Orangeburg, NY 10962, USA

*Correspondence:
Lila Davachi
Schermerhorn Hall,
Rm 401A 1180 Amsterdam Ave,
New York,
NY 10027, USA
Email: ld24@columbia.edu

This file includes:

Main text

Figures 1 to 5

Abstract (146 words)

While time flows uniformly, many factors influence memory for temporal information. The current study examined how cross-event interconnection and goal-directed, agentic engagement impacted the temporal structure of memory for a virtual escape room game played with friends. We found that after controlling for the actual temporal distance, causally-related events were remembered as closer in time than unrelated events. This effect grew stronger after a consolidation period, with causally-connected memories showing greater temporal integration than memories that shared contextual overlap. Notably, within these causally-related events, temporal memory was superior for events where game clues were found. Further, participants who made greater in-game contributions to puzzle-solving showed better temporal memory and increased temporal integration between the contextually-related events compared to participants who contributed less. Together, these findings highlight the role of causal connections in driving memory integration and that agentic task engagement modulates how real-world experiences are remembered.

Keywords: temporal memory; causality; memory integration; agentic task engagement; collaborative problem-solving

Introduction

From going on a date, to having Thanksgiving dinner, to playing games, people engage in all kinds of activities together that involve sequences of events unfolding over time. Memory for these sequences is meaningful not only as a record of how and when the chain of events unfolds, but also as a record of who was responsible for the events. Did Paolo lean in to kiss Victoria at the start or middle of the date? When did Camille spill the wine on Uncle John's shirt while reaching for the turkey? Did Arlene solve a key clue early in the game or at the end? Here, we studied temporal memory in this latter context – when groups play games together. We asked how two factors – cross-event interconnection and agentic task engagement – shape subsequent temporal memory structure in a virtual escape room game.

Accumulating evidence has suggested that memories associated with different temporal information may become integrated (and stored proximally) or differentiated (and stored distantly) (Bein & Davachi, 2024; Clewett et al., 2019, 2025; Ezzyat & Davachi, 2014; Lositsky et al., 2016). Much of this research has focused on investigating temporal integration of information encoded within a continuous event and the disruptive effect of contextual shifts (i.e., event boundaries) (Baldassano et al., 2017; Block & Reed, 1978; DuBrow & Davachi, 2013, 2014, 2016; Sherman et al., 2023; Zacks, 2007; Zakay et al., 1994). However, other findings have highlighted that shared elements between temporally discontinuous events also allow the formation of an integrated memory representation (Schlichting et al., 2015; Schlichting & Preston, 2015; Shohamy & Wagner, 2008; Zeithamova et al., 2012), and that consolidation promotes feature-based integration (Audrain & McAndrews, 2022; Tompary & Davachi, 2017, 2024).

The kind of relationships between memories that fosters integration has not been fully characterized, however. Broadly related to the idea that contextual stability enhances information binding (Clewett et al., 2019; Davachi & DuBrow, 2015; Ezzyat & Davachi, 2011; Swallow et al., 2009), recent studies have demonstrated that events sharing the same encoding context, even when temporally distant, show increased memory integration (Cox et al., 2021; Liu et al., 2024; Yu et al., 2025). Beyond contextual overlap, causal connection has also been proposed to play a key role in linking memories (Chen & Bornstein, 2024). Supporting this notion, narrative comprehension studies have shown that retrieving causally-related prior events facilitates the understanding of the current event, which in turn enhances memory integration between these events (Chang et al., 2021; Cohn-Sheehy et al., 2021; Milivojevic et al., 2015; Song et al., 2025). However, it remains unclear if causality leads to more or less memory integration than that afforded by contextual similarity. Moreover, it is understudied precisely how the representational change triggered by different types of cross-event interconnections may be reflected in the temporal structure of these memories.

As we navigate through daily events, the ebb and flow of our engagement in those experiences also impacts how we behave during these events and how we later remember them (Aly & Turk-Browne, 2016; Chun et al., 2011; DeBettencourt et al., 2018; Li et al., 2024; Murty et al., 2015, 2019; Ruiz et al., 2023; Song et al., 2021). In particular, engagement afforded by agency during a task has been shown to result in strengthened connections among the encoded events in memory (Houser et al., 2022; Li et al., 2024; Ruiz et al., 2023), which points to a potential role of agentic engagement in driving information binding across memories. Notably, majority of this work has measured agency at the level of the self in tasks performed individually by participants. However, when considering how these effects generalize to real-life experiences, it is important to recognize that we often share experiences with other people in daily lives. Yet it remains an open question whether, for those who share the same socially-meaningful experience, differences in agentic engagement across individuals may lead to distinctive memories for the naturally unfolding events. Critically, we ask how these differences, if present, are reflected in the temporal organization of their memories.

To address these questions, the current experiment examined people's temporal memory for a collaborative experience shared with their real-life friends. Specifically, groups of friends played a virtual escape room game where they worked together to solve a series of interrelated puzzles. The format of the game allows for each group to share a unique social experience, while each individual in the group makes different contributions to solving puzzles. Both on the same day of the game and one week later, we asked participants to recall when the milestone events occurred during the game. Leveraging this naturalistic yet structured encoding experience, we investigated whether game events that were contextually related, causally linked, or unrelated might show different levels of memory integration, as reflected in the remember temporal distance between them. Further, we examined whether, despite sharing the same experience and task goal, group members who differed in agentic task engagement, as measured by their in-game contribution, might exhibit distinct temporal memory structures.

Methods

Participants

Thirty groups of 3-5 friends participated in the experiment (total N=142; 140 participants reported demographic information: mean age = 25.83 years; 94 females, 44 males, and 2 reported non-binary gender). All participants provided informed consent and received monetary compensation for their participation. The study protocol was approved by the Columbia University Institutional Review Board.

Exclusion Criteria

We excluded 18 participants from all analyses due to: (a) technical issues recording the game experience (3 groups; N=14); (b) failed completion of both the recent and remote memory tests (N=4). We analyzed data from the remaining 124 participants (27 groups; 123 participants reported demographic information: mean age=26.48 years; 78 females, 43 males, and 1 reported non-binary gender).

We further excluded 11 participants from the analyses of the recent memory test due to delayed (with any overnight interval following the game) or failed completion, and 5 participants from the analyses of the remote memory test due to failed completion. Therefore, 113 participants contributed usable data for the recent memory test, and 119 participants for the remote memory test, with 108 participants contributing data for both. Additional analysis-specific exclusions were noted and justified in Results and Supplementary Materials.

Sample size rationale

By the time of data collection (March 2021), there was no prior literature to our knowledge that had investigated similar research questions with a paradigm involving a real-world group experience as the current study. We thus performed a power analysis based on a finding from a highly-controlled laboratory study showing that events encoded from the same context were remembered as closer in time than those spanning different contexts after a short delay (Clewett et al., 2020). This effect was replicated across three experiments with small-to-medium effect sizes (Cohen's d>0.41). Given that the current study involved more complex event structures and included both short and long delays, we anticipated greater variance across participants, which might lead to a reduced effect size. Accordingly, we conservatively expected a smaller effect (Cohen's d≈0.3) compared to Clewett et al.'s finding. Based on this target effect size, we estimated that a sample size of at least 90 participants (18 to 30 groups of 3 to 5 friends) will be needed to achieve power of 0.8 at an alpha level of 0.05 (two-tailed). Our final sample size exceeded one hundred, providing sufficient statistical power to detect the target effects.

Experimental procedures

Participants met over Zoom and played the virtual escape room game together. All groups played the same game. One week prior to the game, they completed a pre-game survey. On the same day of the game (recent test) and one week after the game (remote test), all participants individually completed a memory test embedded in a post-game survey (Fig. 1A). Note that the pre-game survey and other questions in the post-game surveys were not relevant to the current manuscript, and more information about these surveys can be found in Silver et al., 2025.

Virtual escape room game

The virtual escape room game was designed and administered by a game company (Puzzle Break LLC; https://www.puzzlebreakli.com/), and the game experience for each group was recorded over Zoom. At the beginning of the game, an employee from the game company introduced the game plot and provided game-related files to the participants; they also stayed on the Zoom meeting throughout the game and would be available to provide hints when needed. During the game, participants were expected to collaborate with their group members and solve a series of puzzles until they reached the final goal (i.e., hacking into an evil company's computer system to save the world). On average, the participating groups completed the game in 49 minutes, with a range of 26 to 88 minutes. The game covered a range of different puzzles that occurred in various contexts, including websites (e.g., looking for certain photos on a game character's social media page), images (e.g., decoding the hidden message of a flyer), virtual games (e.g., navigating a virtual maze), etc. (Fig. 1B). Throughout the game, participants obtained clues from some puzzles that could then be used to solve other puzzles. In some contexts (e.g., the virtual maze), participants might find more than one piece of useful information.

Temporal memory tests

Following the game, participants performed a temporal memory test both on the same day of the game (recent test) and one week after (remote test; Fig. 1A). During each memory test, participants were tested on 24 milestone events that occurred during the escape room game in a randomized order. All tested events were selected prior to data collection. These milestone events included the acquisition of key information or clues, the start of particular phases, or the resolution of specific puzzles during the game. On each test trial, participants were presented with a brief description of one event and were asked to recall when (at what time point) the event in question happened during their group's game experience by dragging a slider on a time scale from the beginning of the game (i.e., when the group started screen sharing) to the end of the game (i.e., when the group reached the congratulations page after completing all puzzles). Following each temporal memory question, participants were also tested on their memory for one episodic detail related to the event in a three-alternative forced-choice question. Given that the detail memory test was not directly related to the research questions of the current paper, we report results from this test in the Supplementary Materials. The recent and remote memory tests were identical with the same 24 events tested in the same randomized order. Both tests were self-paced.

Analyses

Scoring game contribution

In order to develop a metric of participant's in-game contribution, we first compiled the list of the 76 events/timepoints that occurred during the escape room game (Fig. S3A). For 50 of these events, an independent researcher identified at least one primary contributor who first reported having solved the step or verbally provided useful information to their group (Fig. 1B). When more than one group member contributed simultaneously to an event, all were coded as equal contributors to that event. We assigned 1 point per event to the primary contributor or equally divided the point when multiple contributors were involved. We then computed the sum of contribution points throughout the game for each participant as their game contribution score, which we operationalized as an index of agentic task engagement. To ensure reliability in contribution coding, three groups' recordings (out of 27) were coded by two researchers separately, who showed high inter-rater agreement (on average 84% of the events [42/50] were attributed to the same contributors). We thus proceeded with having each recording coded by one researcher only. The other 26 events in the game were not attributable to any contributor (e.g., opening a game file, or the start of a puzzle).

Reconstructing game timeline

To examine participants' temporal memory performance, we first gathered the "ground-truth" objective temporal information by recording the timestamps associated with all game events. In order to compare the recalled temporal information of the game events against the ground truth, for the 24 tested events (see Experimental procedures for details about events selected for the test), we separately sorted the timestamps in two distinct ways: first based on the objective timestamps during each game recording and, second, based on each participant's responses on the recent or the remote temporal memory test. Thus, for each group, we reconstructed a "ground-truth" timeline and for each participant, a memory-based timeline for each test (Fig. 1C). Timelines were then normalized to a 0-100 scale and each timestamp recoded as a percentage of total game duration, with 0 representing the start of screen sharing and 100 representing the timepoint reaching the end "congratulations" page. Comparing ground-truth and memory-based timelines allowed us to assess each participant's temporal memory performance with two measures: order memory accuracy, computed with Spearman's correlation between the recalled event order and the ground-truth event order; and timestamp recall offset, defined as the absolute difference between the recalled and the recorded timestamps (Fig. 1C).

Categorizing event pairs

In order to assess memory bias in temporal distance between different events, we first compiled all possible pairwise combinations of two events from the 24 tested events, generating a total of 276 unique event pairs (Fig. S3A & S3B left). We then categorized the event pairs into three types: unrelated events; related events that share puzzle context or source information but did *not* resolve each other's ambiguity (contextually-related); and related events where one resolved the ambiguity of the other, and were thus causally connected (causally-related; Fig. 1B). An example

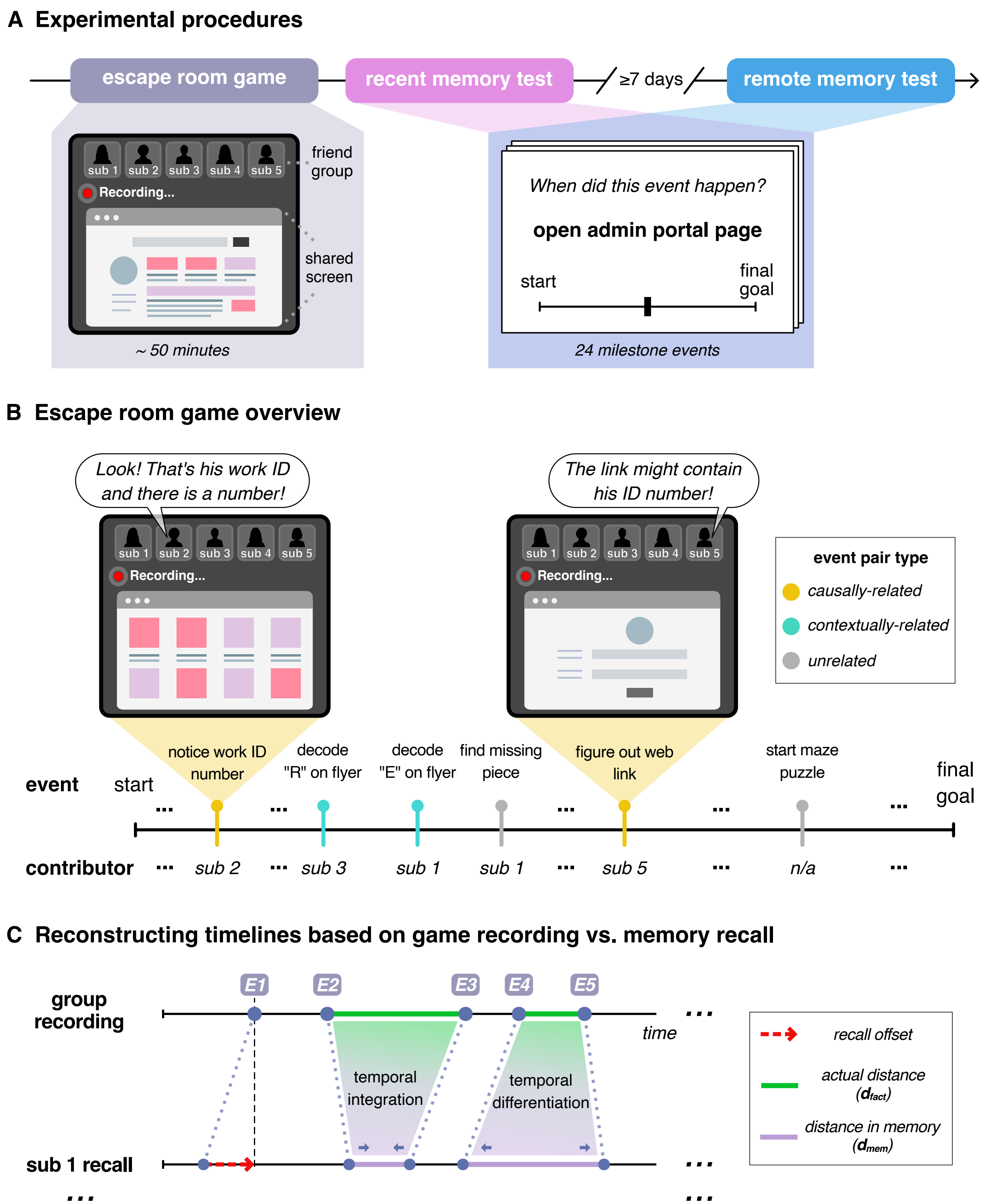


Figure 1. Experimental design. (A) Experimental procedures. Groups of friends played a virtual escape room game together, with the entire game experience recorded. Both on the same day following the game (recent) and one week after (remote), they performed a temporal memory

test during which they were asked to recall when each milestone event happened during the game by dragging a slider on a timescale. **(B)** Escape room game overview. We listed all events that occurred during the game and identified the primary contributor(s) for each event (when applicable; see Methods) who verbally provided information necessary to solve the step. We also coded the timestamps associated with all events based on the game recording. For each pairwise combination of two events that were later tested, we categorized their relationship as one of the following: unrelated events; related events that had shared context or source of information but do not resolve each other (contextually-related); related events where one resolved the ambiguity of the other (causally-related). Examples events shown for clarity and illustration purposes. **(C)** Reconstructing timelines based on game recording vs. memory recall. Across the tested events, we reconstructed separate timelines based on each group's game recording and each participant's responses on each temporal memory test. We normalized all timelines to range from 0 (when screensharing started) to 100 (when the group reached the final goal). For each individual event, we computed the temporal recall offset. For each pair of two events, we computed the temporal distance between the timestamps on the game recording timeline (d_{fact}) and the memory-based timeline (d_{mem}). We aimed to examine whether distinctive types of event pairs would be differentially represented in memory, reflected in being remembered either closer together in time (temporal integration) or farther apart (temporal differentiation).

pair of contextually-related events could be the acquisition of two clues from the same virtual maze, and an example pair of causally-related events could be the acquisition of one clue from an image (clue event) and the application of that clue on a website (target event). Out of all event pairs, there were 215 unrelated pairs, 33 contextually-related pairs, and 28 causally-related pairs (Fig. S3B left). Critically, as a natural result of the game structure, related events tended to happen closer in time than unrelated events, a factor that we took into consideration during data analysis (see Results).

Predicting pair-level temporal distance deviation between memory recall and ground truth
For each pair of tested events, we computed the ground truth temporal distance based on each group's game recording timeline (d_{fact}), and the temporal distance in memory inferred from each participant's recalled timeline on each test (d_{mem}; Fig. 1C). We then computed the memory deviation in temporal distance for that event pair (d_{mem} - d_{fact}) and performed Bayesian multi-level mixed-effects models to predict pair-level temporal distance deviation with different regressors of interest (see Results and Supplemental Tables for model specifications). In all models, we controlled for the actual temporal distance (d_{fact}), as well as the recall offset of each event in that given pair to account for general memory performance. We included all fixed-effects regressors as random-effects regressors, grouped by participants, as well as random intercepts for each friend group. Non-informative priors were specified for all regressors in the models.

To infer the existence and importance of an effect against the corresponding null hypothesis (equivalent to statistical significance in the Frequentist framework), we adopted a highest density credible interval (HDI) + region of practical equivalence (ROPE, which quantifies the range of values of a null effect) decision rule (Kruschke, 2018). Specifically, we considered an effect as "significant" when the corresponding 95% HDI falls completely outside the ROPE ranged from -0.1 to 0.1, i.e., % in ROPE = 0 (Kruschke, 2018). We additionally evaluate each effect in two steps: (1) checking whether the 95% HDI excluded zero to determine the existence of an effect (Hespanhol et al., 2019); and if zero was excluded, (2) computing the probability of direction (p_{dir}) as an index of the likelihood of a directional effect (Makowski et al., 2019), with $p_{dir} > 95\%$ interpreted as evidence for effect directionality. To evaluate the joint effect of a predictor averaged across other predictors with more than two levels, we used a leave-one-out cross-validation (loo-cv)-based model comparison approach (Vehtari et al., 2017) between the full model and a reduced model excluding the effect of interest while retaining all other predictors. We estimated the difference in expected log predictive density (Δelpd) between the full and reduced models and its standard error (SE). We considered differences with $|\Delta elpd| > 4$ as meaningful, and $|\Delta elpd|/SE > 2$ as a noteworthy improvement in model fit attributable to the effect of interest.

We opted to run these models in a Bayesian framework because it better handles complex data structure, whereas the equivalent Frequentist models failed to converge. Nevertheless, all findings were replicated when estimated with reduced Frequentist models including all fixed-effects regressors and only random intercepts for participants and groups (results not reported).

Results

Group members share high temporal memory similarity

We first confirmed whether memory for the timing of game events was more consistent across participants when it was a shared experience. To do this, we correlated each participant's recalled timestamps with those recalled by their own group members, as well as with those from participants in other groups, for both the recent and the remote tests. As expected, inter-subject temporal memory similarity was significant higher within a group than across groups ($F(1, 324)=20.822$, $p<0.001$, $\eta^2_p=0.06$; Fig. S1), which decreased from the recent test to the remote test (Fisher Z transformed Pearson's r: $F(1, 338)=72.864$, $p<0.001$, $\eta^2_p=0.18$). This result showed that while all groups played the same game, the enhanced memory similarity within groups reflected the shared nature of the group experience (see also "Inter-subject temporal memory similarity" in Supplemental Results).

Game contribution positively predicts temporal memory performance

To examine whether and how goal-directed, agentic task engagement influences temporal memory, we first tested the relationship between game contribution scores and temporal memory performance across all participants using Pearson's correlations. Temporal memory performance was measured with order memory accuracy and timestamp recall offset (see also Supplemental Results and Fig. S2 for performance across delays). We found a significant positive relationship between game contribution scores and order memory accuracy during both the recent test ($r(110)=0.232$, $p=0.014$, 95%CI[0.049, 0.401]) and remote test ($r(116)=0.322$, $p<0.001$, 95%CI[0.150, 0.475]; Fig. 2 left). Further, game contribution scores negatively predicted recall offset in both tests (recent: $r(110)=-0.207$, $p=0.029$, 95%CI[-0.378, -0.022]; remote: $r(116)=-0.274$, $p=0.003$, 95%CI[-0.433, -0.098]; Fig. 2 right). Note that one outlier was identified in the game contribution scores with a Rosner's test and thus excluded from this analysis. However, all statistical inferences remained unchanged with or without the outlier (all $p<0.023$ when including the outlier).

In addition to examining correlations across all participants, we also tested whether high and low contributors within a friend group, who shared the game experience, differed in temporal memory performance. To this end, we performed a median split in each group to identified the high (above median) and low (median or below) contributors, and compared their

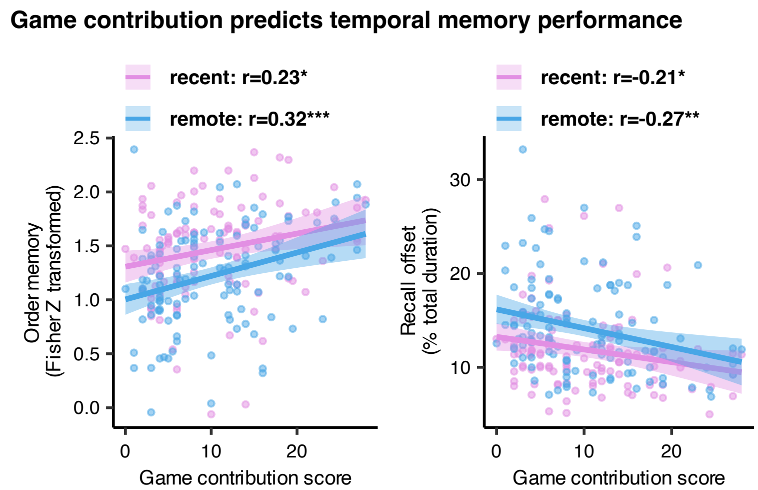


Figure 2. Game contribution predicts temporal memory performance. Pearson correlations between game contribution scores and each of order memory accuracy (left; Fisher Z transformed Spearman rho) and mean recall offset (right) separately for recent and remote tests. Each dot represents one participant; shaded ribbons represent 95% confidence intervals. *p<0.05, **p<0.01, ***p<0.001.

temporal memory performance using paired t-tests. Consistent with previous results, despite playing the game together, high contributors within a group showed significantly higher order accuracy (recent: t(98)=2.32, p=0.022, 95%CI[0.028, 0.356], Cohen's d=0.443; remote: t(103)=3.12, p=0.002, 95%CI[0.093, 0.418], Cohen's d=0.579) and lower recall offset (recent: t(105)=-2.264, p=0.026, 95%CI[-3.397, -0.225], Cohen's d=0.423; remote: t(110)=-2.689, p=0.008, 95%CI[-4.114, -0.623], Cohen's d=0.489) than the low contributors on both the recent and remote tests. Together, these findings highlighted that individuals who showed higher agentic engagement tended to have better temporal memory for the game experience.

Causally-related events were remembered closer in time than other pairs after a long delay

Our central question is whether pairs of events with different types of relations will show diverging temporal structures in memory. Specifically, we hypothesized that the causally-related, ambiguity-resolving events will show greater memory integration and thus be remembered being closer in time compared to other events. Before directly testing this hypothesis, we first assessed whether and how temporal distance in memory reflected the actual temporal distance between two events. This was a critical step because the distribution of the actual temporal distances varied drastically across different types of event pairs as a result of the game design: the majority of the related events occurred close in time (below mean d_{fact}), while unrelated events showed a wider range of temporal distances (Fig. S3B right). Here, we predicted pair-by-pair the memory deviation in temporal distance (d_{mem}-d_{fact}) with the ground truth distance (d_{fact}), test delay (recent, remote), and their interaction, while controlling for the recall offset of individual events. First, we found that distance in memory became closer than the actual distance from the recent to the remote test (negative deviation: $d_{mem} < d_{fact}$), suggesting increased temporal integration between events (posterior median[M]=-1.411, 95%HDI[-2.211, -0.597], % in ROPE=0, p_{dir}=1; Fig. S3B right). Interestingly, on both the recent and the remote tests, deviation in memory for temporal distance was negatively modulated by d_{fact} (recent: M=-7.008, 95%HDI[-7.954, -5.958], % in ROPE=0, p_{dir}=1; remote: M=-9.195, 95%HDI[-10.222, -8.143], % in ROPE=0, p_{dir}=1; Fig. S3B right; see also Table S1 for complete model output). In other words, events that were far apart in time during the game tended to be remembered as closer together than they actually were ($d_{mem} < d_{fact}$; temporal integration), whereas the events that were close in time were remembered as farther apart ($d_{mem} > d_{fact}$; temporal differentiation). Therefore, participants' memory for temporal distance between two events was biased towards the mean distance across all event pairs.

Given the uneven distribution of d_{fact} across pair types and its substantial influence on the distance deviation in memory, we applied a group-specific d_{fact} range filter to select pairs for our main analysis. This was to ensure that all three types of event pairs were well represented in the comparisons while maximizing the number of pairs included for each friend group. Specifically, for each group, we separately computed the d_{fact} range for each pair type and identified the interception range across all three types. We then selected all pairs within the interception d_{fact} range for the given group and focused on these pairs in the subsequent analyses. In this way, we selected between 60 to 120 unrelated event pairs, 27 to 33 contextually-related pairs, and 17 to 26 causally-related pairs per group (Fig. S4). Notably, the selected pairs were on the proximal side of the full d_{fact} range where the majority of related pairs fell, causing the interception range to also lie in the proximal portion of the spectrum. Consequently, the remembered distance for these selected pairs were generally greater than the ground truth distance, due to the aforementioned bias in memory for temporal distance (Fig. S3B right; see also Fig. 3).

We next turned to test our main hypothesis that the causally-related pairs would be remembered being closest in time compared to other pairs. Specifically, within the selected event pairs that were overall matched in d_{fact}, we predicted the distance deviation for each pair (d_{mem}-d_{fact}) with pair type, test delay, and their interaction, while controlling for the pair-specific d_{fact} and the recall offset of individual events. We found that during the recent test, despite d_{fact} being the same, participants remembered two unrelated events as farther apart in time (greater temporal differentiation) compared to both the contextually- and causally- related events (vs. contextually-related: M=4.243, 95%HDI[3.371, 5.090], % in ROPE=0, p_{dir}=1; vs. causally-related: M=4.996, 95%HDI[3.928, 5.963], % in ROPE=0, p_{dir}=1). However, contextually-related events showed no reliable difference from causally-related events in the distance deviation (M=0.728, 95%HDI[-0.116, 1.594], % in ROPE=4.526%; Fig. 3; see also Table S2a for complete model output). Critically, however, this pattern shifted after a one-week delay. On the remote test, unrelated pairs still showed more temporal differentiation than both types of related events (vs. contextually-related: M=1.027, 95%HDI[0.185, 1.957], % in ROPE=0, p_{dir}=0.986; vs. causally-related: M=4.660, 95%HDI[3.656, 5.704], % in ROPE=0, p_{dir}=1). However, in contrast to the recent test results, contextually-related events were remembered as farther apart than the causally-related events (M=3.627, 95%HDI[2.774, 4.478], % in ROPE=0, p_{dir}=1). This finding suggests that, with consolidation, temporal integration for causally connected memories is greater than those that merely shared similar encoding context. To assess whether the pattern observed on the recent test statistically differed from that on the remote test, we performed model comparison between the full model described above and a reduced model that specifically excluded the interactive effect between pair type and test delay. We found that the full model substantially outperformed the reduced model ($|\Delta elpd|$= 45.1, SE=10.4), highlighting

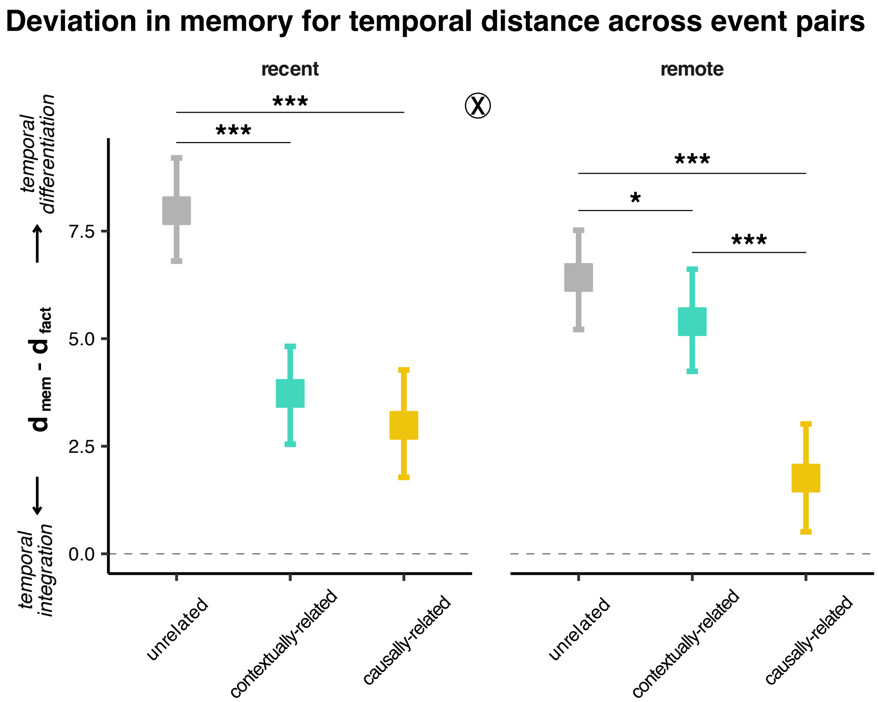


Figure 3. Deviation in memory for temporal distance across event pairs. We performed a Bayesian mixed-effects model predicting the deviation in the distance between the recalled timestamps (d_{mem}) and between the ground truth timestamps (d_{fact}) with event pair type, test delay, and their interaction. We also controlled for the corresponding d_{fact} and recall offset of individual events in the model (not depicted and held constant at the mean level). Squares represent the model-predicted posterior medians; error bars represent 95% highest density credible intervals (HDI). *: % in ROPE=0, 95%HDI excluded 0, and p_{dir}>0.95; ***: % in ROPE=0, 95%HDI excluded 0, and p_{dir}>0.999; ⊗ interaction drove superior model fit indicated by |Δelpd|>4 and |Δelpd|/SE>2.

consolidation-dependent changes in the temporal organizations of memories for different event pairs.

Cross-event memory representations differ between high and low in-game contributors

Thus far, we have demonstrated that game contribution scores predicted temporal memory performance, and that cross-event temporal memory structure varied across event pair types and changed after a long delay. Building on these results, we next asked whether, beyond differences in performance, participants with different levels of in-game contribution might also exhibited distinctive cross-event memory representations. In particular, we examined whether such differences might be observed among group members who shared the same game experience. To test this, we identified the high and low contributors within each group using a median-split, and performed separate models within the high and low contributors predicting memory deviation in temporal distance ($d_{mem} - d_{fact}$) with the same regressors as the model conducted in the full sample.

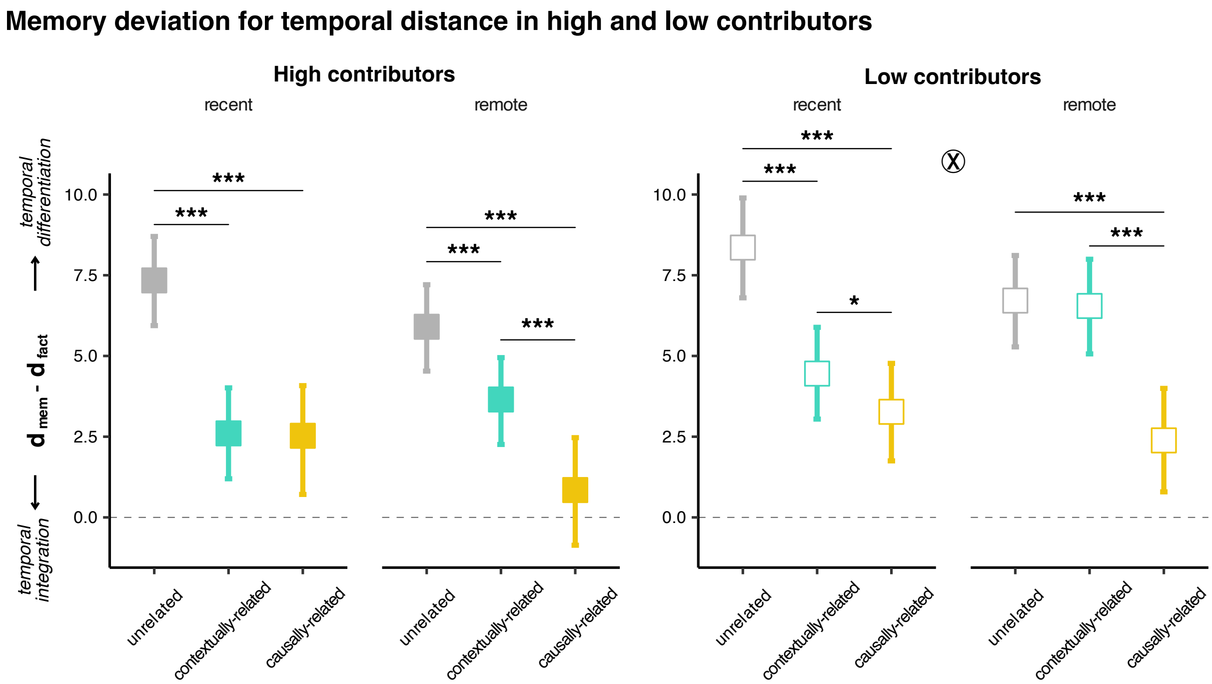


Figure 4. Memory deviation for temporal distance in high vs. low contributors. We performed separate Bayesian mixed-effects models within the high and low contributors of each group predicting the memory deviation in temporal distance (d_{mem} - d_{fact}) with event pair type, test delay, and their interaction. We also controlled for the corresponding d_{fact} and recall offset of individual events in the model (not depicted and held constant at the mean level). Squares represent the model-predicted posterior medians; error bars represent 95% highest density credible intervals (HDI). *: % in ROPE=0, 95%HDI excluded 0, and p_{dir}>0.95; ***: % in ROPE=0, 95%HDI excluded 0, and p_{dir}>0.999; ⊗ interaction drove superior model fit indicated by |Δelpd|>4 and |Δelpd|/SE>2.

We found that the pattern of results in the high contributors resembled that in the full sample: with d_{fact} being the same, on the recent test, unrelated events were remembered as farther apart than the related events (vs. contextually-related: M=4.752, 95%HDI[3.572, 6.013], % in ROPE=0, p_{dir}=1; vs. causally-related: M=4.846, 95%HDI[3.231, 6.665], % in ROPE=0, p_{dir}=1), while no difference was found between the two types of related events (M=0.091, 95%HDI[-1.315, 1.519], % in ROPE=11.842%); on the remote test, the causally-related events were remembered as being closest in time, followed by the contextually-related events, and unrelated events were remembered as being farthest (% in ROPE =0, 95% HDI excluded zero, and p_{dir}>0.999 for all pairwise comparisons; Fig.4; see also Table S3a&S3b for complete model output). However, a different pattern was noted within the low contributors: on the recent test, while they also remembered the unrelated pairs being farther than the related pairs as the high contributors did, their memory for the contextually-related pairs showed relatively less temporal integration than the causally-related pairs (% in ROPE =0, 95% HDI excluded zero, and p_{dir}>0.985 for all pairwise comparisons); in contrast, on the remote test, low contributors’

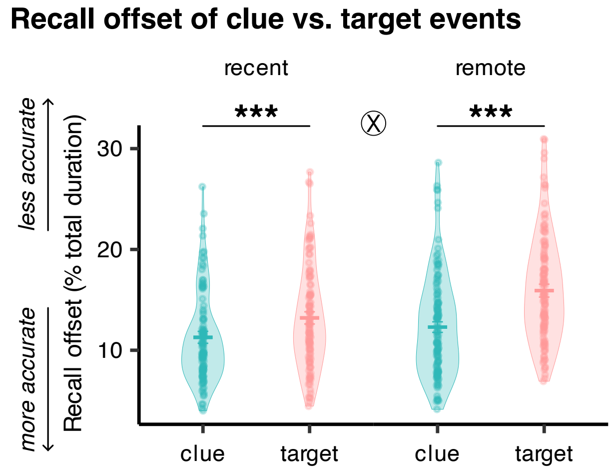


Figure 5. Recall offset of clue vs. target events. Recall offset for clue and target events within the causally-related pairs on the recent and remote tests. Each dot represents one participant; thick short lines represent the mean values; error bars represent 95% confidence intervals. ***p<0.001, ⊗ significant interaction.

memory for the contextually-related events even showed similar levels of temporal differentiation as that of the unrelated events (M=0.173, 95%HDI[-0.921, 1.307], % in ROPE=13.868%), while the causally-related events were still remembered as being closest together (vs. unrelated: M=4.364 95%HDI[3.178, 5.533], % in ROPE=0, p_{dir}=1; vs. contextually-related: M=4.183, 95%HDI[3.215, 5.180], % in ROPE=0, p_{dir}=1; Fig.4; see also Table S3c&S3d for complete model output). These results suggested that compared to the unrelated events, the temporal structure of memories for the contextually similar events in particular depended on the levels of goal-directed, agentic engagement during the game, whereas memory for the causally-linked events consistently showed more temporal integration regardless of in-game performance.

We also examined whether the interaction between pair type and test delay statistically improved model fit separately for the high contributor model and the low contributor model. In both models, predictive performance was better when the interaction term was included, although the evidence for a superior fit was weak in the high contributor model (|Δelpd|=8.3, SE=5.1), but noteworthy in the low contributor model (|Δelpd|=33.5, SE=8.7), suggesting a considerable shift in temporal memory structure across consolidation within the low contributors.

More accurate and persistent temporal memory for clue events than target events

Building upon the finding that, after consolidation, causally-related pairs were remembered closer in time compared to other event pairs, we asked whether within these causally-related events, temporal memory might also differ between the clue events when the clues were obtained and the target events when the clues were used to resolve ambiguity. This analysis may help reveal potential asymmetries in the memory representation changes that led to their relatively greater temporal integration. To this end, we computed the mean recall offset for clue and target events in all causally-related pairs on both the recent and remote tests. We found significantly smaller recall offset (i.e., higher accuracy) for clue events than the target events ($F(1, 337.13)=84.958$, $p<0.001$, $\eta^2_p=0.20$; Fig. 4), both on the recent ($b=-1.929$, $p<0.001$, 95%CI[-2.78, -1.08]) and on the remote test ($b=-3.617$, $p<0.001$, 95%CI[-4.44, -2.79]). Further, while recall offset increased from the recent to the remote test as expected ($F(1, 348.60)=43.814$, $p<0.001$, $\eta^2_p=0.11$), this increase was significantly greater for the target events compared to the clue events (event type × test delay interaction: $F(1, 337.13)=7.868$, $p=0.005$, $\eta^2_p=0.02$), suggesting that temporal memory for the clue events was more persistent than the target events. Given that some events may have simultaneously served as clues in some pairs but targets in other pairs (e.g., a clue found in a previous puzzle may lead to the discovery of a keyword, and the keyword also served as the clue for another puzzle), we also repeated this analysis within the exclusive clue events and exclusive target events during the game, and replicated all effects (event type main effect: $F(1, 336.73)= 89.3966$, $p<0.001$, $\eta^2_p=0.21$; test delay main effect: $F(1, 350.75)= 51.1851$, $p<0.001$, $\eta^2_p=0.13$; event type × test delay interaction: $F(1, 336.73)= 6.2794$, $p=0.013$, $\eta^2_p=0.02$).

Discussion

Our memory for temporal intervals is often distorted. For example, you may be surprised to realize that it has been five years since the last time you saw a friend although it feels like only several months have passed, whereas the hike from last week seems like a long time ago. To study what drives temporal distortions in memory, we examine how two factors – cross-event interconnection and agentic engagement – shape subsequent temporal memory structure in a virtual escape room game. Controlling for actual temporal distance, we found that related events were remembered as being closer in time than the unrelated ones. Critically, after a period of consolidation, memories for causally-related events that resolved each other’s ambiguity became more temporally integrated than other related events that merely shared contextual similarity, suggesting that inferential bridges or causal connections drove greater and more persistent memory integration. Moreover, temporal memory varied across people and was influenced by the degree of their agentic engagement during the game. Specifically, those who contributed more to puzzle-solving showed superior temporal memory performance compared to low contributors. High agentic engagement also enhanced memory integration particularly between contextually-related events. Together, these findings demonstrate that temporal memory structure is shaped by both the objective relationships among events and the degree of agentic engagement experienced during these events.

Existing research has suggested that shared contextual information or event features across memories promotes the formation of an integrated representation (Clewett et al., 2019; Schlichting & Preston, 2015). The current study extends this literature by demonstrating enhanced integration between causally-linked memories for inferential-bridging events. Critically, we highlighted that following consolidation, causally-related events showed greater memory integration than even events related via contextual overlap, evidenced by a stronger bias towards temporal proximity in memory. This increased integration driven by causality may be achieved via two mechanisms. First, effectively retrieving and applying useful information is a core element of an escape room game. Reactivation of relevant information obtained from prior events during the encoding of new events may strengthen the connections between these memories (Hulbert & Norman, 2015; Ritvo et al., 2019; Schlichting & Preston, 2015; Zeithamova et al., 2012). This idea is also supported by recent narrative comprehension studies showing that reactivating causally-related events during new events leads to increased neural activity pattern similarity between these events (Cohn-Sheehy et al., 2021; Milivojevic et al., 2015). Here, we demonstrate in a real-life social game that this representational change is reflected in the temporal organization of these memories. Second, events sharing causal connections may also be prioritized during offline processing via mechanisms such as memory replay. Existing research has shown that events central in the causal chains are considered as more important (Trabasso & Sperry, 1985) and are better remembered than other events (Black & Bern, 1981; Lee & Chen, 2022). It is possible that the ambiguity-resolving events, given their inferential links critical to the puzzle-solving experience, were not only more replayed during consolidation given their importance (Cowan et al., 2021; Tambini & Davachi, 2013), but also replayed in an overlapping manner which facilitated the reorganization of these memories (Lewis & Durrant, 2011). Our results also support and extend beyond the notion that offline consolidation promotes memory integration (Audrain & McAndrews, 2022; Tompary & Davachi, 2017, 2024) by showing that the effect of consolidation on integration varied across memories that shared different types of interconnection.

Interestingly, within causally-related events, participants showed more precise and persistent memory for the timepoints when clues were found compared to their corresponding target events where the clues were used. This finding indicates that increased integration between causally-linked memories may be driven by a greater shift in memory representation of the target events than the clue events. This asymmetry reflects the special status of the clue events in the game. Specifically, clue events tended to carry latent affordances: when participants first encountered the clues during the game, their specific use and the actions they would afford were often not revealed until the target events occurred. Therefore, to be able to readily access the clue information for puzzle-solving, participants might store the associated events as anchor points in memory, hence the superior temporal recall.

While causally-linked memories showed consistently more integration compared to unrelated memories, temporal structure of contextually-related memories depended on agentic engagement. Immediately after the game, high contributors remembered contextually-related events being as close in time as causally-related events, while low contributors already showed greater differentiation for contextually- versus causally- related memories. Following consolidation, high contributors still showed greater temporal integration between contextually-related events than unrelated events, but no difference was found in low contributors between these two pairs types. Extensive research has demonstrated that contextual continuity or overlap helps bridge information across memories (Clewett et al., 2019; Cox et al., 2021; DuBrow & Davachi, 2013, 2014; Ezzyat & Davachi, 2014; Liu et al., 2024; Yu et al., 2025; Zacks, 2007). Here, our finding suggests that low task engagement attenuates this effect. Compared to high contributors, low contributors often passively watched other group members solving puzzles rather than actively participated in decision-making during the game. The reduced agency within the low contributors led to not only worse temporal memory, but also weaker connections between certain memories compared to high contributors. These results are consistent with existing work showing that agency enhances memory and strengthens encoded associations (Houser et al., 2022; Li et al., 2024; Murty et al., 2015, 2019; Ruiz et al., 2023). Notably, high levels of agentic engagement likely involved greater attention to the puzzles, which may also modulate information binding across events (Jayakumar et al., 2024). Prior work has demonstrated that when shared context across temporally distant events was only incidentally encoded, the effect of overlapping context on memory integration was weaker compared to when the contextual information was actively incorporated into the encoding experience (Cox et al., 2021; Yu et al., 2025). It is possible that causal links between events robustly and, perhaps, more automatically drive memory integration, whereas memories that share contextual overlap require extra cognitive resources to enable long-lasting integration.

Unlike conventional in-lab studies with highly-controlled paradigms, the current experiment focused on a real-life experience shared with familiar others. Some recent work has also adopted naturalistic paradigms to study temporal memory, including duration judgements for movie events (Lositsky et al., 2016) and sequential memory for experiences like a museum art tour (Diamond et al., 2025) or a haunted house visit (Cliver et al., 2025). Our study differed from them in a few ways. First, an escape room involves engaging problem-solving tasks that demand more active engagement compared to passively watching or listening to naturalistic stimuli. Second, by collecting memory data from individuals who played the game together, we leveraged the social nature of this activity to study how memories might be similar or different across those who shared the same experience and task goals. In addition to revealing the effect of agentic engagement on temporal memory structure as aforementioned, we showed that participants shared more similar temporal memory with their group members than those from other groups, consistent with the notion that mutual goals within a social context bias memory to align across people (Echterhoff et al., 2009; Higgins et al., 2021). We also demonstrated that temporal memory within a group tended to center around those who contributed more in the group, driven by their precise memory for the experience (see Supplemental Results). The social aspect of the current experiment thus provided insights into collective memory in real-world, collaborative scenarios.

Nevertheless, compared to highly-controlled experiments, the use of a naturalistic paradigm comes with limitations. In our study, due to the game structure, related events tended to occur closer in time than many unrelated events. While we carefully accounted for this and selected pairs matched in temporal distance for our main analyses, these pairs were distributed around the proximal end of the distance spectrums, with distances ranging from 0 to 52% of total game duration (approximately up to 22 minutes apart in real time). Therefore, future work is needed to explore whether our findings generalize to interrelated events spanning a wider range of temporal distances.

In summary, leveraging a real-life social game, we showed that memories for causally-connected events were not only more integrated than unrelated events, but following consolidation, more integrated than other related events that only shared contextual similarity. Interestingly, within these causally-connected, ambiguity-resolving event pairs, temporal memory was more precise for events where clues were found compared to the events where the clues were used. Further, people's temporal memory depended on how much they contributed during the group task. High contributors not only showed more accurate temporal memory than low contributors, but also demonstrated superior temporal binding between contextually-related events. Together, our findings shed light on the temporal organization and representational structure of related memories in real-world experiences.

Acknowledgement
We thank Arielle Clarke, Jessie Lin, Corinna Jones, and Anish Nanjappa for their assistance with coding the game video recordings.

References

Aly, M., & Turk-Browne, N. B. (2016). Attention promotes episodic encoding by stabilizing hippocampal representations. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(4), E420–E429. https://doi.org/10.1073/pnas.1518931113

Audrain, S., & McAndrews, M. P. (2022). Schemas provide a scaffold for neocortical integration of new memories over time. *Nature Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-022-33517-0

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, *95*(3), 709-721.e5. https://doi.org/10.1016/j.neuron.2017.06.041

Bein, O., & Davachi, L. (2024). Event Integration and Temporal Differentiation: How Hierarchical Knowledge Emerges in Hippocampal Subfields through Learning. *Journal of Neuroscience*, *44*(10). https://doi.org/10.1523/JNEUROSCI.0627-23.2023

Black, J. B., & Bern, H. (1981). Causal coherence and memory for events in narratives. *Journal of Verbal Learning and Verbal Behavior*, *20*(3), 267–275. https://doi.org/10.1016/S0022-5371(81)90417-5

Block, R. A., & Reed, M. A. (1978). Remembered duration: Evidence for a contextual-change hypothesis. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(6), 656–665. https://doi.org/10.1037/0278-7393.4.6.656

Chang, C. H. C., Lazaridi, C., Yeshurun, Y., Norman, K. A., & Hasson, U. (2021). Relating the Past with the Present: Information Integration and Segregation during Ongoing Narrative Processing. *Journal of Cognitive Neuroscience*, *33*(6), 1106–1128. https://doi.org/10.1162/jocn_a_01707

Chen, J., & Bornstein, A. M. (2024). The causal structure and computational value of narratives. *Trends in Cognitive Sciences*, *28*(8), 769–781. https://doi.org/10.1016/j.tics.2024.04.003

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A Taxonomy of external and internal attention. *Annual Review of Psychology*, *62*(1), 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

Clewett, D., DuBrow, S., & Davachi, L. (2019). Transcending time in the brain: How event memories are constructed from experience. *Hippocampus*, *29*(3), 162–183. https://doi.org/10.1002/hipo.23074

Clewett, D., Gasser, C., & Davachi, L. (2020). Pupil-linked arousal signals track the temporal organization of events in memory. *Nature Communications*, *11*(1), 4007. https://doi.org/10.1038/s41467-020-17851-9

Clewett, D., Huang, R., & Davachi, L. (2025). Locus coeruleus activation "resets" hippocampal event representations and separates adjacent memories. *Neuron*, *113*(15), 2521-2535.e8. https://doi.org/10.1016/j.neuron.2025.05.013

Cliver, K. G., Gregory, D. F., Martinez, S. A., Mitchell, W. J., Stasiak, J. E., Reisman, S. S., Helion, C., & Murty, V. P. (2025). Temporal memory for threatening events encoded in a haunted house. *Cognition and Emotion*, *39*(1), 65–81. https://doi.org/10.1080/02699931.2024.2338962

Cohn-Sheehy, B. I., Delarazan, A. I., Reagh, Z. M., Crivelli-Decker, J. E., Kim, K., Barnett, A. J., Zacks, J. M., & Ranganath, C. (2021). The hippocampus constructs narrative memories across distant events. *Current Biology*, *31*(22), 4935-4945.e7. https://doi.org/10.1016/j.cub.2021.09.013

Corriveau, A., Chao, A. F., deBettencourt, M. T., & Rosenberg, M. D. (2025). Recognition memory fluctuates with sustained attention regardless of task relevance. *Psychonomic Bulletin & Review*, *32*(2), 714–728. https://doi.org/10.3758/s13423-024-02560-x

Corriveau, A., James Jr., A. R., deBettencourt, M. T., & Rosenberg, M. D. (2025). Sustained attentional state is a floodlight not a spotlight. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001769

Cowan, E. T., Schapiro, A. C., Dunsmoor, J. E., & Murty, V. P. (2021). Memory consolidation as an adaptive process. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-021-01978-x

Cox, W. R., Dobbelaar, S., Meeter, M., Kindt, M., & Ast, V. A. van. (2021). Episodic memory enhancement versus impairment is determined by contextual similarity across events. *Proceedings of the National Academy of Sciences*, *118*(48). https://doi.org/10.1073/pnas.2101509118

Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: The role of prediction and context. *Trends in Cognitive Sciences*, *19*(2), 92–99. https://doi.org/10.1016/j.tics.2014.12.004

DeBettencourt, M. T., Norman, K. A., & Turk-Browne, N. B. (2018). Forgetting from lapses of sustained attention. *Psychonomic Bulletin and Review*, *25*(2), 605–611. https://doi.org/10.3758/s13423-017-1309-5

Diamond, N. B., Simpson, S., Baena, D., Murray, B., Fogel, S., & Levine, B. (2025). Sleep selectively and durably enhances memory for the sequence of real-world experiences. *Nature Human Behaviour*, *9*(4), 746–757. https://doi.org/10.1038/s41562-025-02117-5

DuBrow, S., & Davachi, L. (2013). The influence of context boundaries on memory for the sequential order of events. *Journal of Experimental Psychology: General*, *142*(4), 1277–1286. https://doi.org/10.1037/a0034024

DuBrow, S., & Davachi, L. (2014). Temporal memory is shaped by encoding stability and intervening item reactivation. *Journal of Neuroscience*, *34*(42), 13998–14005. https://doi.org/10.1523/jneurosci.2535-14.2014

DuBrow, S., & Davachi, L. (2016). Temporal binding within and across events. *Neurobiology of Learning and Memory*, *134*, 107–114. https://doi.org/10.1016/j.nlm.2016.07.011

Echterhoff, G., Higgins, E. T., & Levine, J. M. (2009). Shared Reality: Experiencing Commonality With Others' Inner States About the World. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *4*(5), 496–521. https://doi.org/10.1111/j.1745-6924.2009.01161.x

Ezzyat, Y., & Davachi, L. (2011). What constitutes an episode in episodic memory? *Psychological Science*, *22*(2), 243–252. https://doi.org/10.1177/0956797610393742

Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: Pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron*, *81*(5), 1179–1189. https://doi.org/10.1016/j.neuron.2014.01.042

Hespanhol, L., Vallio, C. S., Costa, L. M., & Saragiotto, B. T. (2019). Understanding and interpreting confidence and credible intervals around effect estimates. *Brazilian Journal of Physical Therapy*, *23*(4), 290–301. https://doi.org/10.1016/j.bjpt.2018.12.006

Higgins, E. T., Rossignac-Milon, M., & Echterhoff, G. (2021). Shared reality: From sharing-is-believing to merging minds. *Current Directions in Psychological Science*, *30*(2), 103–110. https://doi.org/10.1177/0963721421992027

Houser, T. M., Tompary, A., & Murty, V. P. (2022). Agency enhances temporal order memory in an interactive exploration game. *Psychonomic Bulletin & Review*, *29*(6), 2219–2228. https://doi.org/10.3758/s13423-022-02152-7

Hulbert, J. C., & Norman, K. A. (2015). Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. *Cerebral Cortex*, *25*(10), 3994–4008. https://doi.org/10.1093/cercor/bhu284

Jayakumar, M., Tsegai-Moore, A., Nemeth, C., & Aly, M. (2024). *Pupil size at encoding predicts the temporal structure of memory*. OSF. https://doi.org/10.31234/osf.io/bzqfu

Kruschke, J. K. (2018). Rejecting or Accepting Parameter Values in Bayesian Estimation. *Advances in Methods and Practices in Psychological Science*, *1*(2), 270–280. https://doi.org/10.1177/2515245918771304

Lee, H., & Chen, J. (2022). Predicting memory from the network structure of naturalistic events. *Nature Communications*, *13*(1), 4235. https://doi.org/10.1038/s41467-022-31965-2

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*, *15*(8), 343–351. https://doi.org/10.1016/j.tics.2011.06.004

Li, X., Ni, N., Born, S., Gualano, R., Lee, I., Bellana, B., & Chen, J. (2024). *Agency personalizes episodic memories*. OSF. https://doi.org/10.31234/osf.io/7evwj

Liu, Z., Johansson, M., Johansson, R., & Bramão, I. (2024). The effects of episodic context on memory integration. *Scientific Reports*, *14*(1), 30159. https://doi.org/10.1038/s41598-024-82004-7

Lositsky, O., Chen, J., Toker, D., Honey, C. J., Shvartsman, M., Poppenk, J. L., Hasson, U., & Norman, K. A. (2016). Neural pattern change during encoding of a narrative predicts retrospective duration estimates. *eLife*, *5*, e16070. https://doi.org/10.7554/eLife.16070

Makowski, D., Ben-Shachar, M. S., Chen, S. H. A., & Lüdecke, D. (2019). Indices of Effect Existence and Significance in the Bayesian Framework. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02767

Milivojevic, B., Vicente-Grabovetsky, A., & Doeller, C. F. (2015). Insight reconfigures hippocampal-prefrontal memories. *Current Biology*, *25*(7), 821–830. https://doi.org/10.1016/j.cub.2015.01.033

Murty, V. P., DuBrow, S., & Davachi, L. (2015). The simple act of choosing influences declarative memory. *Journal of Neuroscience*, *35*(16), 6255–6264. https://doi.org/10.1523/jneurosci.4181-14.2015

Murty, V. P., DuBrow, S., & Davachi, L. (2019). Decision-making Increases Episodic Memory via Postencoding Consolidation. *Journal of Cognitive Neuroscience*, *31*(9), 1308–1317. https://doi.org/10.1162/jocn_a_01321

Ritvo, V. J. H., Turk-Browne, N. B., & Norman, K. A. (2019). Nonmonotonic Plasticity: How Memory Retrieval Drives Learning. *Trends in Cognitive Sciences*, *23*(9), 726–742. https://doi.org/10.1016/j.tics.2019.06.007

Ruiz, N. A., DuBrow, S., & Murty, V. P. (2023). Agency as a bridge to form associative memories. *Journal of Experimental Psychology: General*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/xge0001356

Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, *6*, 1–10. https://doi.org/10.1038/ncomms9151

Schlichting, M. L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, *1*, 1–8. https://doi.org/10.1016/j.cobeha.2014.07.005

Sherman, B. E., DuBrow, S., Winawer, J., & Davachi, L. (2023). Mnemonic Content and Hippocampal Patterns Shape Judgments of Time. *Psychological Science*, *34*(2), 221–237. https://doi.org/10.1177/09567976221129533

Shohamy, D., & Wagner, A. D. (2008). Integrating Memories in the Human Brain: Hippocampal-Midbrain Encoding of Overlapping Events. *Neuron*, *60*(2), 378–389. https://doi.org/10.1016/j.neuron.2008.09.023

Silver, B. M., Yu, W., Bolger, N., Davachi, L., & Ochsner, K. N. (2025). What Are My Friends Really Like? How We Change Our Perceptions of Familiar Others' Traits and Actions. *Personality and Social Psychology Bulletin*, 01461672251320476. https://doi.org/10.1177/01461672251320476

Song, H., Finn, E. S., & Rosenberg, M. D. (2021). Neural signatures of attentional engagement during narratives and its consequences for event memory. *Proceedings of the National Academy of Sciences*, *118*(33). https://doi.org/10.1073/pnas.2021905118

Song, H., Ke, J., Madhogarhia, R., Leong, Y. C., & Rosenberg, M. D. (2025). *Cortical reinstatement of causally related events sparks narrative insights by updating neural representation patterns* (p. 2025.03.12.642853). bioRxiv. https://doi.org/10.1101/2025.03.12.642853

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event Boundaries in Perception Affect Memory Encoding and Updating. *Journal of Experimental Psychology. General*, *138*(2), 236. https://doi.org/10.1037/a0015631

Tambini, A., & Davachi, L. (2013). Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(48), 19591–19596. https://doi.org/10.1073/pnas.1308499110

Tompary, A., & Davachi, L. (2017). Consolidation Promotes the Emergence of Representational Overlap in the Hippocampus and Medial Prefrontal Cortex. *Neuron*, *96*(1), 228-241.e5. https://doi.org/10.1016/j.neuron.2017.09.005

Tompary, A., & Davachi, L. (2024). Integration of overlapping sequences emerges with consolidation through medial prefrontal cortex neural ensembles and hippocampal–cortical connectivity. *eLife*, *13*, e84359. https://doi.org/10.7554/eLife.84359

Trabasso, T., & Sperry, L. L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language*, *24*(5), 595–611. https://doi.org/10.1016/0749-596X(85)90048-8

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

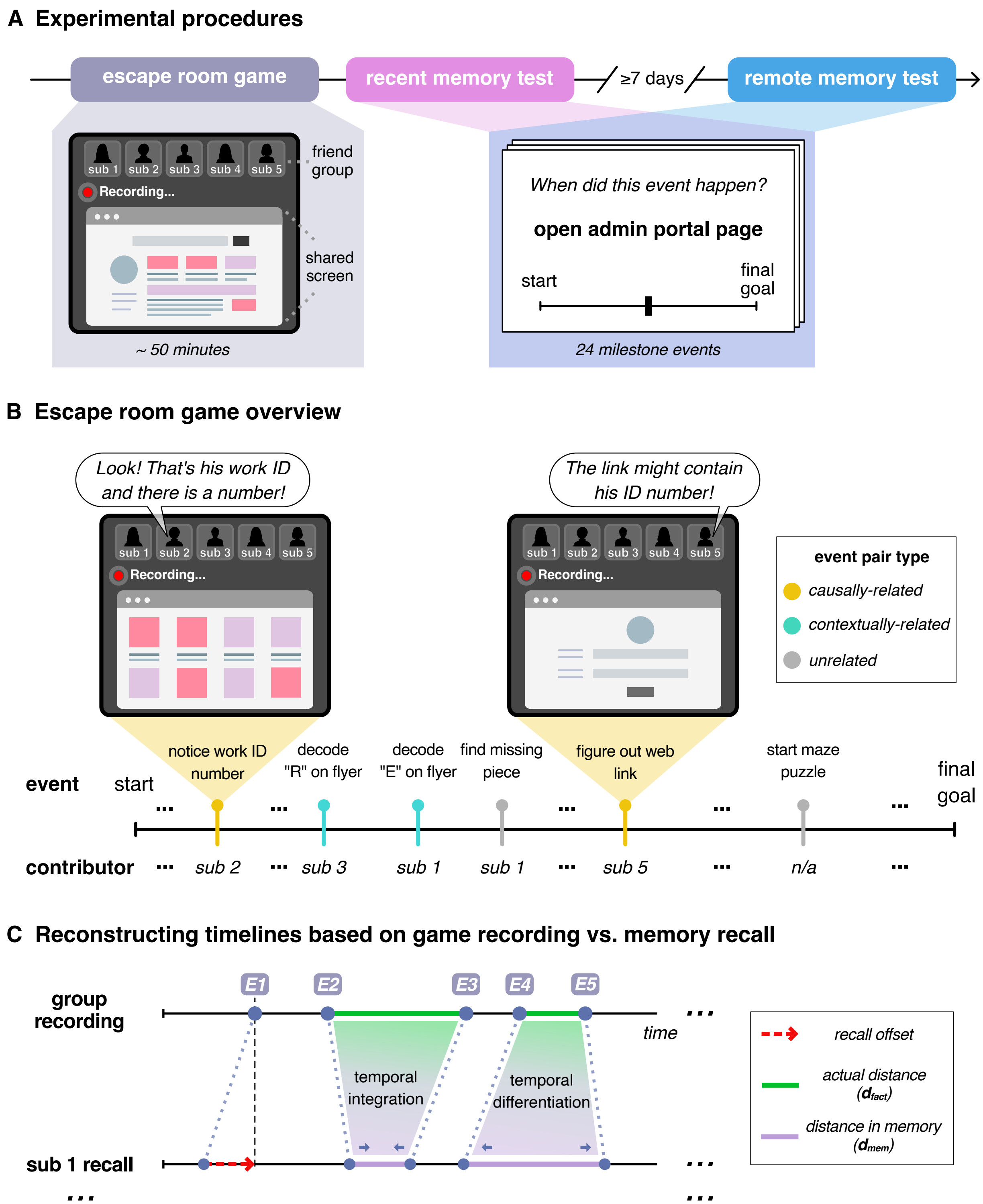
Yu, W., Duncan, K. D., & Schlichting, M. L. (2025). Using retrieval contingencies to understand memory integration and inference. *Memory & Cognition*. https://doi.org/10.3758/s13421-025-01727-8

Zacks, J. M. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, *133*(2), 273. https://doi.org/10.1037/0033-2909.133.2.273

Zakay, D., Tsal, Y., Moses, M., & Shahar, I. (1994). The role of segmentation in prospective and retrospective time estimation processes. *Memory & Cognition*, *22*(3), 344–351. https://doi.org/10.3758/BF03200861

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron*, *75*(1), 168–179. https://doi.org/10.1016/j.neuron.2012.05.010

Figures



A Experimental procedures
escape room game
recent memory test
≥7 days
remote memory test
sub 1
sub 2
sub 3
sub 4
sub 5
friend group
Recording...
shared screen
~ 50 minutes
When did this event happen?
open admin portal page
start
final goal
24 milestone events
B Escape room game overview
Look! That's his work ID and there is a number!
The link might contain his ID number!
event pair type
causally-related
contextually-related
unrelated
event
start
notice work ID number
decode "R" on flyer
decode "E" on flyer
find missing piece
figure out web link
start maze puzzle
final goal
contributor
sub 2
sub 3
sub 1
sub 1
sub 5
n/a
C Reconstructing timelines based on game recording vs. memory recall
group recording
E1
E2
E3
E4
E5
time
temporal integration
temporal differentiation
sub 1 recall
recall offset
actual distance (d_{fact})
distance in memory (d_{mem})

Figure 1. Experimental design. (A) Experimental procedures. Groups of friends played a virtual escape room game together, with the entire game experience recorded. Both on the same day following the game (recent) and one week after (remote), they performed a temporal memory test during which they were asked to recall when each milestone event happened during the game by dragging a slider on a timescale. **(B)** Escape room game overview. We listed all events that occurred during the game and identified the primary contributor(s) for each event (when applicable; see Methods) who verbally provided information necessary to solve the step. We also coded the timestamps associated with all events based on the game recording. For each pairwise combination of two events that were later tested, we categorized their relationship as one of the following: unrelated events; related events that had shared context or source of information but do not resolve each other (contextually-related); related events where one resolved the ambiguity of the other (causally-related). Examples events shown for clarity and illustration purposes. **(C)** Reconstructing timelines based on game recording vs. memory recall. Across the tested events, we reconstructed separate timelines based on each group's game recording and each participant's responses on each temporal memory test. We normalized all timelines to range from 0 (when screensharing started) to 100 (when the group reached the final goal). For each individual event, we computed the temporal recall offset. For each pair of two events, we computed the temporal distance between the timestamps on the game recording timeline (d_{fact}) and the memory-based timeline (d_{mem}). We aimed to examine whether distinctive types of event pairs would be differentially represented in memory, reflected in being remembered either closer together in time (temporal integration) or farther apart (temporal differentiation).

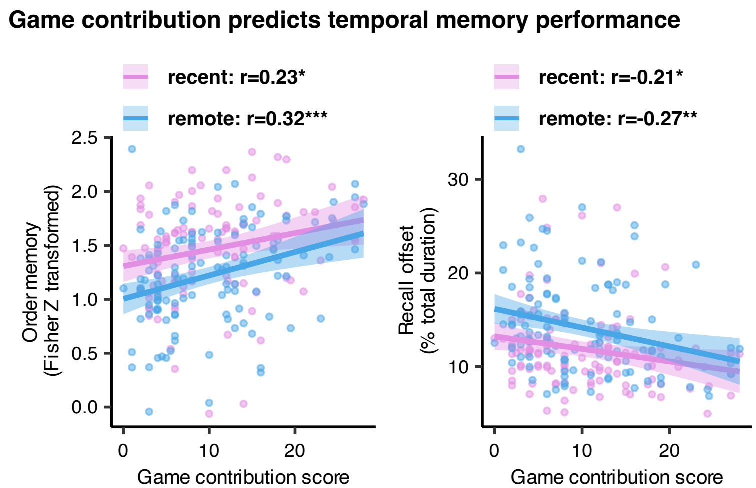


Figure 2. Game contribution predicts temporal memory performance. Pearson correlations between game contribution scores and each of order memory accuracy (left; Fisher Z transformed Spearman rho) and mean recall offset (right) separately for recent and remote tests. Each dot represents one participant; shaded ribbons represent 95% confidence intervals. *p<0.05, **p<0.01, ***p<0.001.

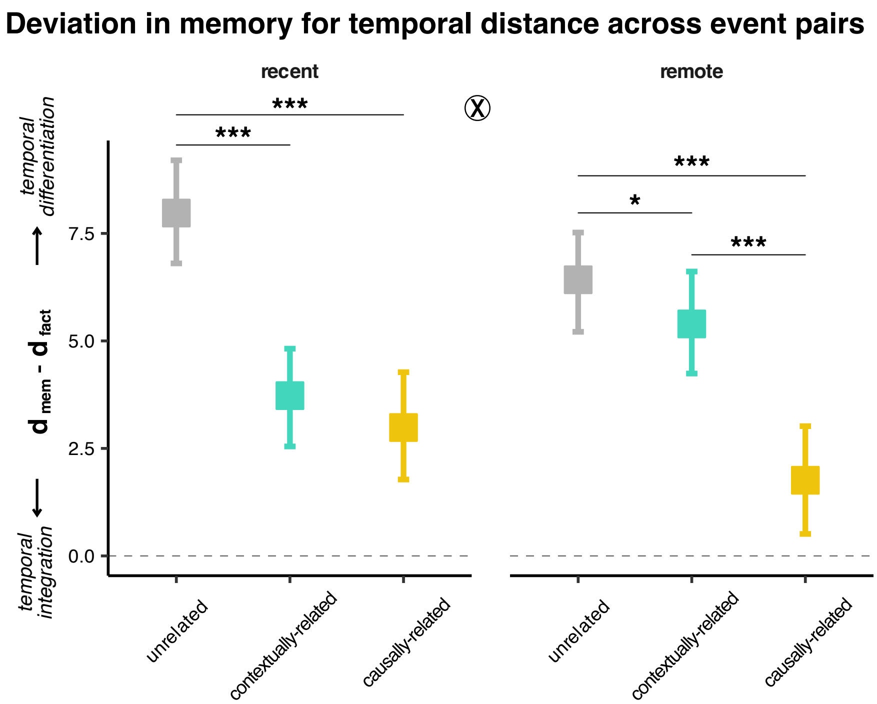


Figure 3. Deviation in memory for temporal distance across event pairs. We performed a Bayesian mixed-effects model predicting the deviation in the distance between the recalled timestamps (d_{mem}) and between the ground truth timestamps (d_{fact}) with event pair type, test delay, and their interaction. We also controlled for the corresponding d_{fact} and recall offset of individual events in the model (not depicted and held constant at the mean level). Squares represent the model-predicted posterior medians; error bars represent 95% highest density credible intervals (HDI). *: % in ROPE=0, 95%HDI excluded 0, and p_{dir}>0.95; ***: % in ROPE=0, 95%HDI excluded 0, and p_{dir}>0.999; ⊗ interaction drove superior model fit indicated by |Δelpd|>4 and |Δelpd|/SE>2.

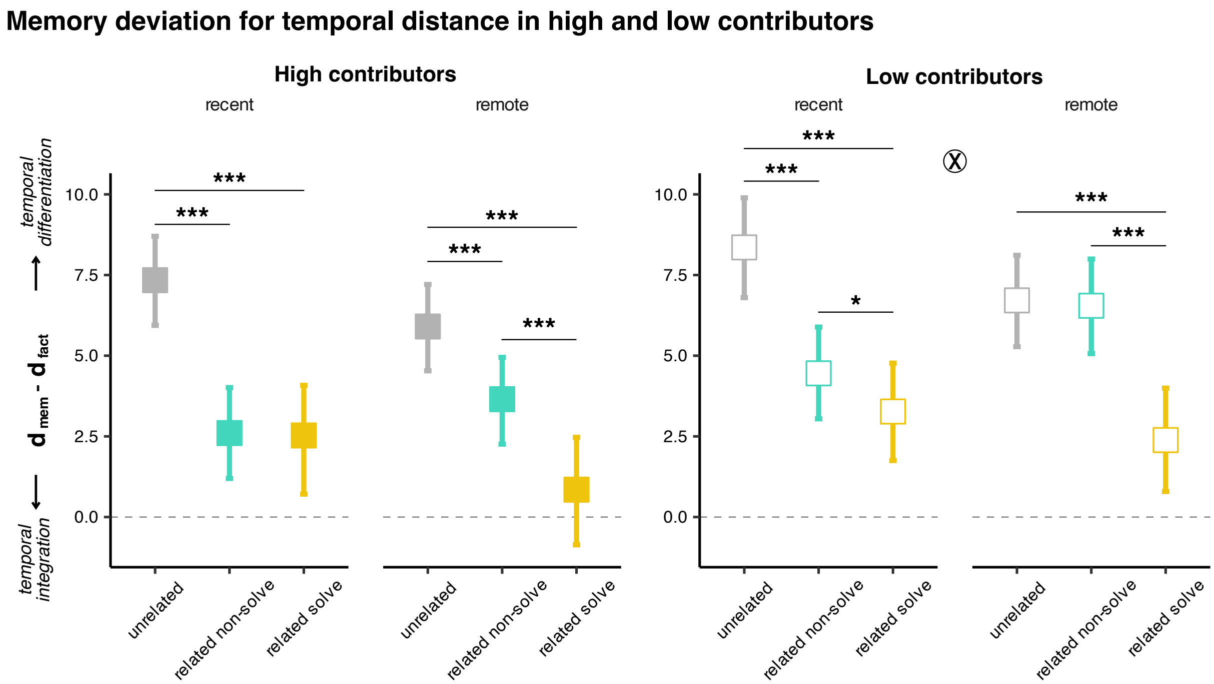


Figure 4. Memory deviation for temporal distance in high vs. low contributors. We performed separate Bayesian mixed-effects models within the high and low contributors of each group predicting the memory deviation in temporal distance (d_{mem} - d_{fact}) with event pair type, test delay, and their interaction. We also controlled for the corresponding d_{fact} and recall offset of individual events in the model (not depicted and held constant at the mean level). Squares represent the model-predicted posterior medians; error bars represent 95% highest density credible intervals (HDI). *: % in ROPE=0, 95%HDI excluded 0, and p_{dir}>0.95; ***: % in ROPE=0, 95%HDI excluded 0, and p_{dir}>0.999; ⊗ interaction drove superior model fit indicated by $|\Delta elpd|>4$ and $|\Delta elpd|/SE>2$.

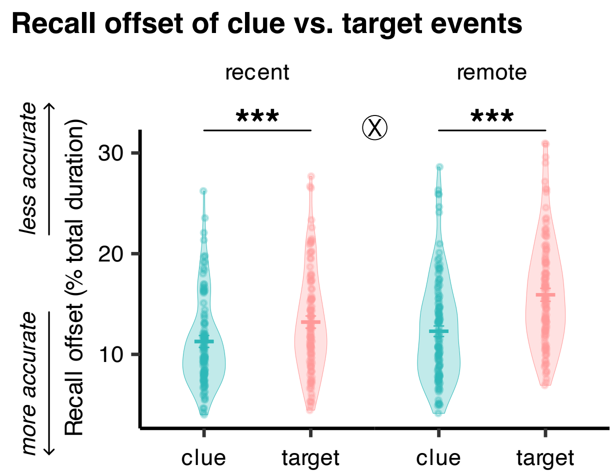


Figure 5. Recall offset of clue vs. target events. Recall offset for clue and target events within the causally-related pairs on the recent and remote tests. Each dot represents one participant; thick short lines represent the mean values; error bars represent 95% confidence intervals. ***p<0.001, ⊗ significant interaction.