



Published in final edited form as:

*J Exp Psychol Gen.* 2022 October ; 151(10): 2448–2465. doi:10.1037/xge0001215.

## Mnemonic vs. Temporality: Distinguishing between components of episodic representations

Johannes B. Mahr,

Daniel L. Schacter

Department of Psychology, Harvard University, 33 Kirkland St., Cambridge, MA 02138, USA

### Abstract

Human beings regularly ‘mentally travel’ to past and future times in memory and imagination. In theory, whether an event is remembered or imagined (its ‘mnemonic’) underspecifies whether it is oriented towards the past or the future (its ‘temporality’). However, it remains unclear to what extent the temporal orientation of such episodic simulations is cognitively represented separately from their status as memory and imagination. To address this question, we investigated to what extent episodic simulations are distinguishable in recall by virtue of both temporal orientation and mnemonic. In three experiments (N = 360), participants were asked to generate and later recall events differing along the lines of temporal orientation (past/future) and mnemonic (remembered/imagined). Across all of our experiments, we found that mnemonic and temporality each contributed to participants’ ability to discriminate different types of event simulations in recall. However, participants were also consistently more likely to confuse in recall event simulations that shared the same temporal orientation rather than the same mnemonic. These results show that the temporal orientation of episodic simulations can be cognitively represented separately from their mnemonic and have implications for debates about the structure of episodic representations as well as the role of temporality in this structure.

### Keywords

episodic simulation; episodic memory; mental time travel; temporality; mnemonic; reality monitoring

---

The human mind has the ability to produce a remarkable variety of event representations. We can generate representations of past and future events (e.g., Schacter et al., 2008), of merely possible and fictional events (e.g., Hassabis et al., 2007), and even ‘vicarious’ events that were experienced by other people (e.g., Pillemer et al., 2015). All of these representations are ‘episodic’ in the sense that they represent *events* and seem to rely on a perhaps unitary capacity for ‘episodic simulation’ (Addis, 2018; 2020; Schacter & Addis,

---

Correspondence concerning this article should be addressed to Johannes B. Mahr [jmahr@fas.harvard.edu](mailto:jmahr@fas.harvard.edu).

None of the experiments presented here were formally pre-registered. Data for these experiments is available upon request. All materials and procedures for Experiment 1, Experiment 2, and Experiment 3 are accessible via Testable.

A preprint of a previous version of this manuscript was posted on PsyArxiv (<https://psyarxiv.com/xawp8/>).

We have no conflict of interest to declare.

2007; Schacter, et al., 2008): a capacity to mentally generate imaginative constructions of events.<sup>1</sup>

Two particularly prominent aspects of such episodic representations are their temporal orientation or temporality (i.e. their orientation towards the future or the past; Tulving, 2002) and their mnemonicity (i.e. their status as remembered or imagined; Michaelian & Sutton, 2017). However, while some research (under the heading of ‘reality monitoring’; e.g., Johnson & Raye, 1981; Simons et al., 2017) has investigated how episodic representations achieve their mnemonicity<sup>2</sup>, less is known about the role of temporal orientation within the architecture of the episodic simulation system and its relationship to event mnemonicity. Here, therefore, we begin addressing the question of whether temporal orientation and mnemonicity rely on separable representational structures. Both mnemonicity (Michaelian, 2016; Johnson & Raye, 1981) and temporal orientation (Boyle, 2020; De Brigard & Gessell, 2016; Mahr, 2020; Mahr et al., 2021; Klein & Steindam, 2015) are generally viewed to rely on processes separate from those generating an episode’s contents (i.e. ‘what is happening’ in a given event). However, temporality and mnemonicity seem to be closely related; after all, one can only imagine the future. As a result, these elements have been difficult to separate experimentally and experimental work on episodic simulations has often conflated temporal orientation and mnemonicity (e.g. McDonough & Gallo, 2010; 2013; Addis et al., 2009). This conflation has made it difficult to determine to what extent both mnemonicity and temporal orientation are indeed represented as distinct elements in episodic simulation. Therefore, Schacter et al. (2012) emphasized the importance of understanding the role of temporal orientation in episodic simulation, and argued that distinguishing it from an event’s status as remembered or imagined “requires careful experimental designs that precisely target specific processes of interest” (p.680).

There are several reasons to expect an episode’s mnemonicity and its temporal orientation to rely on separable representations. First, while closely related, mnemonicity and temporal orientation are not reducible to each other. An episode’s status as remembered or imagined underspecifies its temporal orientation and vice versa: even though all remembered episodes are necessarily past directed, not all past directed episodes are necessarily remembered (one can imagine past events; e.g. Addis et al., 2009; De Brigard & Parikh, 2019). Conversely, even though all future episodes are necessarily imagined, not all imagined episodes are necessarily future-directed.

Second, temporal orientation, independently of mnemonicity, contributes to the functions of episodic simulations. Future-directed simulations likely have importantly different functions

---

<sup>1</sup>The term ‘episodic simulation’ has been used and defined in various ways, most often as applying to the imaginative construction of hypothetical mental events, and in the context of discussions of prospection, hypothetical future events (for discussion, see Schacter et al. 2008, and Szpunar, et al., 2014). However, episodic simulation can also be used in a broader sense to include episodic memories, in line with Bartlett’s (1932, p.213) early characterization of memory as ‘an imaginative construction or reconstruction’. In the present paper, we use episodic simulation in this broader sense that acknowledges that memories themselves are imaginative constructions or simulations (for recent discussions on this point, see Addis, 2018, 2020; De Brigard, 2014; Michaelian, 2016).

<sup>2</sup>As Michaelian, (2016) has pointed out, the status of an episodic representation as memory or imagination ought to be distinguished from questions of reality monitoring: whereas reality monitoring concerns the question of whether a given memory represents past reality or not, mnemonicity is about whether one is *currently* remembering or imagining. Nonetheless, research on reality monitoring, partly because it has tended to conflate these two issues, is relevant to the question when and how people represent the difference between current memories and imaginations.

from past-directed ones (see e.g. Bulley et al., 2020; Epstude & Roese, 2017; Schacter et al., 2017). For example, it has been suggested that confusions in the temporal orientation of episodic representations are a major driver of spontaneous confabulations in neurological patients (e.g. Dalla Barba et al., 1997; Schnider et al., 1996; Nedjam et al., 2000). Relatedly, Irish et al. (2012) have observed that semantic dementia patients seem to regularly confuse the temporal orientation of their episodic representations.

Finally, the phenomenology of episodic simulation is usually taken to involve ‘mental *time* travel’ (Tulving, 1983) involving a sense of subjective time (Tulving, 2002; Klein, 2015; 2018). As a result, the question of the role of subjective time (i.e., temporal orientation) in episodic representations has figured heavily in philosophical debates about whether episodic memory and future imagination should be taken to be one or separate psychological capacities (e.g. Perrin, 2016; Michaelian et al., 2021). This sense of time, moreover, is commonly analyzed as merely one element of the phenomenology of episodic remembering and imagining (e.g. Clayton & Russell, 2009; Boyle, 2019; 2020; Perrin et al., 2020). For example, while the feeling that a given episode belongs to one’s past is part of the phenomenology of remembering, this feeling can occur without episodic contents, as illustrated by experiences of *déjà vu*, where feelings of pastness (or familiarity; Brown, 2003) and feelings of premonition (Cleary & Claxton, 2018) can occur without episodic contents being available.

### **Is there a cognitive role for temporality within the architecture of episodic thought?**

While these considerations provide reasons to suspect that an episode’s temporal orientation and its mnemonicity can be represented separately, there is little research speaking to whether temporality as such plays an identifiable role within the cognitive architecture of episodic simulations. In fact, on standard accounts of the role of temporal information in episodic simulation, temporality itself need not be explicitly represented in episodic memory or episodic future thought at all (Conway et al., 2019; D’Argembeau, 2020). Instead, episodic representations (at least autobiographical ones) are thought to be organized within a hierarchy of autobiographical knowledge structures according to life periods, general events, and specific episodes (Conway & Rubin, 1993). On this view, temporal orientation seems to be a by-product of placing a given simulation within the temporal outline of one’s own life.

One study of particular note by McDonough and Gallo (2010; see also 2013), however, found an interaction effect between participants’ reality monitoring performance and event temporality. In a blocked design, they asked participants to generate autobiographical memories and future imaginations related to object word cues. In two separate tasks, they then asked participants to either judge whether they had generated a past event or to judge whether they had generated a future event for a given cue. Their results suggested that participants committed a higher number of misattributions in the ‘past-task’ compared to the ‘future-task’. McDonough and Gallo interpreted these results as revealing a ‘reality monitoring asymmetry’ in the judgment of past events – that is, an increased tendency to recall imagined events as remembered. However, because the temporality and mnemonicity

factors were confounded in this study, it remains unclear whether this effect was driven by misattributions of mnemonicity or temporality.

In another recent study with a similar design, De Brigard and colleagues (2020) found that the general temporal direction of simulated events (past/future) was recalled well, even though specific temporal information (time/month) of these simulations was not. These authors also found that recall for temporal information differed across event types: recall of temporal information was worse for future and counterfactual (i.e. imagined) than for remembered events. Nevertheless, De Brigard et al. concluded that “a temporal component may not be necessary when generating mental simulations of possible events” (p. 8).

Similarly, research on how episodic memories are dated has found that the specific date of a memory is not part of its contents but instead depends on ‘reconstructive’ mechanisms at retrieval (Friedman, 1993, 2004; 2005; see also Mahr et al., 2021). These mechanisms seem to resemble the kinds of mechanisms described under the source monitoring framework (Johnson et al., 1993) insofar as they also operate at the post-retrieval stage over episodic content. However, instead of relying solely on features of the content of a given simulation (such as perceptual detail), the inference mechanisms responsible for dating memories are usually taken to utilize general knowledge about one’s life history (i.e. autobiographical knowledge) or important world events (‘temporal landmarks’; Shum, 1998).

Finally, evidence from cognitive neuroscience has so far failed to identify a unique signature of the temporality of episodic representations (Schacter et al., 2012; but see Nyberg et al., 2010 for suggestive evidence). While some differences in neural activation have been observed between past- and future-directed episodic simulations (Addis et al., 2007; 2009; Benoit & Schacter, 2015), these differences are commonly ascribed to differing constructive demands between past (i.e., remembered) and future (i.e., imagined) simulations. Future-directed imaginations will commonly be more novel than remembered events and therefore require more extensive recombination of episodic details. These results are thus surprising in view of the potential importance of temporality in episodic thought mentioned above.

## The present study

In order to examine whether the temporal orientation of an episode is separable from its mnemonicity, we utilized a source memory paradigm with two parts, similar to previous research by McDonough and Gallo (2010; 2013) and De Brigard et al. (2020; see also McLelland et al., 2015). In the first part of our procedure (see Figure 1, “Simulation Encoding Task”), we asked participants to generate three kinds of episodic simulations in response to object word cues: (1) imagined future events, (2) imagined past events (i.e. episodic counterfactuals), and (3) remembered events. Then, in a second part (see Fig. 1, “Simulation Retrieval Task”), participants were presented with each object word cue again and asked to recall whether they had generated an event of type (1), (2), or (3) in response to this cue.

In spite of the fact that event mnemonicity and event temporality are not entirely independent dimensions of episodic simulation, crossing these factors allowed us to generate a  $2 \times$

2 design matrix (see Figure 2). However, because “remembered future events” fit less clearly into this conceptualization, Experiments 1 and 2 focused only on remembered past, imagined past, and imagined future events. Experiment 3 aimed to complete the  $2 \times 2$  design by including recast events (i.e. remembered events that were ‘recast’ to be occurring in the same way in the future; see Addis et al., 2009; Thakral et al., 2021) to control for the effects of content-similarities between imagined past and remembered events.

## Rationale

Importantly, whereas imagined future and imagined past events share the same mnemonicity (i.e., ‘imagined’) they differ in temporality. Conversely, imagined past and remembered events share the same temporal orientation (i.e., ‘past’) but differ in mnemonicity. Based on this logic, we reasoned that the way in which participants would confuse different types of simulation in recall (illustrated in Figure 3) would allow us to evaluate to what extent participants had represented an episode’s temporality and mnemonicity. On the one hand, if participants can discriminate simulations mainly according to their status as remembered vs. imagined, we would expect them to primarily confuse events along the lines of temporality (“Temporality Errors” in Figure 3); to be more likely to recall *imagines past* events as *imagined future* than as *remembered* events. This outcome would suggest that participants primarily tended to rely on event mnemonicity and not temporality in recall. On the other hand, if participants are able to discriminate simulations according to their temporal orientation as well, we would expect them to confuse events along the lines of mnemonicity (“Mnemonicity Errors” in Fig. 3); to falsely recall *imagined past* as *remembered* rather than as *imagined future* events.<sup>3</sup>

Two further predictions resulting from our logic bear mentioning. First, in Experiment 1, whereas imagined past events shared both temporality and mnemonicity with other simulation types, remembered events shared only temporality and imagined future events shared only mnemonicity with one other type of simulation. If participants indeed commit memory errors by confusing these elements of their episodic simulations, we would therefore expect them to commit such confusions most often in the imagined past condition. As a result, proportions of correct recall should be lowest in this condition. Second, by a similar logic, memory confusions between remembered events and imagined future events (see Fig. 3, “Temporality + Mnemonicity Errors”) should be rare because these representations share neither the same temporal orientation nor mnemonicity and should therefore be most easily distinguished in recall.

Finally, because we asked participants to rate each event at encoding on a number of rating scales (amount of perceptual detail, event familiarity, emotional intensity, and difficulty), our experimental design also allowed us to check whether (1) participants would indeed generate different types of events in different conditions, (2) event features (as measured

---

<sup>3</sup>Note that our design therefore differs in important ways from those of the studies by McDonough and Gallo (2010; 2013) and De Brigard et al. (2020). While McDonough and Gallo were also interested in asking whether episodic representations can be successfully distinguished along the lines of their temporal orientation, the temporality factor was confounded with mnemonicity in their design. In contrast, De Brigard et al. (2020) explicitly investigated how different elements of different episodic simulations are recalled. Their experiment, however, did not require participants to recall the kind of representation they had previously generated. This feature prevented these authors from comparing participants’ ability to recall temporality and mnemonicity.

by participants' ratings at encoding) would differentiate events better along the lines of temporality or mnemonicity, and (3) ratings at event generation would predict participants' memory performance.

In Experiment 2 we sought to replicate the results of Experiment 1 while ruling out some alternative explanations. Specifically, we wanted to ensure that our results were not caused by participants merely recalling event cues (rather than the generated events themselves) or by failing to generate events within the specified time frame. Finally, in Experiment 3 we controlled for the fact that the previous experiments were imbalanced in their design (i.e., the design required asking participants to generate more imagined than remembered and more past than future events) and for potential effects of similarity between imagined past and remembered events by including recast events as an additional event type.

## Experiment 1

### Methods

**Participants**—We recruited 120 native-English speaking participants from the UK and US between the ages of 18 and 55 via the online testing platform Prolific Academic (Palan & Schitter, 2018). We arrived at this sample size by doubling McDonough and Gallo's (2010) original sample of 60 to account for potentially greater noise caused by online data collection. Three participants had to be excluded because they failed more than one of our attention checks. Thus, 117 participants were included in the analysis for Experiment 1 ( $M_{\text{Age}} = 30.72$  years,  $SD_{\text{Age}} = 8.93$  years; 82 females). All participants provided explicit consent before taking part in the experiment. The procedure and methods for this experiment were approved by Harvard's Institutional Review Board (IRB19–198).

**Procedure**—The procedure for Experiment 1 is depicted in Figure 1. The experiment consisted of two parts. In Part 1, participants completed a *Simulation Encoding Task*, in which they were asked to generate memories and imaginations related to different object word cues and rate them on a number of different scales. In each of 40 'event generation' trials, participants were presented with a different word cue. These cues consisted of common object nouns high in imageability ( $M = 6.6$ ) and concreteness ( $M = 6.5$ ) (scores were computed according to Scott et al., 2019). Participants completed two practice trials before being presented with 30 test trials and eight filler trials at the end of Part 1. Only the 30 test trials were subsequently included in Part 2 (see below). Filler trials were identical for all participants and served as a buffer between Part 1 and the upcoming memory test in Part 2. In all other trials of Part 1, which object word cue was associated with which event type condition in Part 1 was counterbalanced equally across participants according to three counterbalancing orders.

In each event generation trial, participants were instructed to provide a short description (one or two sentences) of each generated event. This task was self-paced, even though participants had to spend a minimum of 15 seconds during each trial to provide an event description. After participants had provided an event description in a given trial, they were then asked to rate their simulation on four 100-point scales according to (1) how difficult it was to generate (0 = extremely easy; 100 = extremely difficult), (2) how many perceptual

details it included (0 = extremely vague; 100 = extremely detailed), (3) how emotionally intense it was (0 = not at all emotional; 100 = extremely emotional), and (4) how familiar it was (0 = I have never experienced anything similar; 100 = I have experienced this exact event).

To make sure that participants used the rating scales intentionally, we added four attention checks at randomly selected points to the event rating phase. In order to pass the attention check, participants were asked to position the slider at exactly the mid-point of the scale (i.e. '50'). If participants failed more than one of these attention checks, they were excluded from analysis.

Participants completed 10 test trials in each event type condition ('Imagine Past', 'Imagine Future', 'Remember'). We counterbalanced event type – cue word pairings across participants (in three counterbalancing orders to which participants were randomly assigned) such that each cue word was paired with each event type and the order of event type trials was randomized. In the 'remember' condition, participants were instructed to generate events that actually occurred in their personal past. By contrast, in the 'imagine past' condition, they were asked to generate events *that could have occurred* (but did not actually occur) in their personal lives in the past. In the 'imagine future' condition, participants were asked to generate events that *could plausibly occur* in their lives in the future. Participants were instructed that each of these events should be related to the event cue presented in a given trial. Events were supposed to occur within five years from the present (i.e. in the last five years in the past case and the next five years in the future case), happen over the span of a few minutes to a few hours, be unrelated across trials, and be constructed from a field (i.e. first-person) perspective ("through your own eyes").

In Part 2, participants were then presented with a *Simulation Retrieval Task*: a surprise source memory test for each of the 30 word cues from test trials in Part 1. For each word cue, participants were sequentially asked to decide whether in Part 1, they had (1) imagined a future event, (2) imagined a past event, or (3) remembered a past event. While the sequence of cues was determined randomly, response options were always presented concurrently. The procedures for all experiments in this study was built via the Testable platform (Rezlescu et al., 2020). The complete procedure for Experiment 1 can be accessed here.

## Results

All analyses for this and the subsequent experiments were carried out in *R* (version 3.6.2; R core team, 2019) with *RStudio* (RStudio Team, 2019). Plots were generated with *ggplot2* (Wickham, 2016). Mixed effects models were computed with the *lme4* package (Bates et al., 2015) and used the maximal random effects structure which still allowed the model to converge (Barr et al., 2014).

**Excluded trials**—Trials in which participants did not provide a description or did not describe an event (consisting at least of a noun + verb phrase) were excluded from analysis (N = 2). We also excluded from analysis event generation trials that were rated to be maximally difficult (difficulty = 100) because we assumed that participants failed

to generate an event simulation in these cases ( $N = 22$ ). Descriptive results for event ratings and recall performance (proportion of correct responses) in Experiment 1 (after these exclusions) are summarized in Table 1.

**Recall performance:** Average performance is depicted in Figure 4 and the distribution of responses in the source memory test across event type conditions is summarized in Figure 5. On average, participants recalled the correct event type in 63.0 % ( $SD = 19.5\%$ ) of trials. A one-way analysis of variance (ANOVA) comparing participants' proportion of correct responses across event type conditions ('Imagine Future', 'Imagine Past', 'Remember') produced a significant effect of condition ( $F(2, 348) = 24.62, p < .0001, \eta^2 = 0.12$ ). Tukey HSD post-hoc tests showed that participants performed significantly worse in the 'imagine past' condition compared to both the 'remember' ( $p < .0001$ ) and the 'imagine future' conditions ( $p < .0001$ ), while performing equally well in the latter two conditions. A one-sample t-test confirmed that participants performed significantly above chance in the 'imagine past' condition ( $t(114) = 6.86, p < .001$ ). This pattern of results is in line with our prediction that memory performance should be lowest in the Imagine Past condition because these events shared temporal orientation as well as mnemonicity each with one other type of simulation. In order to make sure that these performance differences were not caused by differences in how long participants spent with generating different event types, we performed another one-way ANOVA comparing the average amount of time spent describing each event across conditions. This analysis did not produce a significant result ( $F(2, 348) = 0.235, p = .791$ ), suggesting that participants spent a roughly equal amount of time with generating events of each type.

**Error types:** To test what kind of memory errors participants made, we compared participants' proportions of *Mnemonicity Errors*, *Temporality Errors*, and *Mnemonicity + Temporality Errors* via a one-way ANOVA (see Figure 6). This analysis produced a significant effect of error type ( $F(2, 342) = 73.58, p < .0001, \eta^2 = 0.3$ ). Tukey HSD post-hoc tests indicated that participants produced a higher proportion of Mnemonicity Errors ( $M = 0.54, SD = 0.27$ ) than Temporality Errors ( $M = 0.29, SD = 0.27, p < .0001$ ) and a higher proportion of Temporality than Mnemonicity + Temporality Errors ( $M = 0.17, SD = 0.15, p = .0005$ ). This pattern suggests that participants were more successful at distinguishing events according to their temporal orientation than according to their mnemonicity while also being sensitive to both elements in their discrimination performance.

Next, we tested whether Mnemonicity Errors were symmetrically distributed, that is, whether participants were as likely to confuse imagined past events with remembered events as vice versa. To do so, we compared the proportion of remembered events that participants recalled as imagined past events ( $M = 0.2, SD = 0.2$ ) to the proportion of imagined past events that were recalled as remembered ( $M = 0.38, SD = 0.2$ ) via a paired-sample t-test ( $t(116) = -5.1, p < .0001$ ). This analysis suggested that participants were more likely to confuse imagined past events with remembered events than vice versa. The asymmetrical distribution of Mnemonicity Errors replicates a finding by McDonough and Gallo (2010; 2013), who also found that participants were more likely to treat imagined events as remembered than vice versa.



In spite of this asymmetric distribution, if mnemonic errors were indeed driven by shared temporality, we would expect them not to be attributable to only one kind of memory error (i.e., imagined past events being recalled as remembered). Instead, we would expect remembered events to also be more likely to be confused with imagined past than with imagined future events. To test this prediction, we analyzed whether significantly more than half of participants' memory errors in the 'remember' condition were attributable to confusions with imagined past events rather than with imagined future events. That is, we only looked at trials in the 'remember' condition in which participants made a memory error and compared the proportion of 'imagined past' responses in those trials ( $M = 0.72$ ,  $SD = 0.32$ ) to 0.5 via a one-sample t-test ( $t(89) = 6.49$ ,  $p < .0001$ ; excluding 28 participants who did not make any errors in the 'remember' condition). The results of this test suggest that participants were substantially more likely to confuse remembered events with imagined past than with imagined future events.

The same analysis applied to Temporality Errors did not reveal an asymmetrical distribution of errors: participants were roughly equally likely to confuse imagined future with imagined past events as vice versa ( $p = .422$ ).

**Event ratings:** Mean ratings in each condition are summarized in Table 1. Separate one-way ANOVAs for each rating type revealed that participants' detail ( $F(2, 348) = 12.5$ ,  $p < .0001$ ,  $\eta^2 = 0.07$ ), difficulty ( $F(2, 348) = 4.02$ ,  $p = .019$ ,  $\eta^2 = 0.02$ ), and familiarity ratings ( $F(2, 348) = 46.28$ ,  $p < .0001$ ,  $\eta^2 = 0.21$ ), differed between event type conditions in Part 1, while emotionality ratings did not differentiate between conditions ( $F(2, 348) = 2.0$ ,  $p = .136$ ,  $\eta^2 = 0.01$ ). Tukey HSD post-hoc tests showed that remembered events were rated as more detailed (both  $ps < .0008$ ), more familiar (both  $ps < .0001$ ), and easier to generate (both  $ps < .046$ ) than both imagined past and imagined future events. The ratings for imagined past and imagined future events did not differ in any category (all  $ps > .362$ ). Thus, overall, event ratings tended to differentiate events along the lines of mnemonicity rather than temporal orientation. This finding indicates that participants did indeed generate different types of events when asked to generate imagined (past and future) and remembered events in the Simulation Encoding Task.

**Relationship between event ratings and memory errors:** In order to test whether event ratings at encoding predicted whether participants would commit a memory error in the source memory task, we fitted a logistic mixed-effects model (estimated using a ML and Nelder-Mead optimizer) to predict a binomial variable coding for Response Correctness (0 = incorrect, 1 = correct) with fixed effects for Familiarity, Detail, Emotion, and Difficulty. The model included Participant Number and Item as random effects with fixed slopes. Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using the Wald approximation. Neither the effect of Familiarity ( $p = 0.199$ ; Std. beta =  $-0.06$ , 95% CI  $[-0.15, 0.03]$ ), nor of Detail ( $p = 0.346$ ; Std. beta =  $-0.05$ , 95% CI  $[-0.15, 0.05]$ ), nor of Emotionality ( $p = 0.229$ ; Std. beta =  $0.06$ , 95% CI  $[-0.04, 0.15]$ ), nor of Difficulty ( $p = 0.612$ ; Std. beta =  $0.02$ , 95% CI  $[-0.07, 0.12]$ ) reached statistical significance.

However, while event ratings did not predict memory errors, they might still predict whether participants confused an event's mnemonic or temporality. Therefore, we generated two further logistic mixed effects models (in the same manner as above). First, we generated two binomial variables, each respectively coding for whether participants had confused an event's mnemonic and whether they had confused an event's temporality (0 = no confusion, 1 = confusion). That is, all trials in which participants had committed either a Mnemonic or a Mnemonic + Temporality Error were coded as mnemonic confusions. Conversely, all trials in which participants had committed either a Temporality or a Mnemonic + Temporality Error were coded as temporality confusions. On this basis, we generated two models, the first model predicting Mnemonic Confusions, the second predicting Temporality Confusions, each with fixed effects for Familiarity, Detail, Emotion, and Difficulty. Each model included a random effect for Participant Number with fixed slopes.

Regarding mnemonic recall, this analysis showed a significant and positive effect of Familiarity on Mnemonic Confusions ( $p < .001$ ; Std. beta = 0.23, 95% CI [0.13, 0.33]), while no other rating type showed a significant effect (all  $ps > .143$ ). In contrast, regarding temporality recall, we found significant and negative effects of Familiarity ( $p = .001$ ; Std. beta = -0.19, 95% CI [-0.30, -0.07]) and Emotionality ( $p = .015$ ; Std. beta = -0.15, 95% CI [-0.27, -0.03]) on Temporality confusions. Thus, high event familiarity increased the likelihood of mnemonic confusions, while decreasing the likelihood of temporality confusions.

## Experiment 2

Experiment 2 was designed to replicate Experiment 1 while ensuring that (1) participants indeed generated events within the appropriate time frame and (2) participants recalled in Part 2 the events they had generated as opposed to the event cues presented to them in Part 1. To do so, we asked participants in each event generation trial of Part 1 to provide an 'event title' and a 'date' (month/year) in addition to an event description. The provided date allowed us to ensure that participants indeed placed the event within the next (in the future case) or last (in the past case) five years of their lives. The event 'title' was to consist of a one-word or two-word summary of the event whereas the event description was to describe the event in more detail (one or two sentences). For each event cue in Part 2, we then asked participants to recall in as much detail as possible the event they had generated in Part 1 and to provide the title they had chosen for it. This procedure allowed us to assess whether participants made their memory responses based on recall of the initially generated event simulation and thus to assess whether the results of Experiment 1 were attributable to participants failing to recall the event they had originally generated. In Experiment 2 we therefore sought to extend these results by investigating whether participants would produce a similar pattern of memory confusions in Part 2 even for those trials on which they produced evidence of recalling the events they had generated for a given cue in Part 1.

## Methods

**Participants**—We recruited 120 native-English speaking participants between the ages of 18 and 55 from the UK and US via Prolific Academic. Four participants had to be excluded

from analysis because they failed more than one attention check. Thus, the final sample of Experiment 2 consisted of 116 participants ( $M_{\text{Age}} = 29.76$  years,  $SD_{\text{Age}} = 10.06$  years; 72 females). None of the participants taking part in Experiment 2 were part of the sample for Experiments 1 or 3. All participants provided explicit consent before taking part in the experiment. The procedure and methods for this experiment were approved by Harvard's Institutional Review Board (IRB19–198).

**Design**—Experiment 2 was identical in design to Experiment 1 apart from the fact that we asked participants in Part 1 to additionally provide an ‘event title’ and an ‘event date’. For the title, participants were instructed to provide a one- or two-word summary of the event (e.g. ‘grocery shopping’). For the date, participants were instructed to provide the month and year at which the event was set to occur (e.g. “January 2018”). Further, a ‘title recall’ task was added to Part 2: before the source memory test participants were now asked to recall the event in as much detail as possible and then provide its title. Our main interest in the responses for this title-recall task was to assess whether participants showed evidence of recalling the event they had initially generated. Therefore, we treated trials in which participants, rather than providing the original title, provided parts of the original event description as correct responses in this task. In coding the dates that participants provided, we treated all dates that fell outside the last (in the past case) or the next (in the future case) five years from the date at which data collection occurred (September 18<sup>th</sup>, 2020) as incorrect responses. All future dates in past conditions as well as all past dates in future conditions were also treated as incorrect responses. The complete procedure of Experiment 2 can be accessed here.

## Results

**Excluded trials**—As in Experiment 1, we excluded from analysis trials in which participants did not provide an appropriate event description (148 trials) and all events that were rated as maximally difficult to generate (32 trials). As a result of these exclusions one participant did not contribute valid trials to all event type conditions and was therefore excluded from further analysis.

Because our primary interest was in whether participants who succeeded in both generating and recalling events appropriately would show the same pattern of memory errors as in Experiment 1, we focused our analysis only on those trials in which participants had provided an appropriate date and recalled a title in Part 2 that matched the title or the event description they had given in Part 1. On average, participants provided an appropriate date in 85 % of trials ( $SD = 19$  %). As a result of date-based exclusions (413 trials), two further participants did not contribute valid trials to all conditions and were excluded from the below analyses.

Further, participants recalled an appropriate title (arguably indicating recall of the respective event) on average in 94% of trials ( $SD = 12.5$ %). A paired-sample t-test showed that, on average, participants recalled the event title in a significantly higher percentage of trials than event type ( $M = 66.7$ %,  $SD = 17.5$ %,  $t(112) = 16.87$ ,  $p < .0001$ ). Title-recall-based exclusions (187 trials) lead to the exclusions of two further participants. Descriptive results

for the remaining 111 participants (and 2669 trials) of Experiment 2 are summarized in Table 1.

**Recall performance**—Participants' proportions of correct responses in each event type condition are depicted in Figure 4 and distributions of responses across event type conditions are depicted in Figure 5. A one-way ANOVA for differences in the proportions of correct responses across event type conditions produced a significant result ( $F(2, 330) = 40.79, p < .0001, \eta^2 = 0.2$ ). Just as in the previous experiment, Tukey HSD post-hoc tests found that participants performed significantly worse at identifying imagined past events in recall compared to imagined future ( $p < .0001$ ) and remembered events ( $p < .0001$ ) while not differing in recall accuracy across the latter two conditions ( $p = .851$ ).

**Error types**—A one-way ANOVA revealed that participants' proportions of errors differed across error types ( $F(2, 324) = 144.22, p < .0001, \eta^2 = 0.47$ ; see Figure 6). Similar to Experiment 1, Tukey HSD post-hoc tests showed that participants committed a higher proportion of Mnemicity Errors ( $M = 0.64, SD = 0.28$ ) than Temporality Errors ( $M = 0.26, SD = 0.28, p < .0001$ ) and a higher proportion of Temporality than Mnemicity + Temporality Errors ( $M = 0.1, SD = 0.14, p = .0001$ ). This pattern suggests that the enhanced distinctiveness of episodic representations along the lines of temporal orientation compared to mnemicity was not caused by participants' failure to recall the event itself.

Further, a paired-sample t-test found that participants were again more likely to confuse imagined past events with remembered events ( $M = 0.44, SD = 0.36$ ) than vice versa ( $M = 0.21, SD = 0.24, t(110) = 5.61, p < .0001$ ). Nonetheless, significantly more than half of the errors in the 'remember' condition (of the 70 participants that contributed data to and committed errors in that condition) were attributable to 'imagined past' responses ( $M = 0.87, SD = 0.29$ ) as determined by a one-sample t-test ( $t(69) = 10.43, p < .001$ ). As in the previous experiments, participants committed a similar proportion of both types of Temporality Errors ( $p = .088$ ).

**Event ratings**—Replicating Experiment 1, separate one-way ANOVAs for each rating type comparing ratings across encoding conditions showed that participants' detail ( $F(2, 330) = 20.19, p < .0001, \eta^2 = 0.11$ ), and familiarity ratings ( $F(2, 330) = 43.28, p < .0001, \eta^2 = 0.21$ ) differed across conditions. Tukey HSD post-hoc tests revealed that remembered events were again rated as more detailed and familiar than both imagined past ( $p_{\text{Detail}} = .0003, p_{\text{Familiarity}} < .0001$ ) and imagined future events ( $p_{\text{Detail}} = .0001, p_{\text{Familiarity}} < .0001$ ) suggesting that participants did indeed generate different types of events when asked to imagine and remember events respectively. However, in contrast to the previous experiment, imagined past and future events slightly differed in ratings of perceptual detail ( $p = .046$ ).

**Relationship between event ratings and memory errors:** Just as in Experiment 1, we fitted a logistic mixed-effects model (estimated using a ML and Nelder-Mead optimizer) to predict Response Correctness (0 = incorrect, 1 = correct) in the source memory test with fixed effects for Familiarity, Detail, Emotion, and Difficulty. The model included Participant Number as a random effect with fixed slopes. Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs)

and  $p$ -values were computed using the Wald approximation. In contrast to Experiment 1, we found a significant and negative effect of Familiarity on Response Correctness ( $p = .002$ ; Std. beta =  $-0.16$ , 95% CI [ $-0.26, -0.06$ ]). That is, high event familiarity increased the likelihood of memory errors in Experiment 2.

To test whether event ratings predicted mnemonic or temporality confusions, we again generated binomial variables respectively coding for whether participants had confused event mnemonic or temporality in a given trial. Then, we again generated two models, the first model predicting Mnemonic Confusions, the second predicting Temporality Confusions, each with fixed effects for Familiarity, Detail, Emotion, and Difficulty and a random effect for Participant Number with fixed slopes.

Regarding mnemonic confusions, this analysis showed a significant and positive effect of Familiarity on Mnemonic Confusions ( $p < .001$ ; Std. beta =  $0.35$ , 95% CI [ $0.23, 0.47$ ]) and a positive trend for Detail ( $p = .077$ ; Std. beta =  $0.12$ , 95% CI [ $-0.01, 0.24$ ]). Neither Emotionality nor Difficulty showed a significant effect (both  $p$ s  $> .505$ ). Regarding temporality confusions, we only found trends for a negative effects of Familiarity ( $p = .071$ ; Std. beta =  $-0.13$ , 95% CI [ $-0.28, 0.01$ ]) and a negative trend for Emotionality ( $p = .059$ ; Std. beta =  $-0.15$ , 95% CI [ $-0.31, -0.005$ ]) on Temporality Confusions and no effect of Detail ( $p = .657$ ) nor Difficulty ( $p = .634$ ). Therefore, event familiarity might not have helped participants to determine event temporality as much as it did in Experiment 1, which could also explain its relationship to response correctness here.

### Experiment 3

The previous experiments suggest that participants can discriminate different types of episodic simulations separately according to mnemonic and temporality. Further, participants were more successful at distinguishing simulations along the lines of temporality than mnemonic. Nonetheless, Experiments 1 and 2 suffer from a number of limitations. First, these experiments had an imbalanced design: they included more imagined than remembered and more past than future events. This feature could have biased participants' responses. Second, Experiments 1 and 2 could not rule out one important alternative explanation for the observed pattern of memory errors: it might simply be attributable to the fact that imagined past events were more similar (both in terms of their contents as well as in the processes generating them) to remembered events than to imagined future events. In other words, participants might tend to confuse imagined past and remembered events not because of factors intrinsic to their shared temporal orientation. Instead, these confusions might be caused by the fact that generating an imagined past event requires generating a remembered event first. After all, in order to generate counterfactuals, one has to first recall what actually occurred (see Mahr, 2020; Byrne, 2005; Espino & Byrne, 2020).

In fact, this circumstance might also explain the asymmetric distribution of Mnemonic Errors we observed: participants might be more likely to recall imagined past events as remembered than vice versa because (in order to generate the counterfactual) they did in fact generate a remembered event when they were asked to imagine a past event.

Mnemicity Errors might therefore have been driven by a given event's proximity to actual past experiences. Experiment 3 was set up to address these potential concerns.

We wanted to test whether mnemicity confusions would occur more often than temporality confusions even in cases in which the generation of an imagined future event would (similar to imagined past events) also require the prior retrieval of actual past experiences. To do so, we introduced a fourth event type into our design: recast events, that is remembered events that are then 'recast' into the future (Addis et al., 2009; Thakral et al., 2021). In order to generate such events, participants were instructed to first remember an event that they experienced in the past related to the cue word and to then think about it happening again in exactly the same way (involving the same people, places, and objects) in the future. As such, even though these simulations are not in fact 'remembered' future events, they should be closely similar to other remembered events in terms of their contents and the processes producing them in so far as they were experienced before.

As a result, our predictions centered on how participants would recall recast events in the source memory test. If our previous results were indeed merely an outcome of the shared proximity of imagined past and remembered events to actual experiences, we would expect participants to also be more likely to confuse recast with remembered events than with imagined future events. If, however, Mnemicity Errors were not driven by such proximity but by shared temporal orientation instead, we would expect participants to be more likely to confuse recast with imagined future than with remembered events. Figure 7 depicts how the addition of recast ("Remembered Future") events changes how Mnemicity, Temporality and Mnemicity + Temporality Errors were computed.

## Methods

**Participants**—We recruited 120 native English-speaking participants between the ages of 18 and 55 from the UK and the United States via Prolific Academic. Four participants had to be excluded from analysis for failing more than one of our attention checks. Thus, a final sample of 116 participants were included in the analysis for Experiment 3 ( $M_{\text{Age}} = 31.84$  years,  $SD_{\text{Age}} = 11.35$  years, 66 females). None of the participants taking part in Experiment 3 were part of the sample for Experiments 1 or 2. All participants provided explicit consent before taking part in the experiment. The procedure and methods for this experiment were approved by Harvard's Institutional Review Board (IRB19–198).

**Design**—Experiment 3 added a fourth event type condition (recast events) to the Simulation Encoding Task. To generate recast events, participants were instructed to first remember an event that actually happened to them related to the cue word and to then think about the exact same event happening again including the same locations, objects, and people. In order to account for the increased difficulty of the recall task due to the additional event type and in order to present participants with an equal number of events of each type, we reduced the number of test events from 30 to 24. Therefore, participants were presented with 6 events (instead of 10 as in the previous experiments) in each event type condition. Apart from these changes, the procedure and design of Experiment 3 was identical to that of Experiments 2. That is, Experiment 3 also asked participants to provide (1) a title, (2) a

date (within 5 years of the testing date), and (3) an event description for each event. In the Simulation Retrieval Task, participants were then again asked to first recall the event title for the event in question and then recall the event type. The complete procedure for Experiment 3 can be accessed here.

## Results

**Excluded trials**—Just as in Experiment 2, we excluded all trials in which participants did not provide an appropriate event description ( $N = 76$ ), provided an inappropriate date ( $N = 167$ ), and trials that were rated as maximally difficult ( $N = 28$ ). As a result of these exclusions, eight participants did not provide valid trials in all four event type conditions and were therefore excluded from further analysis.

On average, participants recalled an appropriate title in Part 2 (indicating that they were able to recall the events they had generated in Part 1) in 95.9% ( $SD = 9\%$ ) of trials. A paired-sample  $t$ -test comparing the proportions of correct responses in the title recall task and the event type recall task across all event type conditions ( $M = 59.5\%$ ,  $SD = 18.5\%$ ) indicated that participants performed significantly better in recalling event titles than event types ( $t(107) = 20.73$ ,  $p < .0001$ ). To further analyze participants' memories for event type, we excluded all trials in which they failed to recall the correct event title ( $N = 98$ ). As a result of these title-memory-based exclusions, one further participant did not provide valid trials in all four event type conditions and was therefore excluded from further analysis. Descriptive results of the 107 remaining participants of Experiment 3 are summarized in Table 2.

**Recall accuracy**—A one-way ANOVA revealed that participants' proportions of correct responses differed across conditions ( $F(3, 424) = 23.64$ ,  $p < .0001$ ,  $\eta^2 = 0.14$ ; see Figure 8). Tukey HSD post-tests showed that imagined past and recast events were recalled less accurately than both remembered and imagined future events (all  $ps < .0001$ ). However, recall performance did not differ between imagined future and remembered events ( $p = .891$ ) nor between imagined past and recast events ( $p = .398$ ). Note that participants were on average well above chance (0.25) in correctly classifying recast events in recall ( $M = 0.52$ ,  $SD = 0.33$ ), suggesting that they generally succeeded at generating those events.

**Error types**—A one-way ANOVA indicated that average proportions of errors differed across error types ( $F(2, 315) = 85.13$ ,  $p < .0001$ ,  $\eta^2 = 0.35$ ; see Figure 9). Tukey HSD post-tests revealed that participants committed a higher proportion of Mnemicity Errors ( $M = 0.54$ ,  $SD = 0.26$ ) than Temporality Errors ( $M = 0.33$ ,  $SD = 0.26$ ,  $p < .0001$ ) and a higher proportion of Temporality Errors than Mnemicity + Temporality Errors ( $M = 0.12$ ,  $SD = 0.16$ ,  $p < .0001$ ; see Figure 9).

We classified confusions between recast and remembered events as Temporality Errors and confusions between recast and imagined future events as Mnemicity Errors. Thus, this result suggests that participants again primarily confused past with past and future with future events irrespective of their mnemicity. Confirming that this pattern held also in the case of recast events, a paired-sample  $t$ -test ( $t(106) = 6.53$ ,  $p < .0001$ ) showed that participants were significantly more likely to recall recast events as imagined future ( $M = 0.32$ ,  $SD =$

0.29) than as remembered events ( $M = 0.10$ ,  $SD = 0.16$ ). Conversely, another paired-sample t-test ( $t(106) = 3.11$ ,  $p = .002$ ) showed that remembered events were significantly more often recalled as imagined past events ( $M = 0.16$ ,  $SD = 0.23$ ) than as recast events ( $M = 0.07$ ,  $SD = 0.13$ ).

Further, just as in the previous experiments, we found a ‘reality monitoring bias’ in participants’ recall for past events: they were more likely to recall imagined past events as remembered ( $M = 0.27$ ,  $SD = 0.31$ ) than vice versa ( $M = 0.15$ ,  $SD = 0.23$ ,  $t(106) = 3.29$ ,  $p = .001$ ). However, the results of Experiment 3 suggest a corresponding bias in recall for future events: participants were more likely to recall recast events as imagined future events ( $M = 0.32$ ,  $SD = 0.29$ ) than vice versa ( $M = 0.2$ ,  $SD = .21$ ,  $t(106) = 3.82$ ,  $p = .0002$ ). This result speaks against the possibility of a general reality monitoring bias towards recalling imagined events as remembered. Instead, these results suggest that asymmetries in reality monitoring (see McDonough & Gallo, 2010; 2013) interact with the temporal orientation of the simulation in question. Whereas past events tended to be recalled as remembered (i.e. previously experienced), future events tended to be recalled as imagined (i.e. not previously experienced/novel).

**Event ratings**—Mean ratings for each rating type are summarized in Table 2. Separate one-way ANOVAS for each rating type showed that Detail ( $F(3, 424) = 21.16$ ,  $p < .0001$ ,  $\eta^2 = 0.13$ ), Difficulty ( $F(3, 424) = 8.82$ ,  $p < .0001$ ,  $\eta^2 = 0.06$ ), Familiarity ( $F(3, 424) = 71.49$ ,  $p < .0001$ ,  $\eta^2 = 0.34$ ), and Emotionality ratings ( $F(3, 424) = 5.41$ ,  $p < .0001$ ,  $\eta^2 = 0.04$ ) differed across event type conditions. Tukey HSD post-tests showed that, compared to remembered events, recast events were rated as less detailed ( $p = .0002$ ), slightly harder to generate ( $p = .021$ ), and less familiar ( $p = .001$ ). However, recast events were also rated as more detailed than imagined past events ( $p = .019$ ), and more familiar than both imagined future and imagined past events ( $p < .0001$ ). This pattern suggests that recast events were distinguishable from both remembered, imagined past, and imagined future events in terms of their event characteristics.

In order to test whether reported event characteristics would differentiate events better according to their temporal orientation or their mnemonicity, we calculated absolute rating difference scores between past (i.e. remembered and imagined past) events and future (i.e. recast and imagined future) events on the one hand and compared them to difference scores between remembered/previously experienced (i.e. remembered and recast) events and imagined/novel (i.e. imagined future and imagined past) events on the other hand (see Figure 10). Bonferroni-corrected paired-sample t-tests suggested that – compared to differences in temporality – differences in mnemonicity were associated with larger mean absolute difference scores in event Detail ( $t(106) = 5.23$ ,  $p < .0001$ ) and event Familiarity ( $t(106) = 6.59$ ,  $p < .0001$ ) but not with any differences in Emotionality and Difficulty (both  $ps > .141$ ). Similar to the previous experiments, this result suggests that based on event characteristics alone, events should have been more easily distinguishable along the lines of their mnemonicity than their temporal orientation.

**Relationship between event ratings and memory errors**—Just as in the previous experiments, we fitted a logistic mixed effects model (estimated using a ML and Nelder-



Mead optimizer) to predict Response Correctness (0 = incorrect, 1 = correct) in the source memory test with fixed effects for Familiarity, Detail, Emotion, and Difficulty. The model included Participant Number as a random effect with fixed slopes (including Item as an additional random effect caused a singular fit). Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using the Wald approximation. Similar to Experiment 1, we found no significant effect of Familiarity on Response Correctness ( $p = .100$ ; Std. beta =  $-0.09$ , 95% CI  $[-0.20, 0.02]$ ). In contrast to Experiment 1, however, we found a positive effect of Detail on Response Correctness ( $p = .005$ ; Std. beta =  $0.19$ , 95% CI  $[0.06, 0.32]$ ). That is, the more detailed an event was rated to be at encoding, the more likely participants were to later correctly recall its event type. Emotionality ( $p = .435$ ) and Difficulty ( $p = .819$ ) did not predict Response Correctness.

Further, to test whether event ratings predicted mnemonic or temporality confusions, we again generated binomial variables respectively coding for whether participants had confused event mnemonic or temporality in a given trial and generated two mixed effects models: the first model predicting Mnemonic Confusions, the second predicting Temporality Confusions, each with fixed effects for Familiarity, Detail, Emotion, and Difficulty and a random effect for Participant Number with fixed slopes.

Regarding mnemonic confusions, this analysis showed a significant and positive effect of Familiarity ( $p < .001$ ; Std. beta =  $0.20$ , 95% CI  $[0.08, 0.33]$ ) and a significant and negative effect for Detail on Mnemonic Confusions ( $p = 0.019$ ; Std. beta =  $-0.17$ , 95% CI  $[-0.31, -0.03]$ ). Neither Emotionality nor Difficulty showed a significant effect (both  $ps > .566$ ). Further, we only found a significant negative effect of Emotionality ( $p = .043$ ; Std. beta =  $-0.15$ , 95% CI  $[-0.29, -0.005]$ ) and no effect of Familiarity ( $p = .280$ ), Detail ( $p = .657$ ), or Difficulty ( $p = .634$ ) on Temporality Confusions.

## General Discussion

Is the temporal orientation of episodic representations represented separately from their status as remembered vs. imagined (their 'mnemonic')? Further, if people indeed commonly represent both event mnemonic and temporality, which of these elements allows them to more easily distinguish their mental event simulations from each other? In order to answer these questions, we investigated the kinds of memory errors participants would make when distinguishing episodic representations along the lines of temporal orientation and mnemonic in recall. Across three experiments, we found two main results.

First, participants were more likely to confuse their simulations along the lines of *either* mnemonic *or* temporality than along the lines of *both* mnemonic *and* temporality. This pattern suggests that event mnemonic and temporality *each* affected episodic simulations in such a way so as to allow participants to distinguish them in recall. If only one of these elements had driven participants' recall, we would have expected them to be equally likely to confuse simulations along the lines of one element as to commit confusions along the lines of both elements.

Second, we found that participants tended to confuse in memory the mnemonicity of events sharing the same temporal orientation. Importantly, mnemonicity confusions occurred irrespective of temporal orientation: past-oriented events were primarily recalled as past and future-oriented events were primarily recalled as future irrespective of their status as remembered or imagined. This finding suggests that episodic representations were easier to distinguish along the lines of temporal orientation than mnemonicity.

**The representation of mnemonicity and temporality**—These results can be explained in two main ways. First, it is possible that the observed pattern of source memory errors was caused by factors at the event generation stage. On this interpretation, mnemonicity errors were not caused by a ‘retrieval advantage’ for temporal orientation. Instead, differences in temporal orientation might instead have stronger effects than mnemonicity on how events are generated. For example, sharing the same temporality might have caused participants to generate events that were more similar to each other than events sharing the same mnemonicity. In line with this possibility, one can arguably draw more heavily on actual experiences in the generation of past events than in the generation of future events (Benoit & Schacter, 2015; Addis et al., 2009). The resulting differences, for example in event familiarity, might have made past and future events easier to distinguish at recall than imagined and remembered events. However, mnemonicity confusions do not seem to have been driven by differences in the extent to which event generation relied on the retrieval of actual experiences. If that was the case, we would have expected participants to confuse imagined past as well as recast events primarily with remembered events. Instead, in Experiment 3, recast events tended to be misremembered (in line with their temporal orientation) as imagined future events (and vice versa). Importantly, the possibility that, in generating recast events in Experiment 3, participants might have slightly changed these events from their remembered origin in order to transpose them into the future does not speak against this interpretation: even if recast events were slightly altered, if participants succeeded at generating these events, they must have explicitly retrieved an actual past experience to do so. Note also that both the fact that recast events differed from remembered as well as imagined future events in terms of their event characteristics and that participants were well above chance in correctly classifying recast events in recall suggests that they succeeded at generating these events.

More generally, participants seem to have succeeded at generating different types of events in different conditions, as evidenced by the fact that their ratings of event characteristics consistently differentiated between conditions. In particular, participants seem to have distinguished events according to their mnemonicity: in line with previous research (D’Argembeau & van der Linden, 2004; Johnson et al., 1988; McDonough & Gallo, 2010), participants’ ratings of event characteristics consistently distinguished events more clearly according to their mnemonicity than their temporal orientation. Moreover, Experiment 3 showed that event ratings differed more strongly along the lines of mnemonicity than temporality. Indeed, participants seem to have been sensitive to the fact that familiarity ratings distinguished events particularly well according to their mnemonicity. As a result, and as shown by the fact that familiarity ratings tended to predict mnemonicity errors, whenever this heuristic did not hold, participants were more likely to commit a mnemonicity error.

Another possibility is that source memory errors were caused by participants selectively failing to generate the appropriate event type at encoding. When asked to generate an imagined past event, participants might have sometimes failed to generate a counterfactual and instead just generated a remembered event. Similarly, in the case of recast events, instead of retrieving a remembered event, participants might have sometimes simply generated an imagined future event. Therefore, even though participants *generally* seem to have succeeded at generating events of each type, source memory errors might have been caused by *selective* failures to generate the correct events in specific trials. On this interpretation, however, it remains unclear why participants also primarily misremembered imagined future as recast and remembered as imagined past events. To explain this pattern, one would have to assume that, not only did participants tend to selectively generate remembered events in ‘imagine past trials’ and imagined future events in ‘recast trials’ but also that they selectively committed the reverse errors. If this were the case, it would suggest that participants selectively confused the mnemonicity of events in line with their temporal orientation *at encoding* rather than at recall. This confusion, in turn, would suggest that temporal orientation has a larger influence on the generation of episodic representations than mnemonicity does. However, while failures to generate imagined past and recast events in specific trials might be explained by the increased complexity of these event types, it is less clear why participants would have erroneously generated imagined past events in ‘remember trials’ or recast events in ‘imagine future trials’.

In light of these considerations, an explanation of the pattern of source memory errors in the present study based on failure at event generation alone seems unlikely. Nonetheless, more research is required to definitively rule out this possibility. In particular, replicating the present results with an event recombination paradigm similar to the one used by Addis and colleagues (2009) could directly control for similarities between event types and failures to generate different event types.

A second way to explain the present results focuses on retrieval processes: events might have been distinguishable along the lines of both mnemonicity and temporality because of processes occurring *at recall* rather than at the event generation stage. Further, events might have been more easily distinguishable along the lines of temporality due to a ‘retrieval advantage’ of this element: participants might have been more successful at recovering event temporality than mnemonicity because they actually were more likely to retrieve a given event’s temporality than its mnemonicity. This retrieval-based explanation implies that participants might have indeed retrieved and therefore separately represented the temporal orientation and mnemonicity of their simulations as such. The present study might therefore provide the first specific behavioral evidence (to our knowledge) for the explicit representation of temporality in the generation and retrieval of episodic representations. This outcome would support Mahr’s (2020) assertion that episodic simulations might include dedicated representations of both temporal orientation and mnemonicity, and arguments by Schacter et al. (2012) regarding the theoretical importance of distinguishing mnemonicity from temporality.

While our results do not allow us to draw strong conclusions about what caused the retrieval advantage of temporality over mnemonicity, it is possible that temporality

and mnemonicity relied on different retrieval processes. More concretely, in contrast to mnemonicity, temporality might have been explicitly encoded during event generation and could therefore later be retrieved ‘propositionally’. Indeed, the fact that event ratings at encoding reflected mnemonicity (but not temporality) differences together with the fact that such rating differences predicted mnemonicity errors might be taken to suggest that mnemonicity was not explicitly encoded but had to be inferred from event characteristics. In other words, participants might have relied on differences in event characteristics between remembered and imagined events to *infer* mnemonicity at recall while being able to *retrieve* a propositional representation of temporality. Future research should therefore more explicitly consider differences in how different elements of episodic simulations might be cognitively entertained.

Further, since participants in Experiments 2 and 3 were instructed to generate event dates at encoding it is possible that participants might have then retrieved those dates instead of temporality itself. This possibility, however, is unlikely in view of the fact that (1) Experiment 1 (in which we did not ask participants to generate event dates) and Experiment 2 produced closely similar results in terms of temporality recall and (2) De Brigard et al. (2020; see also, Mahr et al., 2021) have shown that participants in a similar paradigm were more successful at recalling the general temporal orientation of their simulations than specific dates (suggesting that temporality recall is likely easier than the recall of more specific times).

**The relationship between mnemonicity and temporality recall**—As mentioned at the outset, our procedure was inspired by earlier work from McDonough and Gallo (2010; 2013). These authors found a ‘reality monitoring asymmetry’ in how participants recalled previously generated events: participants committed more source misattributions when asked to judge whether they had previously remembered a past event related to a given cue than when asked whether they had previously imagined a future event. In other words, McDonough and Gallo’s results suggested that participants were more likely to falsely recall imagined future events as remembered than vice versa. However, because McDonough and Gallo’s studies only included remembered events and imagined future events, their design confounded temporal orientation with mnemonicity. Therefore, it remained unclear whether the observed reality monitoring asymmetry was due to confusions of temporality or mnemonicity.

In Experiments 1 and 2 we also observed a reality monitoring asymmetry: participants were consistently more likely to recall imagined past events as remembered than vice versa. However, the fact that such confusions primarily occurred between imagined past and remembered events (and not between remembered and imagined future events) rules out the possibility that this asymmetry was due to confusions in temporal orientation. Moreover, Experiment 3 suggests that this asymmetry was specific to past events. Whereas participants tended to judge past events to be remembered (i.e. previously experienced), for future events they showed the opposite tendency: judging future events preferably to be novel, ‘imagined’ events rather than previously experienced, ‘recast’ events. This pattern suggests that reality monitoring decisions seem to be sensitive to temporal orientation. Temporal orientation might therefore provide information to (possibly subsequent) reality monitoring processes

due to the fact that only past (and not future) events seem to be candidates for being judged as ‘remembered’.

Both the fact that event temporality was consistently better recalled than mnemonicity and the fact that event temporality seems to have influenced mnemonicity-attributions is consistent with the hypothesis that temporality and mnemonicity rely on different representations.

**Memories for the future:** Finally, our results speak to research on ‘memories for the future’ (e.g. Szpunar et al., 2013) and the idea that a crucial function of episodic thought is the generation of simulations of behavior to be recalled and executed at a later point in time (Ingvar, 1979; 1985). Consistent with the idea of a highlighted role of memory for future imaginations, across all of our experiments, imagined future events were consistently more likely to be accurately classified than other imagined events. As De Brigard et al. (2020) have suggested, from this perspective one might think that ‘memory for the future’ does not require the explicit representation of the temporal orientation of a previously generated event simulation. After all, when one draws on a plan made in the past, it is hardly relevant that (in the past) it was directed towards the future. Instead, whether said plan has been already executed or not (i.e., information related to the episode’s mnemonicity) is of primary importance.

In a similar vein, a number of authors have argued that the contents of episodic simulation should be viewed as a-temporal (Boyle, 2020; De Brigard & Gessell, 2016; Klein & Steindam, 2016). According to this view, there is nothing intrinsically past or future about the content of any given simulation (see Mahr et al., 2021 for evidence speaking to this point). Nonetheless, the present study shows that the temporal orientation of previously generated imaginations seem to be better retained than other types of ‘source’ information. While these results do not contradict the claim that episodic contents are not intrinsically temporal, they nonetheless suggest that temporality can play an important role in the generation and retrieval of episodic information. Such a representation of temporality is likely to be conceptually complex (Hoerl & McCormack, 2019). Future research should therefore seek to elaborate both on the details of the processes underlying event temporality and the nature of the conceptual representation of temporality in episodic simulation.

## Acknowledgments

This research was supported by a Mind, Brain, and Behavior Faculty Award by Harvard University, and a Walter Benjamin Fellowship (MA 9499/1–1) by the German Research Foundation to J.B.M. and by grant R01 AG008441 from the National Institute on Aging to D.L.S. We thank Adam Bulley and Felipe De Brigard for comments on an earlier version of this manuscript.

## References

- Addis DR (2018). Are episodic memories special? On the sameness of remembered and imagined event simulation. *Journal of the Royal Society of New Zealand*, 48(2–3), 64–88.
- Addis DR (2020). Mental time travel? A neurocognitive model of event simulation. *Review of Philosophy and Psychology*, 11(2), 233–259.
- Addis DR, Pan L, Vu MA, Laiser N, & Schacter DL (2009). Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering. *Neuropsychologia*, 47(11), 2222–2238. [PubMed: 19041331]

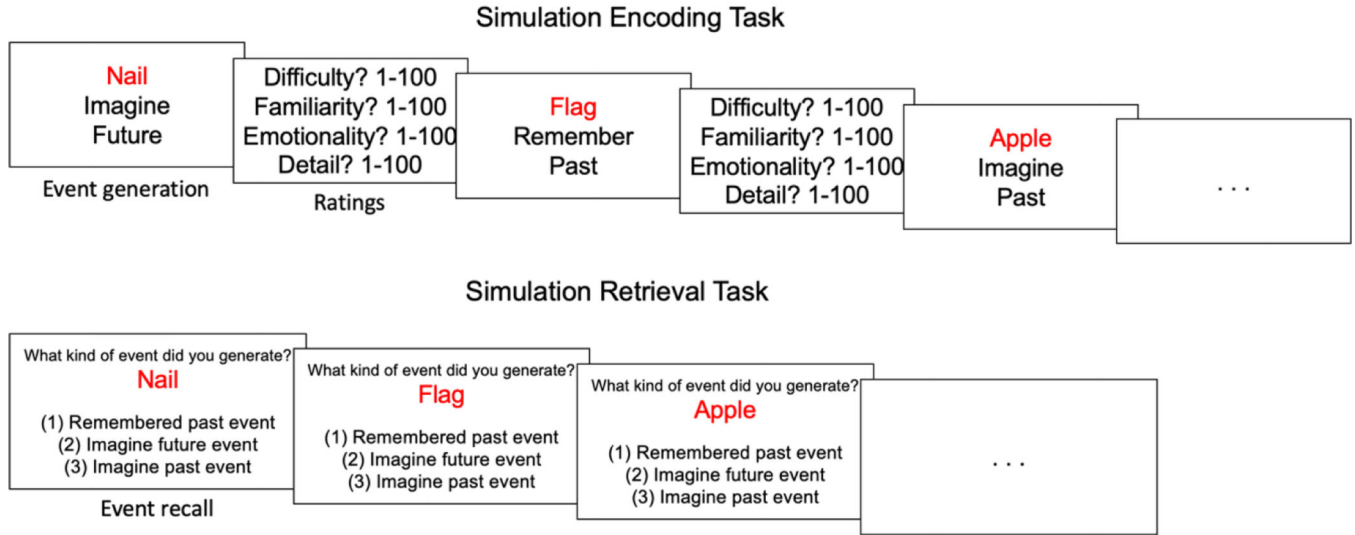
- Addis DR, Wong AT, & Schacter DL (2008). Age-related changes in the episodic simulation of future events. *Psychological Science*, 19(1), 33–41. [PubMed: 18181789]
- Addis DR, Wong AT, & Schacter DL (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, 45(7), 1363–1377. [PubMed: 17126370]
- Barr DJ, Levy R, Scheepers C, & Tily HJ (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Bartlett FC (1932). *Remembering*. Cambridge, UK: Cambridge University Press.
- Bates D, Maechler M, Bolker B, & Walker S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Benoit RG, & Schacter DL (2015). Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. *Neuropsychologia*, 75, 450–457. [PubMed: 26142352]
- Boyle A. (2019). The impure phenomenology of episodic memory. *Mind & Language*, 35(5), 641–660.
- Boyle A. (2020). Remembering events and representing time. *Synthese*. 10.1007/s11229-020-02896-6
- Brown AS (2003). A review of the déjà vu experience. *Psychological Bulletin*, 129(3), 394. [PubMed: 12784936]
- Bulley A, Redshaw J. & Suddendorf T. (2020). The future-directed functions of the imagination: From prediction to metaforesight. In Abraham A. (Ed.), *The Cambridge Handbook of the Imagination*. Cambridge University Press.
- Byrne R. (2005). *The rational imagination: How people create alternatives to reality*. MIT Press.
- Conway MA, Justice LV, & D'Argembeau A. (2019). The self-memory system revisited. In John M. (Ed.), *The Structure and Function of Autobiographical Memory*. Oxford University Press.
- Conway MA, & Rubin DC (1993). The structure of autobiographical memory. In Collins AF, Gathercole SE, Conway MA, & Morris PE (Eds.), *Theories of memory* (p. 103–137). Lawrence Erlbaum Associates.
- Clayton NS, & Russell J. (2009). Looking for episodic memory in animals and young children: Prospects for a new minimalism. *Neuropsychologia*, 47(11), 2330–2340. [PubMed: 18984000]
- Cleary AM, & Claxton AB (2018). Déjà vu: An illusion of prediction. *Psychological Science*, 29(4), 635–644. [PubMed: 29494276]
- Dalla Barba G, Cappelletti JY, Signorini M, & Denes G. (1997). Confabulation: Remembering 'another' past, planning 'another' future. *Neurocase*, 3(6), 425–436.
- D'Argembeau A. (2020). Zooming in and out on one's Life: Autobiographical representations at multiple time scales. *Journal of Cognitive Neuroscience*, 32(11), 2037–2055. [PubMed: 32163320]
- D'Argembeau A, & Van der Linden M. (2004). Predicting the phenomenology of episodic future thoughts. *Consciousness and Cognition*, 21(3), 1198–2106.
- De Brigard F. (2014). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese*, 191(2), 155–185.
- De Brigard F. & Gessell B. (2016). Time is not of the essence. In: Michaelian K, Klein S, & Szpunar KK (Eds.), *Seeing the future: Theoretical Perspectives on Future-oriented Mental Time Travel* (pp. 153–179). Springer.
- De Brigard F, Gessell B, Yang BW, Stewart G, & Marsh EJ (2020). Remembering possible times: Memory for the details of past, future, and counterfactual simulations. *Psychology of Consciousness: Theory, Research, and Practice*, 7(4), 331–339..
- De Brigard F, & Parikh N. (2019). Episodic counterfactual thinking. *Current Directions in Psychological Science*, 28(1), 59–66.
- Epstude K, & Roese NJ (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, 12(2), 168–192. [PubMed: 18453477]
- Espino O, & Byrne RM (2021). How people keep track of what is real and what is imagined: The epistemic status of counterfactual alternatives to reality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(4), 547. [PubMed: 33090843]
- Friedman WJ (1993). Memory for the time of past events. *Psychological Bulletin*, 113(1), 44–66.
- Friedman WJ (2004). Time in autobiographical memory. *Social Cognition*, 22(5), 591–605.

- Friedman WJ (2005). Developmental and cognitive perspectives on humans' sense of the times of past and future events. *Learning and Motivation*, 36, 145–158.
- Hassabis D, Kumaran D, & Maguire EA (2007). Using imagination to understand the neural basis of episodic memory. *Journal of Neuroscience*, 27(52), 14365–14374.
- Hoerl C. & McCormack T. (2019). Thinking in and about time: A dual systems perspective on temporal cognition. *Behavioral and Brain Sciences*, 42(e244), 1–69.
- Ingvar DH (1985). Memory for the future: an essay on the temporal organization of conscious awareness. *Human Neurobiology*, 4, 127–136. [PubMed: 3905726]
- Ingvar DH (1979). “Hyperfrontal” distribution of the cerebral grey matter flow in resting wakefulness: on the functional anatomy of the conscious state. *Acta Neurologica Scandinavica*, 60(1), 12–25. [PubMed: 495039]
- Irish M, Addis DR, Hodges JR, & Piguet O. (2012). Exploring the content and quality of episodic future simulations in semantic dementia. *Neuropsychologia*, 50(14), 3488–3495. [PubMed: 22982607]
- Johnson MK, Hashtroudi S, & Lindsay DS (1993). Source monitoring. *Psychological Bulletin*, 114(1), 3. [PubMed: 8346328]
- Johnson MK, & Raye CL (1981). Reality monitoring. *Psychological Review*, 88(1), 67.
- Johnson MK, Suengas AG, Ann Foley M, & Raye CL (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: Applied*, 117(4), 371–376.
- Klein SB, & Steindam C. (2016). The role of subjective temporality in future-oriented mental time travel. In: Michaelian K, Klein SB, & Szpunar K. (Eds.), *Seeing the future: Theoretical perspectives on future-oriented mental time travel*. (pp. 135–53). Oxford University Press.
- Klein SB (2018). Remembering with and without memory: A theory of memory and aspects of mind that enable its experience. *Psychology of Consciousness: Theory, Research, and Practice*, 5(2), 117.
- Klein SB (2016). Auto-noetic consciousness: Reconsidering the role of episodic memory in future-oriented self-projection. *Quarterly Journal of Experimental Psychology*, 69(2), 381–401.
- Kuznetsova A, Brockhoff PB, & Christensen RH (2017). lmerTest package: tests in linear mixed effects models. *Journal of statistical software*, 82(13), 1–26.
- Mahr JB (2020). The dimensions of episodic simulation. *Cognition*, 196, 104085.
- Mahr JB, Greene JD, & Schacter DL (2021). A long time ago in a galaxy far, far away: How temporal are episodic contents? *Consciousness & Cognition*, 96, 103224.
- McDonough IM, & Gallo DA (2010). Separating past and future autobiographical events in memory: Evidence for a reality monitoring asymmetry. *Memory and Cognition*, 38(1), 3–12. [PubMed: 19966233]
- McDonough IM, & Gallo DA (2013). Impaired retrieval monitoring for past and future autobiographical events in older adults. *Psychology and Aging*, 28(2), 457. [PubMed: 23795764]
- McLelland W, Devitt A, Schacter D, & Addis DR (2015). Making the future memorable: The phenomenology of remembered future events. *Memory*, 23(8), 1255–1263. [PubMed: 25359389]
- Michaelian K, Perrin D, & Sant’Anna A. (2021). Continuities and discontinuities between imagination and memory: The view from philosophy. In: Abraham A. (ed.), *The Cambridge Handbook of the Imagination*. Cambridge University Press.
- Michaelian K. & Sutton J. (2017). Memory. In Zalta EN (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition). <https://plato.stanford.edu/archives/sum2017/entries/memory/>
- Michaelian K. (2016). *Mental time travel*. MIT Press.
- Mitchell KJ, & Johnson MK (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin*, 135(4), 638–677. [PubMed: 19586165]
- Nedjam Z, Dalla Barba G, & Pillon B. (2000). Confabulation in a patient with fronto-temporal dementia and a patient with Alzheimer’s disease. *Cortex*, 36(4), 561–577. [PubMed: 11059455]
- Nyberg L, Kim ASN, Habib R, Levine B, & Tulving E. (2010). Consciousness of subjective time in the brain. *Proceedings of the National Academy of Sciences*, 107(51), 22356–22359.

- Palan S, & Schitter C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22–27.
- Perrin D. (2016). Asymmetries in subjective time. In: Michaelian K, Klein SB, & Szpunar KK (eds.), *Seeing the future: Theoretical perspectives on future-oriented mental time travel*. (pp. 39–61). Oxford University Press.
- Perrin D, Michaelian K, & Sant’Anna A. (2020). The phenomenology of remembering is an epistemic feeling. *Frontiers in Psychology*, 11, 1531. [PubMed: 32719642]
- Pillemer DB, Steiner KL, Kuwabara KJ, Thomsen DK, & Svob C. (2015). Vicarious memories. *Consciousness and Cognition*, 36, 233–245. [PubMed: 26172521]
- R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing.
- RStudio Team (2019). RStudio: Integrated Development for R. RStudio, Inc.
- Rezlescu C, Danaila I, Miron A, & Amariei C. (2020). More time for science: Using Testable to create and share behavioral experiments faster, recruit better participants, and engage students in hand-on research. *Progress in Brain Research*, 253, 243–262. [PubMed: 32771126]
- R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Schacter DL & Addis DR (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society (B)*, 362, 773–786.
- Schacter DL, Addis DR, & Buckner R. (2008). Episodic simulation of future events: Concepts, data, and applications. *Annals of the New York Academy of Sciences*, 1124, 39–60. [PubMed: 18400923]
- Schacter DL, Benoit RG, & Szpunar KK (2017). Episodic future thinking: Mechanisms and functions. *Current opinion in behavioral sciences*, 17, 41–50. [PubMed: 29130061]
- Schacter DL, & Madore KP (2016). Remembering the past and imagining the future: Identifying and enhancing the contribution of episodic memory. *Memory Studies*, 9(3), 245–255. [PubMed: 28163775]
- Schacter DL, Addis DR, Hassabis D, Martin VC, Spreng RN, & Szpunar KK (2012). The future of memory: Remembering, imagining, and the brain. *Neuron*, 76(4), 677–694. [PubMed: 23177955]
- Scott GG, Keitel A, Becirspahic M, Yao B, & Sereno SC (2019). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior research methods*, 51(3), 1258–1270. [PubMed: 30206797]
- Shum MS (1998). The role of temporal landmarks in autobiographical memory processes. *Psychological bulletin*, 124(3), 423. [PubMed: 9849113]
- Simons JS, Garrison JR, & Johnson MK (2017). Brain mechanisms of reality monitoring. *Trends in Cognitive Sciences*, 21(6), 462–473. [PubMed: 28462815]
- Szpunar KK, Spreng RN, & Schacter DL (2014). A taxonomy of prospection: Introducing an organizational framework for future-oriented cognition. *Proceedings of the National Academy of Sciences USA*, 111, 18414–18421.
- Szpunar KK, Addis DR, McLelland VC, & Schacter DL (2013). Memories for the future: new insights into the adaptive value of episodic memory. *Frontiers in Behavioral Neuroscience*, 7(47).
- Thakral PP, Yang AC, Addis DR, & Schacter DL (2021). Divergent thinking and constructing future events: Dissociating old from new ideas. *Memory*, 29(6), 729–743. [PubMed: 34182887]
- Tulving E. (1983). *Elements of episodic memory*. Oxford University Press.
- Tulving E. (2002). Chronesthesia: Conscious awareness of subjective time. In Stuss DT & Knight RT (Eds.), *Principles of frontal lobe function* (pp. 311–325). Oxford University Press.
- Wickham H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer.



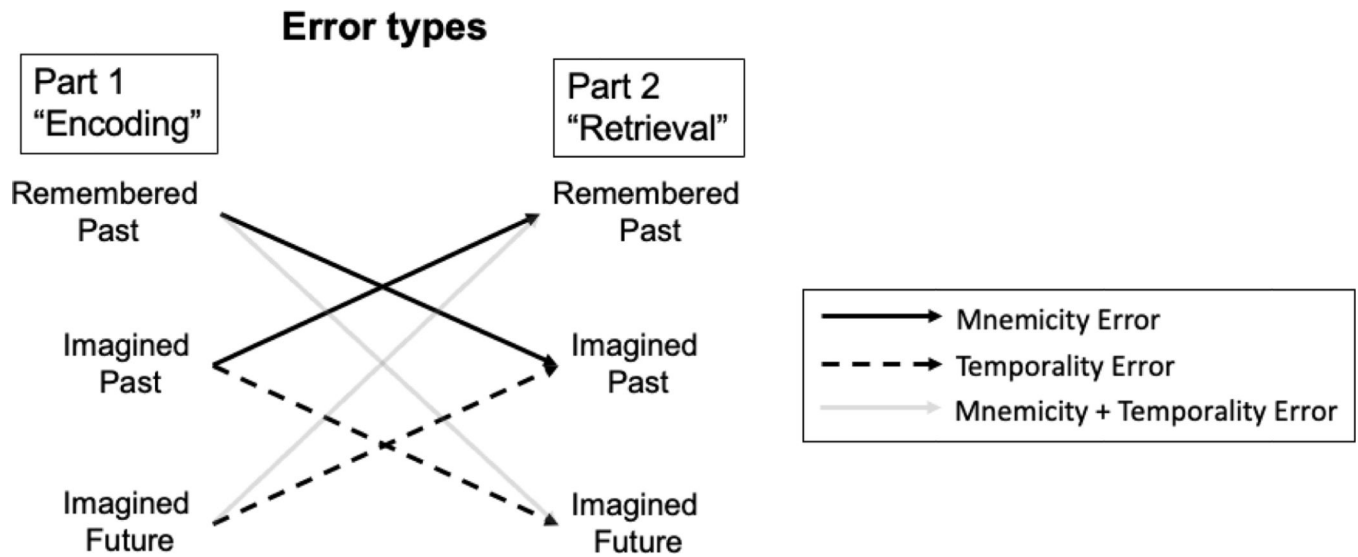
### Experiment 1



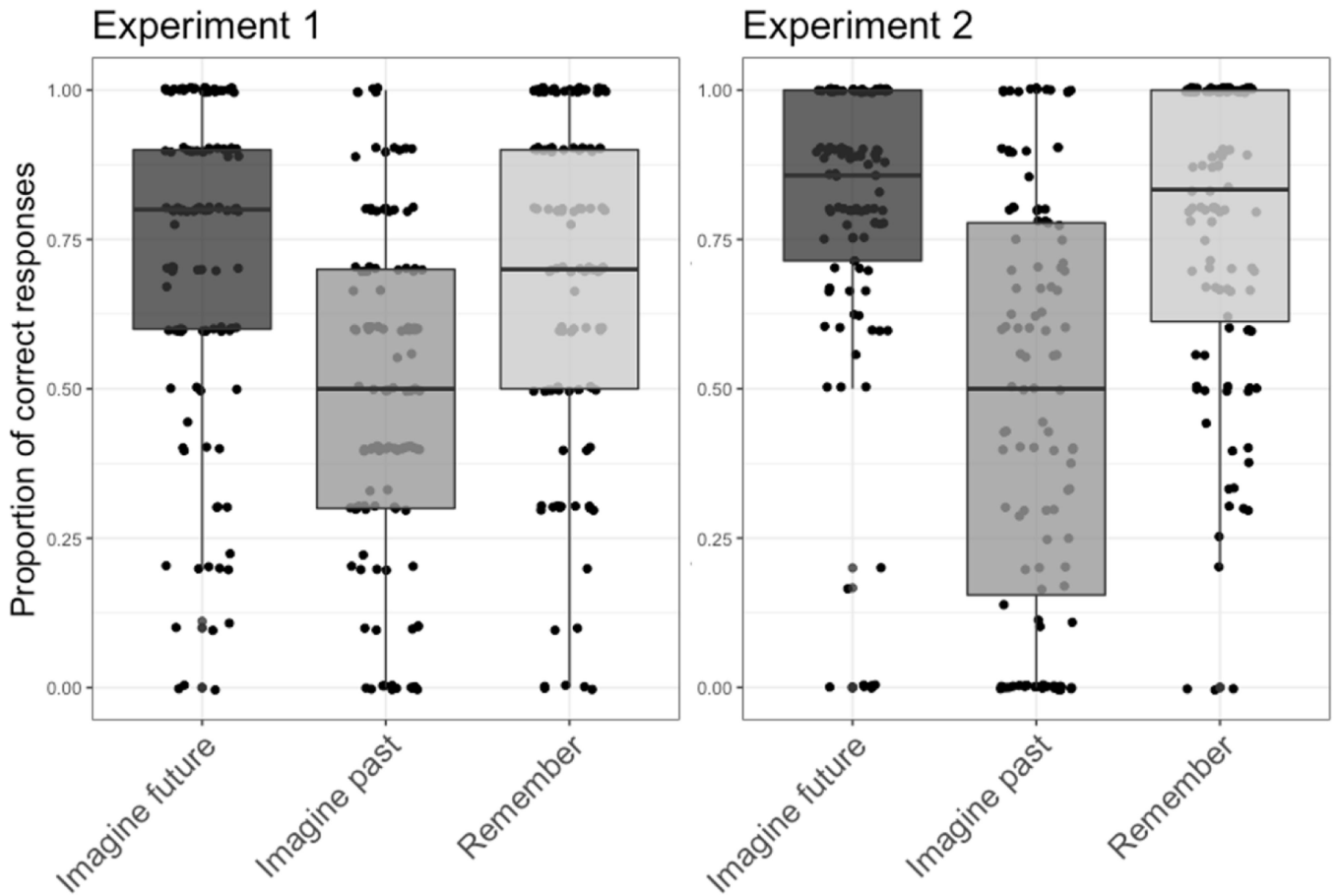
**Figure 1:** Outline of the procedure of Experiment 1. In the first part of the experiment, participants completed a Simulation Encoding Task in which they were asked to generate and describe remembered, imagined past, or imagined future events related to an object word cue. Participants were also asked to rate each event for difficulty, amount of perceptual detail, familiarity, and emotional intensity. In the second part of the experiment, (Simulation Retrieval Task), participants were presented again with each cue word and asked to recall what type of event they had generated for that cue in part 1.

		<b>Mnemicity</b>	
		<i>Remembered</i>	<i>Imagined</i>
<b>Temporality</b>	<i>Past</i>	Remembered Past	Imagined Past (Counterfactual)
	<i>Future</i>	“Remembered” Future (Recast)	Imagined Future

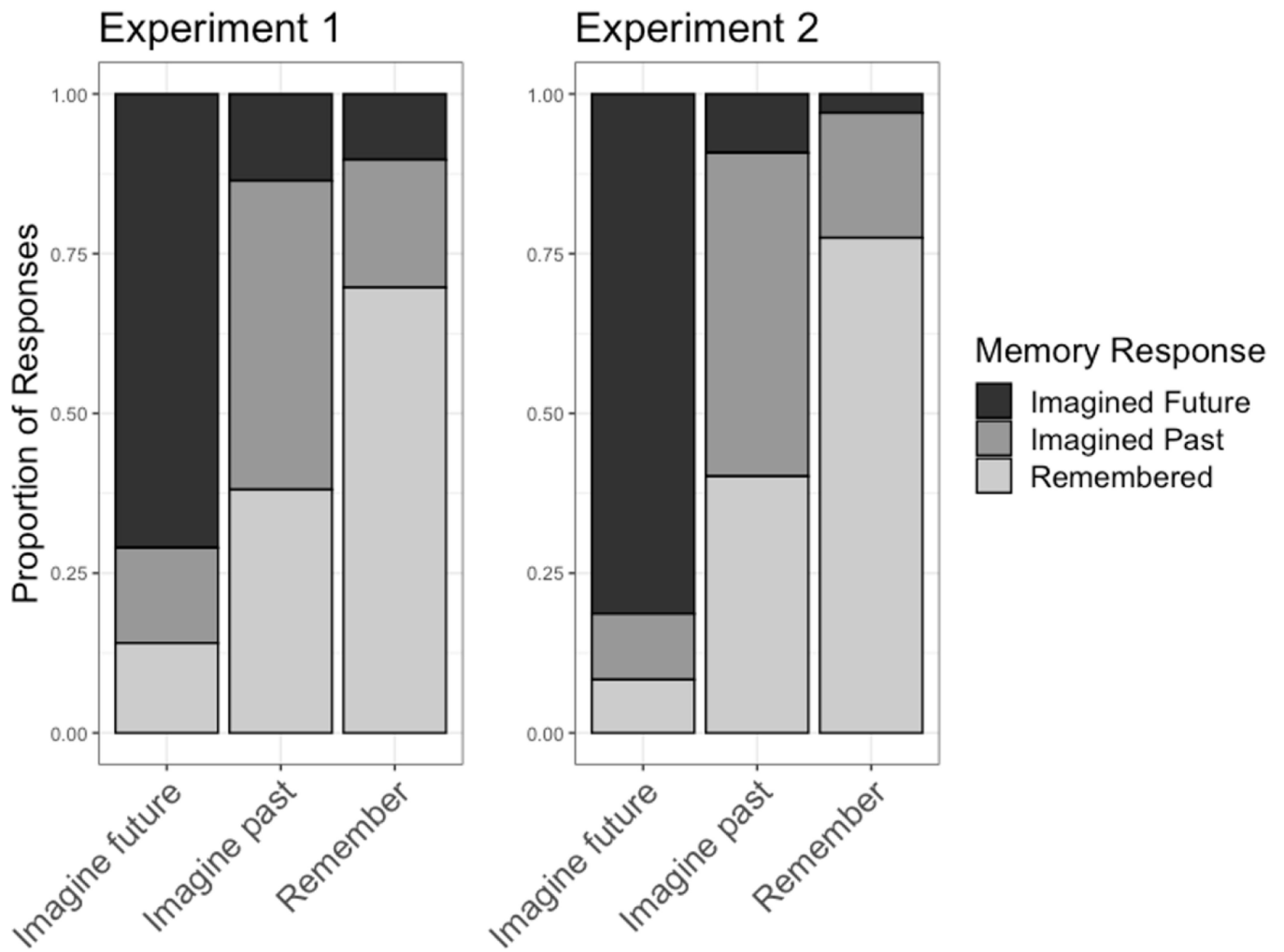
**Figure 2:** Possible representations resulting from combining differences in temporality (past/future) and mnemicity (remembered/imagined). While Experiments 1 and 2, only included Remembered Past, Imagined Past (i.e. counterfactual), and Imagined Future events, Experiment 3 also included “Remembered” Future (i.e. Recast) events.

**Figure 3:**

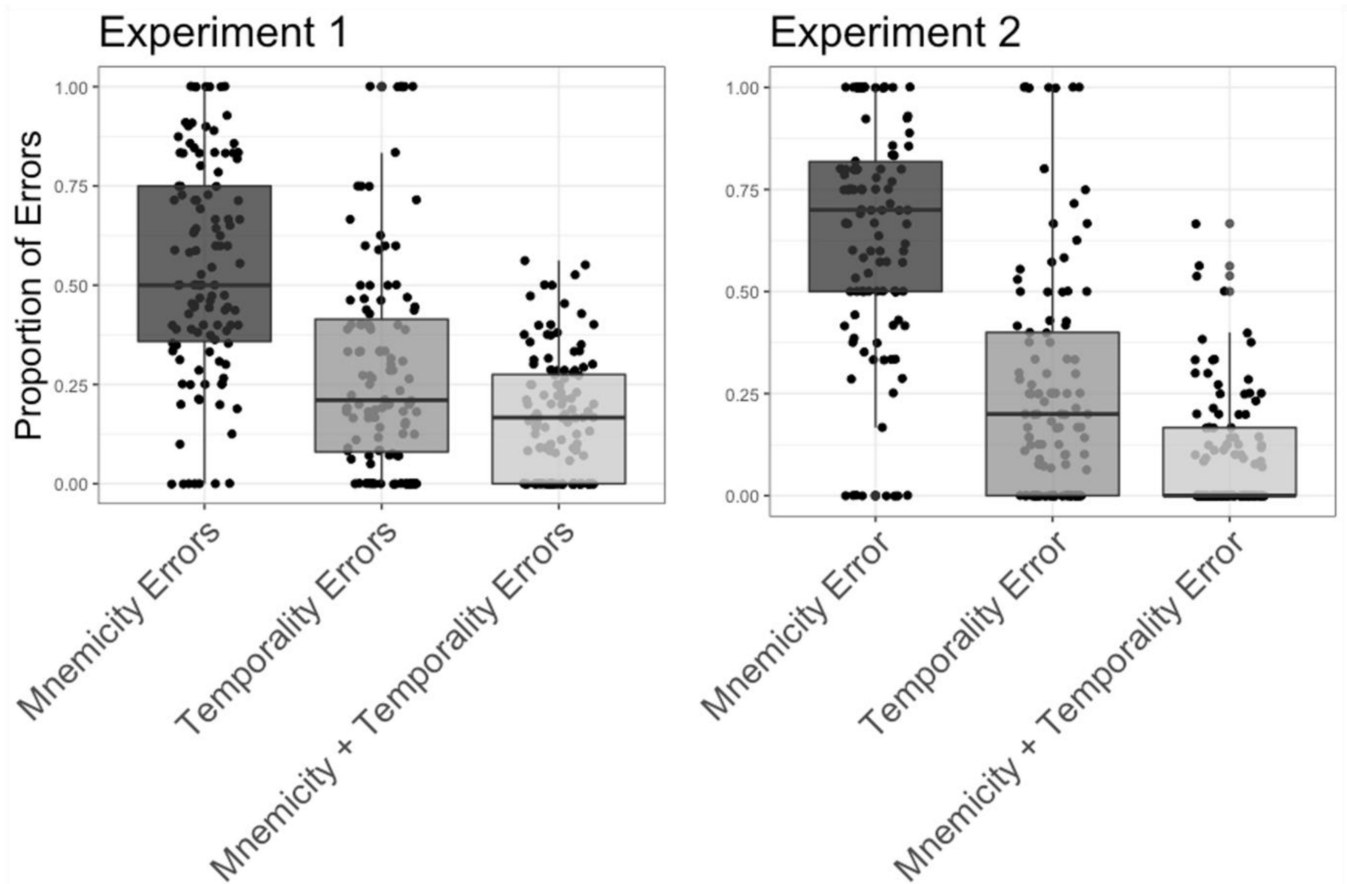
Error types differentiating between the role of temporality and mnemicity in the distinctiveness of different kinds of episodic simulations. Mnemicity Errors (i.e. confusions between remembered past and imagined past events) indicate successful retention of temporality (since these events share the same temporal orientation). In contrast, Temporality Errors (i.e. confusions between imagined past and imagined future events) indicate successful retention of Mnemicity (since these events are both imagined). Mnemicity + Temporality Errors (i.e. confusions between remembered and imagined future events indicate retention of neither mnemicity nor temporality since these events share neither mnemicity nor temporal orientation).



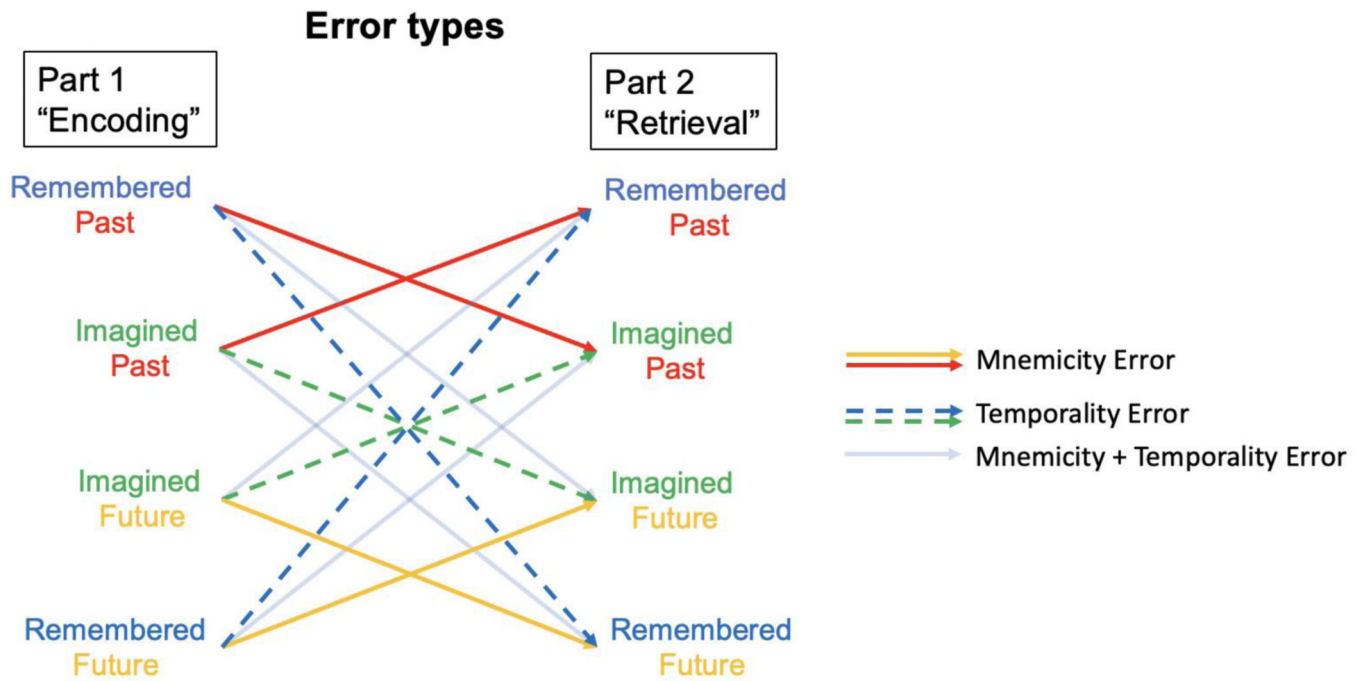
**Figure 4:** Recall performance (proportion of correct responses) in the Simulation Retrieval Task in each event type condition in Experiments 1 and 2. Scatter plots depict proportions for each participant. Boxplot middle bars represent medians, upper and lower hinges correspond to the first and third quartiles (i.e. 25<sup>th</sup> and 75<sup>th</sup> percentiles), and whiskers extend from the hinge to the largest value no further than 1.5 \* IQR from the hinge in either direction (where IQR is the inter-quartile range/the distance between the first and third quartiles).



**Figure 5:** Distribution of different responses in the Simulation Retrieval Task in each event type condition across participants in Experiments 1 and 2.

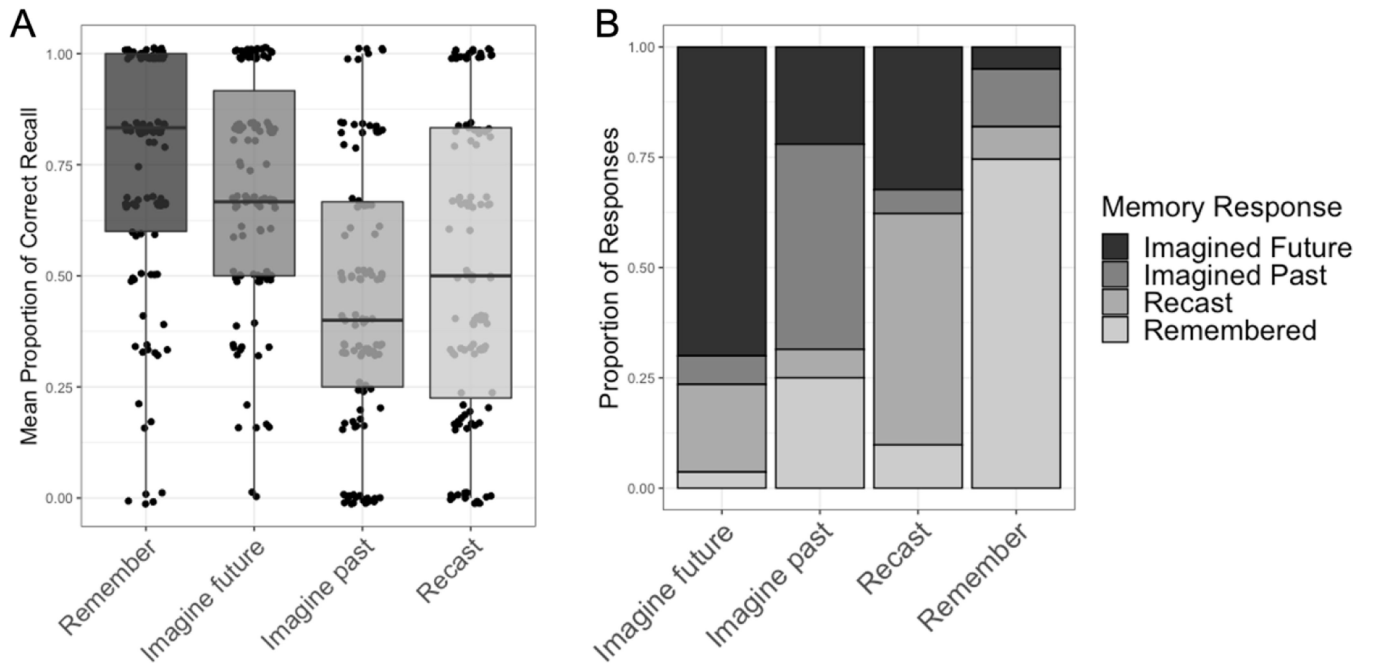


**Figure 6:** Proportions of Mnemicity, Temporality, and Mnemicity + Temporality Errors in the Simulation Retrieval Task of Experiments 1 and 2. In both experiments, participants committed significantly more Mnemicity than Temporality Errors, and more Temporality than Mnemicity + Temporality Errors.



**Figure 7:** Updated error matrix for Mnemicity, Temporality, and Mnemicity + Temporality Errors after the addition of recast/"remembered future" events. Confusions between recast and imagined future events are categorized as Mnemicity Errors, while confusions between recast and remembered events are categorized as Temporality Errors.

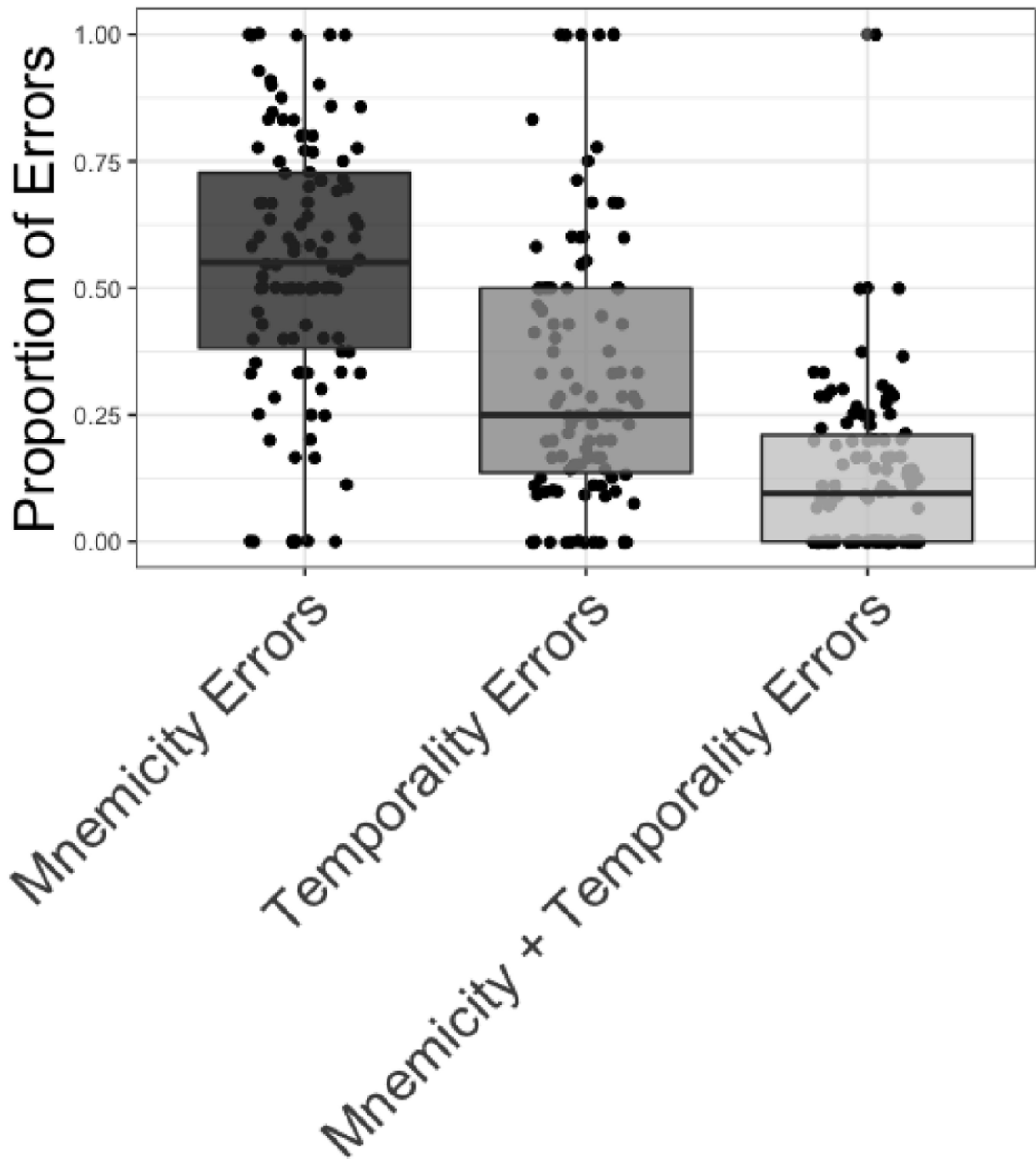
### Experiment 3



**Figure 8: A:** Recall accuracy (proportion of correct responses) in each event type condition in Experiment 3. **B:** Proportion of different memory responses in each event type condition of Experiment 3.

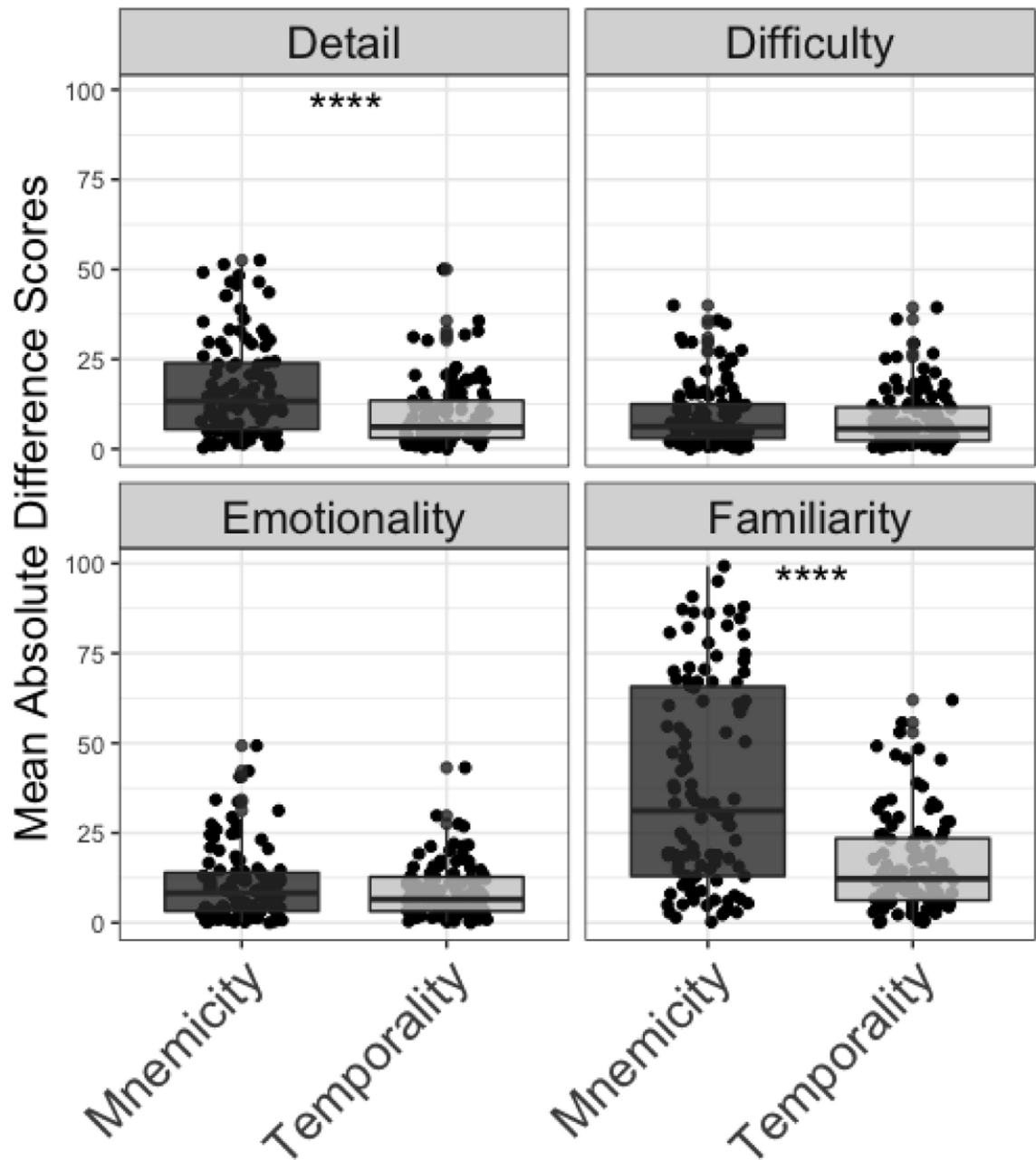


# Experiment 3



**Figure 9:** Distribution of error proportions across different error types in Experiment 3. Confusions between recast and imagined future events were coded as Mnemicity Errors whereas confusions between recast and remembered events were coded as Temporality Errors.

## Experiment 3 - Rating Differences



**Figure 10:** Absolute mean difference scores in event ratings for differences in Mnemicity and Temporality in Experiment 3. Detail and familiarity ratings differed more strongly across remembered (remembered + recast) and imagined (imagined past + future) events than across past (remembered + imagined past) and future (imagined future + recast) events. Differences in difficulty and emotionality ratings did not differentiate between event mnemicity and temporality.

**Table 1:**

Descriptive results for event ratings and memory performance (proportion of correct responses) in Experiments 1, and 2.

Experiment	Condition	Difficulty		Emotionality		Detail		Familiarity		Recall accuracy	
		M	SD	M	SD	M	SD	M	SD	M	SD
Experiment 1	Imagine future	22.2	23.9	29.0	28.3	50.7	29.7	43.5	35.8	0.71	0.45
	Imagine past	23.6	24.0	27.9	28.4	53.6	30.8	47.8	37.8	0.48	0.50
	Remember	17.6	21.8	32.5	29.9	63.3	29.8	71.5	36.8	0.7	0.46
Experiment 2	Imagine future	23.0	24.8	32.0	30.4	53.5	29.5	46.7	37.5	0.81	0.39
	Imagine past	23.5	25.3	31.8	30.4	57.8	29.8	48.6	37.8	0.51	0.50
	Remember	18.4	23.5	36.2	31.4	70.1	28	77.1	32.9	0.77	0.42

**Table 2:**

Descriptive results for Experiment 3. Recall accuracy was calculated as the proportion of correct responses in a given event type condition..

Experiment	Condition	Difficulty		Emotionality		Detail		Familiarity		Recall accuracy	
		M	SD	M	SD	M	SD	M	SD	M	SD
Experiment 3	Imagine future	23.7	23.9	28.8	28.4	52.6	29.6	39.6	37.8	0.70	0.26
	Imagine past	24.5	25.1	24.3	26.2	50.7	28.5	42.3	38.3	0.47	0.3
	Recast	20.6	29.1	32.6	29.1	59.8	29.3	70.7	34.8	0.52	0.33
	Remember	14.4	29.8	35.8	29.8	72.5	26.1	85.8	27.8	0.75	0.28