Journal of Memory and Language 151 (2026) 104796

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Retroactive facilitation and change awareness across lags and retention intervals

Christopher N. Wahlheim [a,*], Geoff Ward [b], Samuel W. Cox [a]

[a] Department of Psychology, University of North Carolina at Greensboro, USA
[b] Department of Psychology, University of Essex, UK

ARTICLE INFO

Keywords:
Retroactive interference
Retroactive facilitation
Study-phase retrieval
Integration
Lag
Retention interval

ABSTRACT

Memory for an earlier event can be impaired by a later event (retroactive interference). However, when later events cue retrieval of earlier events, memory for the earlier event can be enhanced (retroactive facilitation). Two experiments using continuous A-B, A-C paired-associate tasks examined these retroactive memory effects across lags and retention intervals. Associations between change awareness and first-response recall were used as indices of study-phase retrieval and related encoding processes. Participants studied A-B, A-C, and retention-interval-matched A-B pairs and indicated when they detected response changes. The lag between A-B and A-C and the retention interval between A-C and test trials varied across 4, 16, and 48 intervening items. Cued-recall test formats differed: Experiment 1 prompted retrieval of first responses, with participants reporting recollected changes. Experiment 2 used modified-modified free recall in which participants reported responses as they came to mind and indicated whether each response was first, second, or only; the first two decisions indexed recollection of changes. Retroactive facilitation increased across retention intervals. Experiment 1 showed a negatively curvilinear, non-monotonic lag pattern for A-B, A-C recall at the longest retention interval, but this pattern was not supported in Experiment 2. Facilitation occurred on trials in which change awareness persisted from study to test. These findings suggest that retroactive facilitation reflects instances in which study-phase retrieval supports detection and recollection of change, consistent with accounts in which integrative encoding preserves ordered memories. The variation in lag patterns tentatively implicates a contribution from contextual variability, but this awaits confirmation.

Memory for earlier events often depends on how later, related events are encoded. When a later event preserves prior details, it can support memory for the earlier event, but when it introduces a change, it can either impair or enhance memory depending on whether the earlier event is retrieved during subsequent study. Classic interference theories propose that learning a related association with changed details (A-C) impairs memory for an earlier association (A-B), producing retroactive interference (for a review, see MacLeod, 2024). However, later events can also cue retrieval of and improve memory for earlier events (Bruce & Weaver, 1973), a phenomenon known as retroactive facilitation. More recent evidence suggests that these divergent outcomes depend on the extent to which changes are noticed and remembered (for reviews, see Wahlheim et al., 2021; Wahlheim & Zacks, 2025). When A-C pairs prompt study-phase retrieval of A-B pairs, shared and distinctive features can be coactivated and encoded in relation to one another (e.g., "A was paired with B, but now with C"), and directing retrieval during change perception promotes this process and subsequent facilitation (Jacoby et al., 2015). The present work builds on this framework of memory for changes by characterizing how self-reported change awareness relates to retroactive facilitation across lags and retention intervals, with implications for explaining variations in retroactive memory effects over time.

Retroactive interference over time

Early studies reported greater retroactive interference when tests followed shortly after interpolated learning (Briggs, 1954; Underwood, 1948). Retroactive interference was initially attributed to response competition between A-B and A-C associations and to unlearning, whereby intruding A-B responses during A-C learning were assumed to

* Corresponding author at: Department of Psychology, University of North Carolina at Greensboro, 296 Eberhart Building, P. O. Box 26170, Greensboro, NC 27402-6170, USA.
E-mail address: cnwahlhe@uncg.edu (C.N. Wahlheim).

https://doi.org/10.1016/j.jml.2026.104796
Received 15 March 2025; Received in revised form 21 July 2026; Accepted 23 July 2026
Available online 30 July 2026

be extinguished (Melton & Irwin, 1940). Subsequent evidence challenged unlearning and suggested that interference reflects competition between independently stored traces whose relative accessibility changes over time (Koppenaal, 1963). This view was later formalized by proposing that exponentially decaying trace strengths determine interference magnitude, such that retroactive and proactive effects depend on the relative strengths of A-B and A-C at test (Baddeley, 1976). As the retention interval increases, retroactive interference diminishes as the interpolated trace decreases in relative strength, whereas proactive interference increases as competing trace strengths converge.

An alternative account adapted the Search of Associative Memory (Raaijmakers & Shiffrin, 1981) model but attributed time-dependent interference patterns to contextual fluctuation rather than trace decay (Mensink & Raaijmakers, 1988). In this view, the same cue can be encoded in different contexts across occasions, and interference reflects retrieval competition operating against a drifting context representation. Using retrieval competition and time-varying context alone, the model accounted for decreasing retroactive interference and increasing proactive interference across retention intervals. Like the decay-based model, this model assumes independent A-B and A-C traces, and does not require integrative mechanisms.

Related temporal context accounts provide a complementary perspective on time-dependent interference. Early models assumed that temporal context drifted gradually over time, thus modifying cue-context match at retrieval (e.g., Glenberg & Swanson, 1986). The Temporal Context Model (Howard & Kahana, 2002) advanced this view by proposing that context evolves as a function of studied and retrieved items, allowing the model and later retrieved-context variants to capture a range of free recall dynamics and multi-list interference effects (e.g., Lohnas et al., 2015). However, these retrieved-context models have not been adapted to account for retroactive interference differences across retention intervals, and they primarily address interference rather than facilitation. Thus, although contextual updating mechanisms offer tools for understanding time-dependent interference effects, additional processes would be required to explain retroactive facilitation effects, which we consider further in the General Discussion.

In sum, these accounts explain time-dependent interference effects framed as competition between independent traces whose accessibility changes due to decay or contextual dynamics. Whether framed in terms of trace strength or cue-context match, the critical mechanism involves relative accessibility at test rather than interaction between the traces during study. Notably, these frameworks do not incorporate retrieval practice during A-C study or awareness of change as mechanisms that could modify the accessibility of earlier associations. Consequently, although they successfully explain temporal modulation of interference, they do not directly address conditions under which later learning enhances memory for earlier associations. This limitation motivates consideration of additional mechanisms that support retroactive facilitation.

Retroactive facilitation at various lags and retention intervals

Although dominant accounts of retroactive memory effects emphasize interference, interpolated learning can also enhance memory for original associations. Early reports of retroactive facilitation showed that learning A-C pairs sometimes improved later recall of A-B pairs, especially when the responses were semantically similar (e.g., afraid-scared), whereas dissimilar responses produced interference (Barnes & Underwood, 1959). These findings were explained in terms of mediated associations, but they also suggest that semantic similarity may have increased the likelihood that A-C pairs evoked retrieval of the earlier A-B association during study. When A-B was elicited in this manner, the resulting coactivation of shared cues and related responses may have supported strengthened or more elaborated memory representations.

Subsequent studies identified additional conditions under which retroactive facilitation emerged. Using short lists of paired associates, Bruce and Weaver (1973) observed facilitation in first-response recall when A-C pairs followed A-B pairs, especially when the changed pairs repeated. Although the magnitude of facilitation varied across retention intervals, the pattern suggested that repeated cues served as reminders that reinstated A-B pairs during A-C learning. Similarly, in a continuous paired-associate procedure, (Robbins & Bray, 1974) found larger facilitation at shorter lags between A-B and A-C pairs, consistent with the idea that recently encoded A-B pairs were more likely to be retrieved during A-C study. They further proposed that participants' awareness that responses had changed encouraged relational processing, increasing the likelihood that A-C pairs served as retrieval cues for A-B pairs. As they noted:

"Let us assume that the stimulus, A, represents a retrieval cue. The second presentation of A may serve as a cue to retrieve the association A-B, and this may be a function of both the A-B, A-C lag and the retention interval. Further, assume that the act of retrieval results in additional, different retrieval cues being formed with regard to A-B. This view would then lead to an expectation of retroactive facilitation." (p. 425).

Across these studies, retroactive facilitation was most evident under conditions that plausibly increased study-phase retrieval, such as semantic relatedness, repeated cues, and shorter lags, suggesting that facilitation reflects trace interactions during encoding rather than competition at test alone. This provides a foundation for examining how retrieval and awareness of change determine whether later learning impairs or enhances memory for earlier associations.

Study-phase retrieval and change awareness in retroactive facilitation effects

More recent work has directly tested whether retroactive facilitation arises from study-phase retrieval of earlier associations and awareness that responses have changed. In a key experiment, Jacoby et al. (2015) manipulated whether A-C pairs prompted retrieval of earlier A-B pairs during study. Participants first learned a list of A-B pairs followed by a second list containing both repeated A-B pairs and changed A-C pairs, with unique A-B pairs in each list. Critically, half of the A-C pairs corresponded to A-B pairs from the first list and half corresponded to A-B pairs from the second list, producing longer lags for between-list than within-list changes. Study-phase retrieval was manipulated through instructions given before the second list. One group (N-Back) was asked to indicate whether each response had changed from earlier in the experiment, which required retrieving A-B pairs from the first list when studying between-list A-C pairs. A second group (Within-List Back) indicated only whether responses had changed relative to the second list, discouraging retrieval of first-list responses. Compliance with these instructions was verified through change-detection reports with the N-Back group detecting more between-list changes. Critically, retroactive facilitation in later recall of first-list A-B pairs emerged only in the N-Back group, suggesting that retrieval of earlier responses during A-C study enhanced memory for the prior associations.

A collection of studies examining how awareness of change associates with memory provide converging support for memory-for-change frameworks (Wahlheim & Jacoby, 2013; Wahlheim & Zacks, 2019), which were inspired by the recursive reminding hypothesis (Hintzman, 2010). In A-B, A-C paradigms assessing retroactive memory effects, these accounts propose that studying an A-C pair can cue retrieval of the earlier A-B association. When this retrieval occurs, participants may detect that the response has changed by comparing the retrieved and current responses ($B \neq C$). Accordingly, this comparison serves as an encoding event that strengthens accessibility of the earlier response through retrieval practice and allows the coactivated B and C responses to become integrated within a representation of the change. Retroactive facilitation can arise in part because retrieval practice increases the accessibility of the earlier response. Moreover, integrated representations that include information about how responses changed can also guide retrieval by allowing participants to recollect that the earlier response preceded the later one, thereby avoiding erroneous reporting of the updated response. The extent to which task conditions promote these processes can be assessed by measuring awareness of change at study (detection) and memory for those changes at test (recollection).

This account does not propose that recollecting change at test causally improves memory for initially learned responses. Instead, recollecting change is as an empirical marker that is correlated with instances in which A-B associations were retrieved and integrated with A-C pairs earlier during study. Operationally, recollecting change occurs when participants report remembering that responses differed across study episodes, which can be assessed in various ways. However, recollecting change does not guarantee successful recall of the original response. For example, someone might remember that a password changed without recalling the earlier password. Rather, it indicates occasions when study-phase retrieval likely occurred and when the earlier and later responses were encoded in relation to one another. Because such integration depends on successful retrieval of A-B during A-C study, more accessible initial memories are more likely to be retrieved and therefore more likely to benefit from additional retrieval practice and integrative encoding. Consequently, recollection of change indexes conditions under which earlier associations had sufficient accessibility to be retrieved and received additional encoding support from comparison with current perceptions that can increase the probability of later recall.

Importantly, this conditional relationship is not tautological because it can be modulated by factors that affect the likelihood of retrieving A-B during A-C study. For example, extending study time for A-C pairs increases opportunities for retrieval and comparison processes, leading to higher rates of change recollection and greater retroactive facilitation (Garlitch & Wahlheim, 2020; Negley et al., 2018). Similarly, providing interpolated retrieval practice with feedback for A-B pairs prior to A-C study increases the probability that those associations are retrieved when A-C pairs are encountered, which enhances subsequent memory for A-B associations (Wahlheim et al., 2023). Converging evidence also comes from manipulations of encoding strategy. For example, instructing participants to form integrative images linking A-B and A-C pairs produces statistical dependence between B and C responses and retroactive facilitation, whereas distinctive encoding leads to statistical independence and retroactive interference (Burton et al., 2017). Together, these findings indicate that retroactive facilitation emerges when task conditions increase the likelihood that earlier associations are retrieved and encoded in relation to later ones.

Of particular relevance here, similar principles appear to govern retroactive memory effects across retention intervals. Retroactive facilitation in A-B, A-C paradigms has been examined while varying the semantic relationship between responses and the delay before test (Antony et al., 2022). In that work, facilitation occurred only for related responses and was greater after a 48-h than a 5-min retention interval. This likely reflected semantic relatedness increasing retrieval of A-B pairs during A-C study, which promoted integrative encoding when changes were detected, with the benefits of integrative encoding emerging over time. Consistent with this interpretation, statistical dependence between recall of both responses, an indicator of integrative encoding (see Horner & Burgess, 2013), has been shown to be greater at the delayed interval. Related work using a similar task showed that when change detection during A-C study was not measured, a 5-min interval produced retroactive interference that was eliminated at a 24-h interval, whereas requiring overt change detection and retrieval of earlier responses during A-C study produced facilitation at a 24-h interval (Cox et al., 2026). Conditional analyses showed that facilitation occurred when participants recollected that responses had changed, that its magnitude increased with retention interval, and that overt detection increased the frequency of change recollection. Collectively, these findings suggest that retroactive facilitation is most evident at longer test delays when study-phase retrieval and detection of changes occur, because reductions in retroactive interference over time allow the benefits of integrative encoding to emerge.

In sum, these studies suggest that retroactive memory effects depend on whether study events cue retrieval of earlier associations and whether changes are encoded. When A-B pairs are retrieved during A-C study, retrieval practice increases the accessibility of the A-B association, and comparison of responses can promote encoding that integrates them within a representation of change. Although much of the memory benefit for the A-B association likely reflects retrieval practice, recollection of change can also guide accurate memory reports. These processes occur selectively across items, depending on whether the A-B association is sufficiently accessible during A-C study to be retrieved and compared, which in turn determines whether change is detected and later recollected. The likelihood of this cascade varies with intrinsic differences in item memorability and task conditions that support study-phase retrieval, such as semantic associations, cue repetition, and lag between study events. Moreover, because retroactive interference diminishes with increasing retention interval, the benefits of retrieval practice and integrative encoding should become most apparent at longer delays. Thus, facilitation should be most evident when task conditions support study-phase retrieval and detection of change during A-C study, especially when sufficient time has elapsed for interference to diminish.

The present study

Building on the preceding framework, the present study examined how change awareness and retroactive memory effects vary across lags and retention intervals. The aim was to establish empirical regularities that any account of retroactive facilitation over time must explain. To this end, we examined how awareness of A-C changes, which often reflect study-phase retrieval of A-B pairs, relates to retroactive facilitation across within-subject combinations of lag and retention interval using continuous paired-associate paradigms with intermixed study and test trials. Participants reported whether they detected response changes during study and whether they recollected those changes at test. Because detecting changes can cue retrieval of earlier associations, requiring overt detection judgments was expected to stimulate retrieval of A-B pairs during A-C study and thereby promote retroactive facilitation. Consistent with evidence that retroactive interference diminishes with increasing delay, we further expected facilitation to become more apparent at longer retention intervals. We also expected greater change awareness when earlier associations remained accessible at shorter lags and retention intervals, and that retroactive facilitation would occur primarily for detected and recollected changes.

Although no existing model was developed specifically to explain the outcomes from the present A-B, A-C task, the lag-repetition literature provides a useful heuristic for generating predictions about how lag and retention interval may relate to recall and change awareness. Increasing the lag between repetitions of a stimulus generally improves memory, although the magnitude of this benefit varies across retention intervals (Cepeda et al., 2008; Glenberg, 1976). A theoretically meaningful feature of these lag functions is their frequent non-monotonicity, with optimal lag lengths increasing as retention intervals lengthen. Contextual variability accounts propose that repeated items are encoded in more varied contexts at longer lags, which improves later retrieval cues, but such models predict only monotonic gains in memory (e.g., Glenberg, 1976). Study-phase retrieval accounts predict non-monotonic lag functions because retrieval practice benefits are greatest when retrieval is challenging but successful, which is most likely to occur at intermediate lags (e.g., Greene, 1989). Dual-mechanism accounts proposes that lag effects reflect contributions from both contextual variability and study-phase retrieval (Benjamin & Tullis, 2010). In this view, contextual fluctuation across lags influences the likelihood of successful retrieval, whereas memory durability reflects strengthening that occurs when earlier occurrences are successfully retrieved during study.

Accordingly, recall benefits at intermediate lags and longer retention intervals arise because a sufficient number of memories strengthened by challenging yet successful retrieval remain accessible when interference has diminished.

A key difference between spaced repetition and A-B, A-C paradigms is that repetition paradigms present stimuli twice, typically producing performance gains as lag increases, whereas A-B, A-C paradigms repeat only the cues of earlier pairs and often produce performance losses across lags due to forgetting of the original responses. In the present paradigm, A-C pairs were expected to cue retrieval of A-B pairs, producing lag functions for first-response recall and change recollection that generally decline as lag increases. However, to the extent that contextual variability and challenging study-phase retrieval promote durable memories, these lag functions may show non-monotonicity, with optimal performance at intermediate lags and longer retention intervals, when the most durable memories remain accessible after retroactive interference has diminished. Accordingly, retroactive facilitation should vary with lag and retention interval and be most evident on trials in which changes are detected and later recollected. Moreover, such variations in recall enhancement should be non-uniform, reflecting differences in interactions of the challenge and success of study-phase retrieval and the retroactive interference still present at test.

Data availability

The deidentified data upon which the study conclusions are based, the code necessary to reproduce the analyses, and the study materials are available here: https://osf.io/cvmqx/.

Experiment 1

Experiment 1 examined how retroactive facilitation and change awareness vary across lags and retention intervals using a continuous paired-associate A-B, A-C paradigm. The lag between A-B and A-C pairs and the retention interval before test were manipulated within subjects to vary the likelihood that A-C pairs would cue retrieval of A-B associations and the degree to which retroactive interference had diminished at test. The procedure included cue- response word pairs presented on study trials with interspersed cued-recall tests. On study trials, some cues reappeared later with a different response (e.g., ocean-*spray*, then ocean-*salty*). Other cues appeared only once (e.g., pasture-graze) and provided a baseline for assessing retroactive memory effects on A-B, A-C trials. During study, participants indicated whether the response had changed from an earlier pair, thus measuring change detection. On test trials, participants recalled the first or only response and then indicated whether they remembered that responses had changed. These measures allowed us to characterize the variations in retroactive facilitation across lags and retention intervals in association with change awareness during study and test.

Materials and methods

Participants

The final sample included 120 participants who contributed usable data (60 from University of North Carolina at Greensboro [UNCG], and 60 from University of Essex [UofE]). An initial stopping rule was established at UNCG to collect usable data from at least 60 and at most 75 participants in one semester. This rule was based on timing and available resources. If the target sample of 60 participants was not reached, then testing would continue into the next semester until that target was reached. Usable data from 60 participants at UNCG were collected in just over one semester. Data from three additional participants were excluded (one received the wrong program, and two indicated recollecting change across items types for ≥98% of trials, suggesting a lack of task engagement). The comparisons of interest included conditional cells with sparse observations. We extended the stopping rule without considering specific statistical outcomes to mitigate the risk of false positives. We collected usable data from 60 participants at UofE. Data from eight additional participants were excluded (five because the full counterbalancing had been reached, and three because they were observed not making a genuine effort).

Because this was the first experiment of its kind, we did not have existing data upon which to conduct an a priori power analysis. Instead, we conducted a sensitivity analysis on the highest-order interaction from the model of primary interest, namely, comparing first-response recall across lags (quadratic) and retention intervals for both item types. We based the analyses on this outcome because it provides the most direct index of whether A-B memories were retained despite the potential of interference from A-C pairs. The change detection and recollection measures offer valuable information about the observed retroactive effects of memory, but they are the outcomes of processes reflected in first-response recall. Focusing on first-response recall thus allowed us to ground simulations in the outcome most central to our hypotheses about study-phase retrieval, integrative encoding, and the predicted non-monotonic lag functions across retention intervals. The analysis indicated that the sample provided 80% power to detect a small effect size ($OR = 1.19$). Specific details about the model and sensitivity analysis, including power curves, are provided in the Supplementary Material (Section 1).

Design

Fig. 1A illustrates the two main features of this within-subjects design. First, for A-B, A-C items, we factorially-combined two independent variables: the Lag, defined as the number of trials intervening between the A-B and A-C pairs with three levels (4, 16, and 48), and the Retention Interval, defined as the number of trials intervening between the A-C pairs and test prompts (A -?) with three levels (4, 16, and 48). This produced nine within-subject conditions.

The second feature allowed for the comparison of single presentations of A-B control pairs with A-B, A-C pairs, matched for the total retention interval (i.e., the number trials intervening between A-B items and test trials). There were six total retention intervals for the A-B, A-C pairs, which were calculated by adding one (the presentation of A-C) to the sum of the lag and retention interval (i.e., 9, 21, 33, 53, 65, or 97 intervening items). These were the retention intervals for the six A-B control pair conditions. Overall, this design produced 15 within-subject conditions, which resulted in the same number of counterbalanced versions of the experiment.

Materials

The materials included 255 cue words (e.g., cheese) that could be paired with one or two response words (e.g., cheddar; cracker). The stimuli were designated as primacy buffers (27), fillers (78), or critical items (150). Primacy buffers were arranged into A-B, A-C and A-B item types that oriented participants to task before starting the actual experiment. Fillers appeared as single study items in the actual experiment to fill gaps among critical items. Buffers and fillers appeared in the same positions in all versions of the experiment. Critical items appeared in the 15 within-subject conditions. These were counterbalanced by rotating 15 groups of 10 unique three-word sets so that each set appeared equally often in each condition across participants.

For primacy buffers, cues ranged from 3 to 9 letters ($M = 5.73$, $SD = 1.80$), first responses from 3 to 11 letters ($M = 5.53$, $SD = 1.72$), and second responses from 3 to 10 letters ($M = 5.68$, $SD = 1.61$). Word frequency, indexed by log hyperspace analogue to language values (Lund & Burgess, 1996), was comparable across cues ($M = 8.44$, $SD = 1.52$), first responses ($M = 9.19$, $SD = 1.80$), and second responses ($M = 9.03$, $SD = 2.01$). The average forward associative strength, according to the University of South Florida (USF) free association norms (Nelson et al., 2004), between cues and first responses was 0.04 ($SD = 0.04$, range = 0.00–0.15), and between cues and second responses was 0.07 ($SD = 0.09$, range = 0.00–0.44). Excluding five cues that were not in the

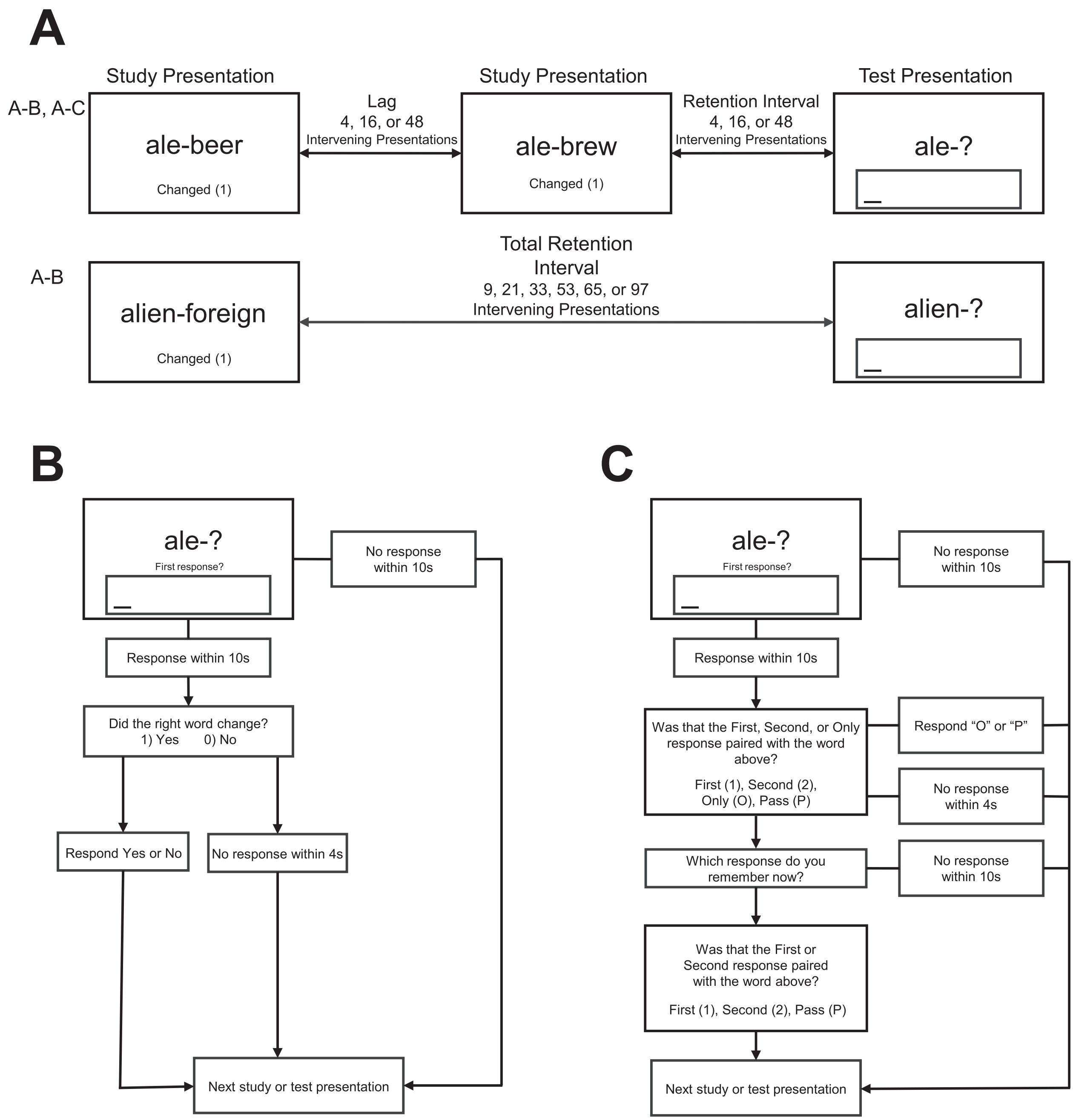

Fig. 1. Schematic of the experimental conditions and procedure.
Note. (A) Continuous paired associate learning task in Experiments 1 and 2. (B) Test trial structure in Experiment 1. (C) Test trial structure in Experiment 2.

USF norms, the average forward associative strength between first and second responses was 0.01 (SD = 0.05, range = 0.00–0.27), and the average backward associative strength was 0.02 (SD = 0.07, range = 0.00–0.36). For filler items, cues and responses were 3–9 letters long (M = 5.36, SD = 1.42; M = 5.10, SD = 1.47, respectively). Word frequency was comparable for cues (M = 8.39, SD = 1.47) and responses (M = 8.66, SD = 1.96). Excluding three cues that were not in the USF norms, the average forward associative strength between cues and responses was 0.06 (SD = 0.07, range = 0.00–0.33).

For critical items, cues were 3–9 letters long (M = 5.37, SD = 1.57) and responses were 3–8 letters long (M = 4.73, SD = 1.07). Within each critical item set, first and second responses were identical in length. Word frequency was comparable across cues (M = 9.35, SD = 1.58), first responses (M = 9.38, SD = 1.65), and second responses (M = 9.35, SD = 1.79). Forward associative strength values were available for 147 of the 150 cues. The average forward associative strength between cues and first responses was 0.07 (SD = 0.10, range = 0.00–0.60) and between cues and second responses was 0.04 (SD = 0.08, range = 0.00–0.64). Excluding 10 responses without forward associative strength values, first and second responses had an average forward associative strength of 0.02 (SD = 0.06, range = 0.00–0.42); excluding 11 responses without backward associative strength values first and second responses had an average backward associative strength of 0.02 (SD = 0.07, range = 0.00–0.46).

Procedure

All participants were tested individually with an experimenter present. At UNCG, the experimenter remained in the room throughout the experiment. At UofE, the experimenter remained in the room until they were assured that participants understood the instructions (~ 10 min), then watched from an observation room. The stimulus presentation was controlled by *E*-prime software (Version 2, *Psychology Software Tools*, 2012). All stimuli appeared centered on the screen in white font against a black background. A continuous paired associate learning task included a single list of 518 intermixed study and test presentations. Primacy buffers appeared for the first 50 presentations and included seven tested A-B, A-C items (21 presentations), nine tested A-B items (18 presentations), and 11 untested A-B items (11 presentations). The actual experimental items appeared next, comprising 468 critical item and filler presentations. The critical items included 90 tested A-B, A-C items (270 presentations) and 60 tested A-B items (120 presentations). The fillers included 78 untested A-B items (78 presentations). The primacy portion transitioned seamlessly into the experiment portion without informing participants, and experimenters used the first few minutes to ensure that participants understood the instructions.

The study trials (Fig. 1 A) displayed word pairs for 6 s each followed by a 0.5 s blank screen interstimulus interval (ISI). Pairs appeared in a pre-randomized order for all participants, with the constraint that no more than two pairs from the same condition appeared consecutively. Participants were told that their tasks would be to read the pairs aloud, study them for an upcoming test, and indicate whether the pairs included changed response words. For each pair, participants were told to press "1" key to indicate that the right-hand word changed or do nothing when there was no such change. All pairs remained on the screen for the entire 6 s study interval.

The test trials (Fig. 1 B) displayed a cue from a study trial paired with a question mark for a self-paced duration of up to 10 s. Participants were told to type the first (or only) response that was paired with the cue. After a response was entered, it was replaced by a prompt asking if participants remembered the response changing earlier (by pressing "1") or if it did not change (by pressing "0"). They had up to 4 s to respond, then the program advanced automatically.

Statistical methods

All analyses were conducted in R (R Core Team, 2025). Logistic mixed-effects models were fit with the *glmer* function from the *lme4* package (Bates et al., 2015) using the *bobyqa* optimizer (maxfun = 20,000). Lag and retention interval were treated as continuous fixed effects and z-scored (mean-centered, unit SD) to place them on a common scale and enable model convergence. This linear transformation does not change the functional form or substantive interpretation because differences among 4, 16, and 48 items are preserved on the standardized scale. To allow curvature, which was necessary to examine the potential non-monotonicity of lag functions, every model included a quadratic lag term ($\text{Lag}^2 = [\text{z-lag}]^2$). Models used crossed random intercepts for subjects and items, (1 | Subject) + (1 | Cue), reflecting the counterbalanced assignment of items to conditions. Potential differences between participants from the two universities were evaluated with likelihood-ratio tests comparing a full model including a fixed effect of Site to a reduced model without that fixed effect. We retained the reduced model unless including Site significantly improved fit. For clarity and generalizability, effects collapsed over Site are reported in the main text. When the full model was retained, site-specific effects are reported in the Supplementary Material (Section 2). Including Site improved model fit in some instances, but it never changed patterns relevant to the primary research questions.

For the following models, when Item type (A-B, A-C vs. A-B) is the only categorical fixed effect (two levels), we report Wald odds ratios (*ORs*) with 95% confidence intervals (*CIs*) from the fitted model. When analyses include change awareness as the only categorical fixed effect (three levels, which we describe below), we report Wald χ^2 omnibus tests for main effects and interactions using the *Anova* function from the *car* package (Fox & Weisberg, 2019). To contextualize effect magnitudes, we interpret *ORs* using standard guidelines (Chen et al., 2010), where *ORs* of 1.22, 1.86, and 3.00 correspond to small, medium, and large effects, respectively. Where relevant, we follow up with pairwise contrasts from the *emmeans* function from the *emmeans* package (Lenth, 2024), reported as *ORs* with 95% CIs. Linear and quadratic lag trends were estimated with the *emtrends* function from *emmeans* at standardized retention intervals. Model-estimated probabilities and *CIs* displayed in figures come from fixed-effects-only estimates with estimated marginal means on the probability scale. The figures also show observed means as white-filled diamonds, computed as the raw proportion of the outcome, pooled across all trials and participants contributing to that cell.

To descriptively characterize the shape of the model-estimated lag functions, we examined the linear and quadratic Lag terms. Because the models included both of these terms, we calculated the quadratic vertex and evaluated slopes at the endpoints of the observed lag range. Functions were classified as monotonic when the endpoint slopes had the same sign and the vertex fell outside the observed lag range, and as non-monotonic when the vertex fell within the observed range. These classifications summarize the shapes of the fitted functions and do not constitute an additional inferential test beyond the regression model.

Finally, we summarized explanatory power with Nakagawa-Schielzeth pseudo-R^2 effect size summaries derived using the *r2* function from the *performance* package (Lüdecke et al., 2021). Marginal R^2 (R^2_M) reflects variance explained by the fixed effects only, and conditional R^2 (R^2_C) reflects variance explained by the full model (fixed + random effects). Model specifications are provided in Wilkinson–Rogers notation for each model description below and in the analysis scripts on OSF (https://osf.io/cvmqx/). The alpha level for significance was 0.05 (two-tailed).

Results

First-response recall

We expected retroactive facilitation to increase across retention intervals and be optimal at an intermediate lag and long retention interval. To characterize retroactive effects of memory, we compared first-response recall across item types, lags, and retention intervals with the model: logit[P(recall)] ~ Item type × Retention interval × Lag + Item type × Retention interval × Lag^2 + (1 | Subject) + (1 | Item), which included 21,600 observations ($R^2_M = 0.04$, $R^2_C = 0.35$). Table 1 (left) reports fixed-effect *ORs*; Table 2 (top) reports slope parameters and monotonicity of lag functions; Fig. 2 (top) shows the model-estimated probabilities. A significant Item type × Retention interval interaction indicated higher recall for A-B, A-C than A-B items that increased across retention intervals. A significant Retention interval × Lag interaction indicated that negative lag functions became shallower across retention intervals. Finally, a significant Item type × Retention interval × Lag^2 interaction showed that the curvilinearity of lag functions increased across retention intervals. Specifically, functions for A-B, A-C items became more negatively curvilinear, shifting from monotonic to non-monotonic between lags 16 and 48, whereas functions for A-B items became more positively curvilinear, shifting from monotonic to non-monotonic between lags 4 and 16. Forgetting slowed across retention intervals, and retroactive facilitation was greatest at the intermediate lag and longest retention interval.

Change detection on study trials

We expected change detection on study trials to decrease across increasing lags, reflecting decreasing study-phase retrieval at longer intervals. Participants reported whether the response had changed from the first to the second pair. Change responses were coded as correct for the second pair of A-B, A-C items and incorrect for the only pair of A-B items. These change-detection responses were then compared across item types, lags, and retention intervals with the model: logit[P(change detection)] ~ Item type × Retention interval × Lag + Item type × Retention interval × Lag^2 + (1 | Subject) + (1 | Item), which included 21,600 observations ($R^2_M = 0.63$, $R^2_C = 0.75$). Table 3 (left) reports fixed-

Table 1
Odds ratios from mixed-effects logistic regression models of first-response recall.

	Experiment 1		Experiment 2	
Predictor	*OR (95% CI)*	*p*	*OR (95% CI)*	*p*
Item type	**0.53 (0.46, 0.61)**	**< 0.001**	**0.45 (0.37, 0.55)**	**< 0.001**
Retention interval	1.03 (0.94, 1.14)	= 0.508	0.95 (0.83, 1.08)	= 0.421
Lag	**0.73 (0.69, 0.78)**	**< 0.001**	**0.71 (0.65, 0.77)**	**< 0.001**
Lag^2	0.95 (0.87, 1.03)	= 0.194	0.91 (0.80, 1.02)	= 0.107
Item type × Retention interval	**0.71 (0.62, 0.81)**	**< 0.001**	**0.81 (0.67, 0.98)**	**= 0.034**
Item type × Lag	0.94 (0.86, 1.02)	= 0.125	1.00 (0.88, 1.12)	= 0.938
Retention interval × Lag	**1.14 (1.07, 1.21)**	**< 0.001**	**1.25 (1.15, 1.36)**	**< 0.001**
Item type × Lag^2	**1.23 (1.09, 1.39)**	**< 0.001**	**1.22 (1.03, 1.45)**	**= 0.021**
Retention interval × Lag^2	**0.87 (0.80, 0.94)**	**< 0.001**	**0.87 (0.78, 0.99)**	**= 0.029**
Item type × Retention interval × Lag	1.00 (0.92, 1.09)	= 0.996	1.02 (0.91, 1.15)	= 0.724
Item type × Retention interval × Lag^2	**1.16 (1.02, 1.30)**	**= 0.019**	1.10 (0.93, 1.30)	= 0.279

Note. Predictors for Retention interval and Lag were standardized (mean-centered and scaled to unit variance) prior to model estimation to address model convergence issues. Odds ratios (ORs) reflect the multiplicative change in the odds of first-response recall associated with a one standard deviation increase in the predictor, holding other variables constant. Interaction terms represent combined effects of standardized predictors. For categorical contrasts, A-B, A-C items served as the reference level for Item type. Statistically significant ORs are displayed in bold.

effect *ORs*; Table 4 (top) summarizes the shapes of the model-estimated lag functions descriptively; Fig. 3 (top) shows the model-estimated probabilities. A significant Item type × Lag interaction indicated high change detection accuracy that declined across lags and was invariant across retention intervals (which occurred after the manipulation).

Change recollection on test trials

We expected change recollection on test trials to decrease across increasing lags, partly reflecting differences in change detection. However, change recollection is also based partly on first-response accessibility, which is sensitive to interactions of lags and retention intervals. Consequently, we expected change recollection functions to resemble first-recall functions, more than change detection functions, and potentially show non-monotonicity at the longest retention interval. Similar to the change detection measure above, participants reported whether they remembered that response had changed from the first to the second pair across earlier study trials. Change responses were coded as correct for A-B, A-C items and incorrect for A-B items. These change-recollection responses were then compared across item types, lags, and retention intervals with the model: logit[P(change recollection)] ~ Site × Item type × Retention interval × Lag + Site × Item type × Retention interval × Lag^2 + (1 | Subject) + (1 | Item), which included 21,600 observations ($R^2_M = 0.39$, $R^2_C = 0.52$). Site is included because the full model improved fit, $\chi^2(12) = 30.80$, $p = .002$ (for Site-specific effects, see Supplementary Materials, Section 2.1).

Table 5 (left) reports fixed-effect *ORs*; Table 6 (top) summarizes the shapes of the model-estimated lag functions descriptively; Fig. 4 (top) shows the model-estimated probabilities. A significant Item Type × Retention Interval × Lag interaction showed that the negative linearity of lag functions for A-B, A-C items became shallower across intervals and flat at the longest interval. Conversely, the lag functions for A-B items were essentially flat. This yielded increasingly smaller significant differences between accurate and inaccurate responses across intervals 4, 16, and 48 at lag 4 (*ORs*: 48.56 > 30.39 > 19.02) and lag 16 (*ORs*: 30.54 > 24.29 > 19.31), but not lag 48 (*ORs*: 19.21 ≈ 19.41 ≈ 19.61), all *ps* < 0.001. Importantly, the interaction involving Item type, Lag^2, and Retention interval was not significant (Table 5), indicating that the predicted change in curvature across retention intervals was not statistically supported. Although the estimated function for A-B, A-C items appears flatter at the longest retention interval (Table 6; Fig. 4), the model does not provide statistical evidence for the non-monotonic pattern we predicted.

First-response recall conditionalized on change classifications

We examined how awareness of change was associated with first-response recall for A-B, A-C items by conditionalizing recall on three levels of change awareness. Importantly, these analyses are not intended to imply that recollecting change at test causally improves memory for the earlier response. Rather, recollection of change serves as an empirical marker that indexes occasions when the earlier A-B association was likely retrieved during A-C study and encoded in relation to the changed response. Detecting change requires some accessibility of the earlier A-B association. Consequently, items for which change is detected and recollected are expected to include cases in which the earlier response was relatively accessible. However, detection also reflects an additional study-phase event in which the retrieved A-B association is compared with perceived A-C pair. This can produce retrieval-practice benefits for

Table 2
Slope parameters and monotonicity of lag functions in first-response recall.

Item type	Retention interval	*Lag b (95% CI)*	*Lag b^2 (95% CI)*	Slope min	Slope max	Vertex (x*)	Monotonicity
Experiment 1							
A-B, A-C	4	−0.44 (−0.53, −0.36)	0.09 (−0.03, 0.21)	−0.62	−0.20	69.70	Monotone decreasing
A-B, A-C	16	−0.31 (−0.38, −0.25)	−0.06 (−0.14, 0.03)	−0.20	−0.47	−29.10	Monotone decreasing
A-B, A-C	48	−0.19 (−0.27, −0.10)	−0.20 (−0.32, −0.08)	0.22	−0.73	14.06	Non-monotone
A-B	4	−0.51 (−0.60, −0.43)	0.15 (0.03, 0.27)	−0.81	−0.10	54.19	Monotone decreasing
A-B	16	−0.38 (−0.44, −0.32)	0.15 (0.07, 0.24)	−0.69	0.03	46.14	Non-monotone
A-B	48	−0.25 (−0.34, −0.17)	0.15 (0.03, 0.27)	−0.56	0.16	38.14	Non-monotone
Experiment 2							
A-B, A-C	4	−0.49 (−0.58, −0.41)	0.02 (−0.10, 0.15)	−0.54	−0.43	208.31	Monotone decreasing
A-B, A-C	16	−0.36 (−0.42, −0.30)	−0.05 (−0.14, 0.03)	−0.25	−0.50	−41.26	Monotone decreasing
A-B, A-C	48	−0.22 (−0.30, −0.14)	−0.13 (−0.25, −0.01)	0.04	−0.57	6.69	Non-monotone
A-B	4	−0.47 (−0.55, −0.38)	0.10 (−0.02, 0.22)	−0.67	−0.20	66.21	Monotone decreasing
A-B	16	−0.32 (−0.38, −0.26)	0.08 (−0.01, 0.16)	−0.48	−0.11	61.32	Monotone decreasing
A-B	48	−0.18 (−0.26, −0.09)	0.05 (−0.07, 0.18)	−0.29	−0.03	52.45	Monotone decreasing

Note. b = linear slope in log-odds; b^2 = quadratic slope in log-odds. Slope min and slope max represent the estimated slopes at the boundaries of the observed lag range. Functions were classified as monotone decreasing if both were negative. If slopes differed in sign, monotonicity depended on the vertex (x*): when the turning point fell within the observed lag range, the function was classified as non-monotone; when it fell outside the range, the function was classified as monotone.

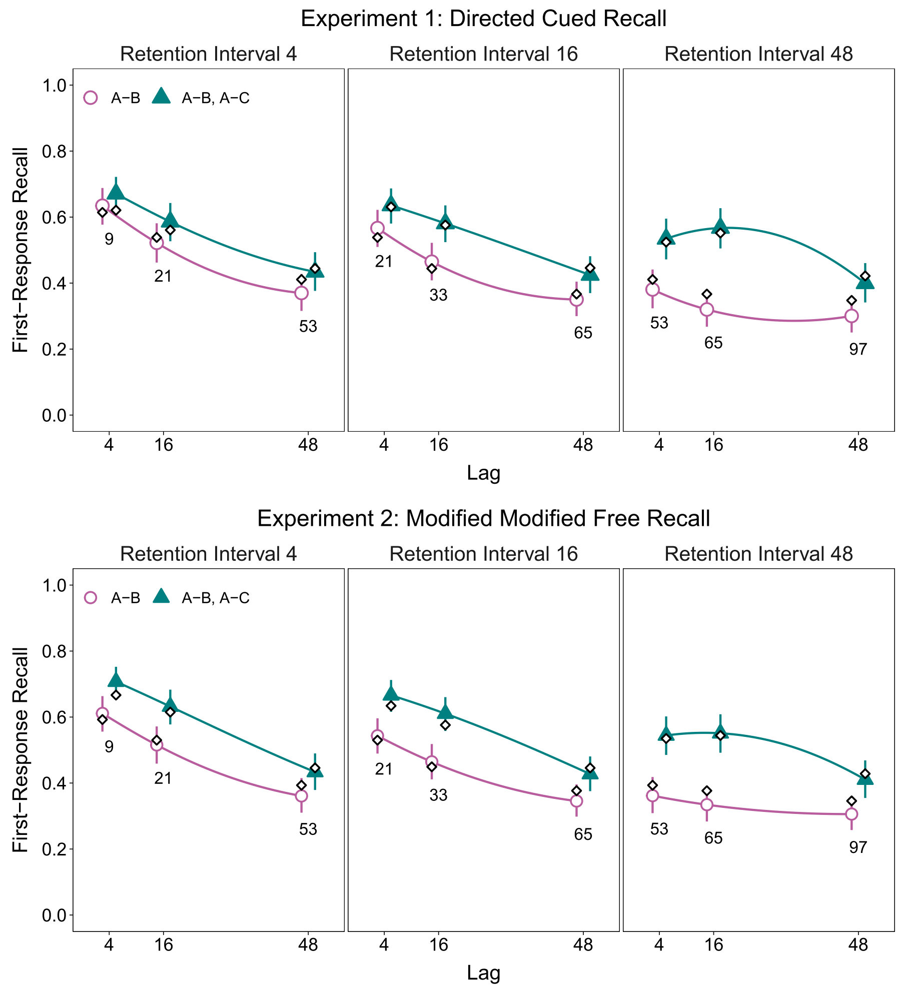

Fig. 2. Probabilities of first-response recall.
Note. Estimates are model-derived from a mixed-effects analysis; error bars represent 95% confidence intervals. Observed means appear as white-filled diamonds. Numeric values beneath the A-B points indicate total retention intervals between study and test. Corresponding odds ratios for fixed effects and slope estimates of the lag functions are reported in Tables 1 and 2.

the earlier response and promote encoding of diagnostic information about how the responses changed. These processes are expected to occur only on trials in which the earlier response is retrieved during A-C study and therefore provide additional encoding support beyond the initial A-B study episode. Because our counterbalancing scheme distributes initial accessibility across lags and retention intervals, differences associated with change awareness cannot be attributed solely to baseline memory accessibility. Rather, it also reflects whether the study episode engendered retrieval and integration processes that singly presented A-B control items do not receive.

To examine how these processes relate to first-response recall, we categorized A-B, A-C items according to three levels of change awareness. *Change Detected & Recollected* included items for which participants detected the changed A-C pair during study and later recollected that the cue had been associated with different responses. These trials are most likely to reflect instances when the earlier response was retrieved and integrated with the changed response. *Change Detected & Not Recollected* included items for which change was detected during study but not later recollected, which likely reflects instances when retrieval occurred but change information was encoded less effectively. *Change Not Detected* included items for which change was not detected during study and therefore was least likely to involve retrieval of the earlier response during A-C study. Any recollection responses for these items reflects non-diagnostic bases for change judgments (e.g., cue familiarity or guessing).

We expected first-response recall to scale with change awareness, being highest when changes were detected and recollected, intermediate when changes were detected but not recollected, and lowest when changes were not detected, akin to prior findings (Negley et al., 2018). However, we had no specific predictions about how the shapes of the lag functions would vary across these levels of change awareness. Indeed, one goal of the present study was to provide the first characterization of how these associations unfold over time. The associations between "changed" classifications and first-response recall were compared across lags and retention intervals with the model: logit[P(recall)] ~ Classification × Retention interval × Lag + Classification × Retention interval × Lag^2 + (1 | Subject) + (1 | Item), which included 10,800 observations ($R^2_M = 0.26$, $R^2_C = 0.44$). Table 7 (left) reports fixed-effect *ORs*; Table 8 (top) summarizes the shapes of the model-estimated lag functions descriptively; Fig. 5 (top) shows the model-estimated probabilities.

A significant effect of Classification indicated that first-response recall was significantly higher when detected changes were recollected than not recollected, $OR = 4.26$, $p < .001$; recall was also significantly higher for both instances when changes were detected than not detected, smallest $OR = 4.58$, $p < .001$. A significant Classification × Lag interaction showed steeper negative linear slopes when changes were not detected than for the other classifications. A significant Retention Interval × Lag interaction also showed shallower negative

Table 3
Odds ratios from mixed-effects logistic regression models of change detection on study trials.

	Experiment 1		Experiment 2	
Predictor	*OR (95% CI)*	*p*	*OR (95% CI)*	*p*
Item type	**0.00 (0.00, 0.00)**	**< 0.001**	**0.00 (0.00, 0.00)**	**< 0.001**
Retention interval	0.97 (0.87, 1.09)	= 0.640	1.05 (0.94, 1.18)	= 0.373
Lag	**0.55 (0.51, 0.60)**	**< 0.001**	**0.58 (0.54, 0.63)**	**< 0.001**
Lag^2	1.01 (0.91, 1.12)	= 0.801	1.04 (0.93, 1.15)	= 0.497
Item type × Retention interval	0.96 (0.70, 1.32)	= 0.818	1.01 (0.75, 1.36)	= 0.967
Item type × Lag	**1.87 (1.53, 2.28)**	**< 0.001**	**1.79 (1.48, 2.16)**	**< 0.001**
Retention interval × Lag	0.99 (0.92, 1.07)	= 0.777	1.05 (0.98, 1.13)	= 0.186
Item type × Lag^2	0.96 (0.73, 1.27)	= 0.789	1.09 (0.83, 1.43)	= 0.547
Retention interval × Lag^2	1.04 (0.94, 1.15)	= 0.497	0.95 (0.86, 1.05)	= 0.326
Item type × Retention interval × Lag	0.98 (0.81, 1.20)	= 0.868	0.87 (0.72, 1.04)	= 0.132
Item type × Retention interval × Lag^2	1.05 (0.79, 1.39)	= 0.728	1.10 (0.85, 1.43)	= 0.476

Note. Predictors for Retention interval and Lag were standardized (mean-centered and scaled to unit variance) prior to model estimation to address model convergence issues. Odds ratios (ORs) reflect the multiplicative change in the odds of first-response recall associated with a one standard deviation increase in the predictor, holding other variables constant. Interaction terms represent combined effects of standardized predictors. For categorical contrasts, A-B, A-C items served as the reference level for Item type. OR values displayed as 0.00 for Item type reflect R's default rounding and indicate small but nonzero estimated odds ratios. Such values mean that the likelihood of recall for that predictor level (A-B) was near zero compared with the reference category (A-B, A-C). Statistically significant ORs are displayed in bold.

linear slopes across intervals. In addition, significant Classification × Lag^2 and Retention Interval × Lag^2 interactions indicated that the curvature of the lag functions varied across classifications and retention intervals. As summarized descriptively in Table 8, lag functions showed little, if any, curvature when changes were detected and recollected, relatively stronger negative curvature when changes were detected but not recollected, and positive curvature when changes were not detected. Across increasing retention intervals, lag functions also became more negatively or less positively curved. Together, these results show that the slopes and curvature of lag functions differed across change classifications and retention intervals, suggesting that the aggregate lag functions reflect a mixture of patterns associated with different levels of change awareness.

Discussion

Experiment 1 provided evidence for both retroactive interference and retroactive facilitation. Consistent with interference accounts, recall decreased with the number of intervening events between study and test (in the single A-B pairs), and recall of A-B, A-C pairs was lower than retention-interval-matched A-B pairs when study-phase retrieval did not occur (as indexed by lack of change detection). These patterns are consistent with accounts in which A-B and A-C events are encoded as independent traces whose relative accessibility at test determines recall (Baddeley, 1976; Mensink & Raaijmakers, 1988). Although such accounts explain reductions in retroactive interference across retention intervals, they do not predict conditions in which interpolated learning improves memory for earlier A-B associations.

By contrast, Experiment 1 showed superior recall on A-B, A-C trials relative to retention-interval-matched A-B pairs, and this facilitation increased across retention intervals, consistent with attenuated interference and amplified facilitation at longer delays (Antony et al., 2022; Briggs, 1954; Cox et al., 2026). Conditional analyses showed that this facilitation was largely driven by trials in which participants detected changes at study and later recollected those changes at test. Recall scaled with the persistence of change awareness from study to test, being highest when changes were detected and recollected, intermediate when detected but not recollected, and lowest when changes were not detected. This pattern is consistent with accounts in which study-phase retrieval supports encoding of the prior response as well as relations between responses and improves later access to earlier information (Jacoby et al., 2015).

Finally, recall at the longest retention interval was highest at an intermediate lag. Such non-monotonic lag functions have been attributed to dual contributions of contextual variability and study-phase retrieval (Benjamin & Tullis, 2010), with benefits emerging when retrieval is both successful and sufficiently demanding. An intermediate lag may therefore maximize recall at longer intervals by balancing retrieval success and difficulty. However, although this pattern appeared descriptively in both experiments, it was only reliably observed here in Experiment 1.

Table 4
Slope parameters and monotonicity of lag functions in change detection on study trials.

Experiment 1							
Item type	**Retention interval**	***Lag b (95% CI)***	***Lag b^2 (95% CI)***	**Slope min**	**Slope max**	**Vertex (x*)**	**Monotonicity**
A-B, A-C	4	−0.58 (−0.69, −0.48)	−0.02 (−0.17, 0.12)	−0.54	−0.64	−218.77	Monotone decreasing
A-B, A-C	16	−0.59 (−0.67, −0.52)	0.01 (−0.09, 0.12)	−0.62	−0.56	440.26	Monotone decreasing
A-B, A-C	48	−0.60 (−0.71, −0.50)	0.05 (−0.10, 0.19)	−0.70	−0.47	137.61	Monotone decreasing
A-B	4	0.06 (−0.20, 0.33)	−0.11 (−0.48, 0.25)	0.28	−0.24	27.85	Non-monotone
A-B	16	0.03 (−0.15, 0.22)	−0.02 (−0.28, 0.23)	0.08	−0.03	35.45	Non-monotone
A-B	48	0.01 (−0.25, 0.26)	0.06 (−0.31, 0.43)	−0.11	0.17	21.63	Non-monotone
Experiment 2							
Item type	**Retention interval**	***Lag b (95% CI)***	***Lag b^2 (95% CI)***	**Slope min**	**Slope max**	**Vertex (x*)**	**Monotonicity**
A-B, A-C	4	−0.59 (−0.70, −0.48)	0.09 (−0.06, 0.23)	−0.76	−0.35	85.34	Monotone decreasing
A-B, A-C	16	−0.54 (−0.61, −0.46)	0.04 (−0.07, 0.14)	−0.61	−0.44	163.08	Monotone decreasing
A-B, A-C	48	−0.49 (−0.59, −0.39)	−0.02 (−0.16, 0.13)	−0.46	−0.53	−262.66	Monotone decreasing
A-B	4	0.13 (−0.12, 0.39)	0.07 (−0.30, 0.44)	−0.02	0.34	5.94	Non-monotone
A-B	16	0.04 (−0.13, 0.22)	0.12 (−0.13, 0.37)	−0.20	0.37	19.26	Non-monotone
A-B	48	−0.05 (−0.27, 0.18)	0.16 (−0.16, 0.49)	−0.38	0.40	25.35	Non-monotone

Note. b = linear slope in log-odds; b^2 = quadratic slope in log-odds. Slope min and slope max represent the estimated slopes at the boundaries of the observed lag range. Functions were classified as monotone decreasing if both were negative. If slopes differed in sign, monotonicity depended on the vertex (x*): when the turning point fell within the observed lag range, the function was classified as non-monotone; when it fell outside the range, the function was classified as monotone.

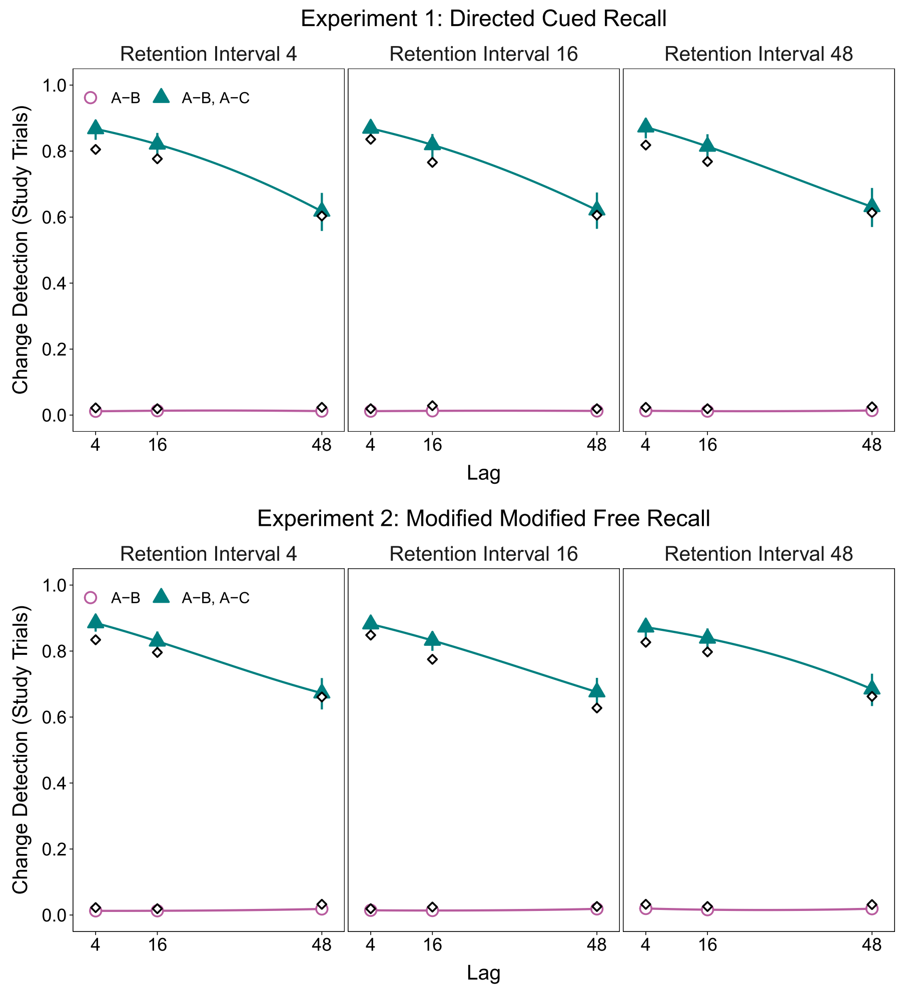

Fig. 3. Probabilities of change detection on study trials.
Note. Estimates are model-derived from a mixed-effects analysis; error bars represent 95% confidence intervals. Observed means appear as white-filled diamonds. Corresponding odds ratios for fixed effects and slope estimates of the lag functions are reported in Tables 3 and 4.

Experiment 2

In Experiment 1, after studying pairs with one or two responses, participants attempted to recall only the first response. Retroactive facilitation was obtained for A-B, A-C items, and this was associated with detecting and recollecting changes. If integrative encoding contributed to facilitation, an appropriate testing methodology will show superior recall of both responses and memory for their order. However, two limitations make it premature to be fully convinced of this interpretation. First, the test trials required recall of first responses while avoiding intrusions of second responses. Participants may have thus focused less on encoding the second responses, thus undermining integrative encoding. If so, when changes were detected and recollected, retroactive facilitation would have reflected retrieval of first responses during the presentation of A-C pairs, while ignoring the second responses and instead rehearsing the first responses. This is similar to rehearsal borrowing in item-method directed forgetting (Bjork, 1970) and would result in superior first-response recall at the expense of second-response recall. This is plausible because the Experiment 1 method did not incentivize encoding of the second responses when detecting changes beyond telling participants that they would later report if they remembered changes. Second, if the retroactive facilitation arose partly through the ordered integration, then the Experiment 1 method precluded participants from being able to show order memory.

Experiment 2 addressed these limitations by indexing the accessibility of both responses using modified-modified free recall (MMFR; Barnes & Underwood, 1959). MMFR instructs participants to report responses as they come to mind and indicate the order of each response. MMFR indexed response accessibility more purely than the Experiment 1 directed cued recall method because MMFR did not require participants to engage in response set suppression of second responses. Moreover, the requirement to report both responses at test expands the focus of encoding to include the second responses when changes are detected. If retroactive facilitation occurs due to rehearsal borrowing, then recall of first responses should be enhanced when changes are detected and recollected, but at the expense of second-response recall. Conversely, if integration contributes to retroactive facilitation, then recall of both responses, as well as memory for their order, should be enhanced when change is detected and recollected.

Materials and methods

Participants

Experiment 2 matched the sample size from Experiment 1, including usable data from 120 participants (60 participants from each site). Data from eight additional UNCG participants were excluded (six never indicated change detection, and two were observed not engaging in the task). Data from five additional UofE participants were excluded (one left the session early, one fell asleep, one produced first-response recall of 0%, one's data were lost due to program failure, and one was observed

Table 5
Odds ratios from mixed-effects logistic regression models of change recollection on test trials.

	Experiment 1		Experiment 2	
Predictor	*OR (95% CI)*	*p*	*OR (95% CI)*	*p*
Item type	**0.05 (0.04, 0.06)**	**< 0.001**	**0.05 (0.04, 0.07)**	**< 0.001**
Retention interval	**0.83 (0.72, 0.96)**	**= 0.013**	**0.83 (0.73, 0.95)**	**= 0.007**
Lag	**0.77 (0.70, 0.84)**	**< 0.001**	**0.84 (0.77, 0.92)**	**< 0.001**
Lag^2	0.98 (0.86, 1.12)	= 0.746	0.90 (0.80, 1.01)	= 0.077
Item type × Retention interval	1.19 (0.96, 1.49)	= 0.116	1.41 (1.12, 1.78)	= 0.004
Item type × Lag	**1.31 (1.13, 1.51)**	**< 0.001**	**1.17 (1.01, 1.36)**	**= 0.036**
Retention interval × Lag	**1.22 (1.11, 1.33)**	**< 0.001**	**1.24 (1.14, 1.35)**	**< 0.001**
Item type × Lag^2	0.95 (0.77, 1.16)	= 0.596	1.14 (0.92, 1.41)	= 0.224
Retention interval × Lag^2	0.94 (0.83, 1.07)	= 0.380	**0.89 (0.79, 1.00)**	**= 0.046**
Item type × Retention interval × Lag	**0.78 (0.68, 0.90)**	**< 0.001**	**0.76 (0.65, 0.88)**	**< 0.001**
Item type × Retention interval × Lag^2	1.03 (0.84, 1.25)	= 0.794	0.96 (0.78, 1.18)	= 0.689

Note. Predictors for Retention interval and Lag were standardized (mean-centered and scaled to unit variance) prior to model estimation to address model convergence issues. Odds ratios (ORs) reflect the multiplicative change in the odds of first-response recall associated with a one standard deviation increase in the predictor, holding other variables constant. Interaction terms represent combined effects of standardized predictors. For categorical contrasts, A-B, A-C items served as the reference level for Item type. Statistically significant ORs are displayed in bold.

not engaging in the task). We performed a sensitivity analysis of first-response recall regressed on the interaction of lags (quadratic), retention intervals, and item types, as in Experiment 1. The sample provided 80% power to detect a small effect size ($OR = 1.27$); see the Supplementary Material (Section 1) for more details.

Design and materials

The design and materials were the same as in Experiment 1.

Procedure

The study trials were the same as in Experiment 1. On test trials, a cue appeared, and participants typed all responses that appeared with the cue as they came to mind. After participants entered a first response, it was replaced by a prompt asking if the response was the first (by pressing "1"), second (by pressing "2"), or only (by pressing "O") to appear with the cue. They could also pass by pressing the "P" key. When participants responded that the response was the first or second to appear, they were prompted to type the next response that came to mind. After participants made a second response, it was replaced by a prompt asking if the response was the first (by pressing "1") or second (by pressing "2") to appear. They could also pass by pressing the "P" key. Participants had up to 10 s for each recall response and 4 s for each order response before the program automatically advanced to the next trial.

Results

First-response recall

We first examined whether the findings of retroactive facilitation increasing across retention intervals and being optimal at the intermediate lag and longest retention interval would replicate under the less constraining MMFR method. To characterize retroactive effects of memory, we compared first-response recall across item types, lags, and retention intervals with the model: logit[P(recall)] ~ Site × Item type × Retention interval × Lag + Site × Item type × Retention interval × Lag^2 + (1 | Subject) + (1 | Item), which included 21,600 observations ($R^2_M = 0.05$, $R^2_C = 0.33$). Site is included here because the full model improved fit, $\chi^2(12) = 34.63$, $p < .001$ (for Site-specific effects, see Supplementary Material, Section 2.2). Table 1 (right) reports fixed-effect *ORs*; Table 2 (bottom) reports slope parameters and monotonicity of lag functions; Fig. 2 (bottom) shows the model-estimated probabilities. A significant Item type × Retention interval interaction indicated higher recall for A-B, A-C than A-B items that increased across retention intervals. A significant Retention interval × Lag interaction indicated that negative lag functions became shallower across retention intervals. A significant Item type × Lag^2 interaction indicated that lag functions were flat or more negatively curvilinear for A-B, A-C compared to A-B items, with the latter being mostly flat. A significant Retention interval × Lag^2 interaction showed that the curvature of the lag function changed across retention intervals. Quadratic coefficients were positive or near zero at the shortest interval and became smaller, and in some cases negative, at longer intervals, indicating increasing concavity of the lag function with delay. Descriptively, this shift was larger in the A-B, A-C condition than in the A-B condition, but unlike Experiment 1, the interaction with Item type was not significant.

Table 6
Slope parameters and monotonicity of lag functions in change recollection on test trials.

Experiment 1							
Item type	**Retention interval**	***Lag b (95% CI)***	***Lag b^2 (95% CI)***	**Slope min**	**Slope max**	**Vertex (x*)**	**Monotonicity**
A-B, A-C	4	−0.38 (−0.48, −0.28)	0.03 (−0.11, 0.16)	−0.43	−0.31	161.38	Monotone decreasing
A-B, A-C	16	−0.23 (−0.29, −0.16)	−0.06 (−0.15, 0.03)	−0.11	−0.40	−12.15	Monotone decreasing
A-B, A-C	48	−0.08 (−0.16, 0.01)	−0.15 (−0.27, −0.03)	0.22	−0.48	17.85	Non-monotone
A-B	4	0.08 (−0.03, 0.20)	−0.10 (−0.27, 0.06)	0.29	−0.20	30.00	Non-monotone
A-B	16	−0.00 (−0.09, 0.08)	−0.03 (−0.15, 0.09)	0.06	−0.09	21.27	Non-monotone
A-B	48	−0.09 (−0.21, 0.02)	0.04 (−0.12, 0.21)	−0.18	0.03	42.31	Non-monotone
Experiment 2							
Item type	**Retention interval**	***Lag b (95% CI)***	***Lag b^2 (95% CI)***	**Slope min**	**Slope max**	**Vertex (x*)**	**Monotonicity**
A-B, A-C	4	−0.37 (−0.46, −0.28)	0.04 (−0.09, 0.17)	−0.45	−0.26	105.80	Monotone decreasing
A-B, A-C	16	−0.17 (−0.23, −0.10)	−0.10 (−0.19, −0.01)	0.04	−0.45	7.71	Non-monotone
A-B, A-C	48	0.04 (−0.04, 0.12)	−0.25 (−0.37, −0.13)	0.53	−0.63	24.12	Non-monotone
A-B	4	0.08 (−0.03, 0.19)	0.15 (−0.01, 0.30)	−0.21	0.48	17.54	Non-monotone
A-B	16	0.00 (−0.07, 0.08)	0.01 (−0.10, 0.11)	−0.01	0.02	15.43	Non-monotone
A-B	48	−0.07 (−0.18, 0.04)	−0.14 (−0.28, 0.01)	0.20	−0.44	17.71	Non-monotone

Note. b = linear slope in log-odds; b^2 = quadratic slope in log-odds. Slope min and slope max represent the estimated slopes at the boundaries of the observed lag range. Functions were classified as monotone decreasing if both were negative. If slopes differed in sign, monotonicity depended on the vertex (x*): when the turning point fell within the observed lag range, the function was classified as non-monotone; when it fell outside the range, the function was classified as monotone.

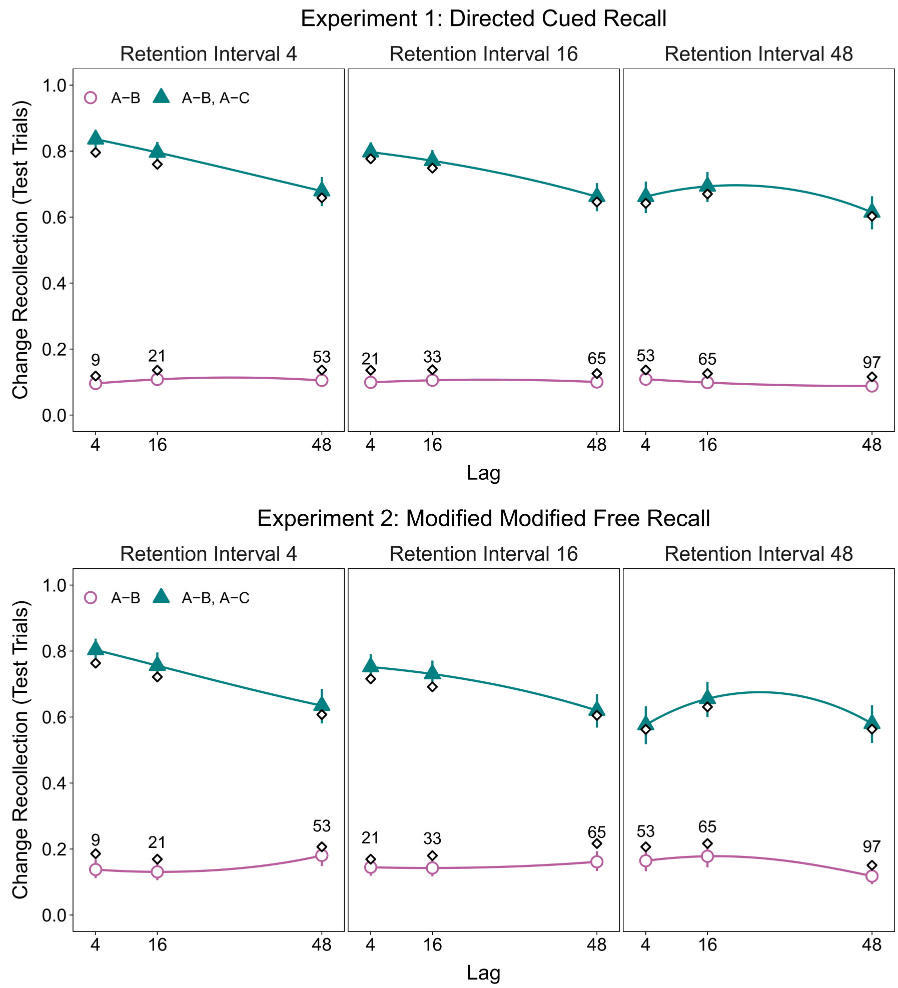

Fig. 4. Probabilities of change recollection on test trials.
Note. Estimates are model-derived from a mixed-effects analysis; error bars represent 95% confidence intervals. Observed means appear as white-filled diamonds. Numeric values beneath the A-B points indicate total retention intervals between study and test. Corresponding odds ratios for fixed effects and slope estimates of the lag functions are reported in Tables 5 and 6.

Change detection on study trials

Because the study trial procedure was the same as in Experiment 1, we expected change detection to again decrease across increasing lags. The differences between accurate and inaccurate responses were compared across lags and retention intervals with the model: logit[P(change detection)] ~ Item type × Retention interval × Lag + Item type × Retention interval × Lag^2 + (1 | Subject) + (1 | Item), which included 21,600 observations ($R^2_M = 0.64$, $R^2_C = 0.73$). Table 3 (right) reports fixed-effect *ORs*; Table 4 (bottom) summarizes the shapes of the model-estimated lag functions descriptively; Fig. 3 (bottom) shows the model-estimated probabilities. A significant Item type × Lag interaction showed that change detection accuracy declined linearly across lags, replicating the pattern observed in Experiment 1. This pattern suggests that memory for A-B pairs, which supports change detection, decreased as lag increased.

Change recollection on test trials

The change of test in Experiment 2 necessitated a different operational definition for change recollection. Here, change recollection was assumed to occur when participants indicated that that their first retrieval attempt was the first or second response to appear with cue, and hence, not the only response to appear. We examined whether, under this definition, the change recollection lag functions would replicate the pattern from Experiment 1. The differences between accurate and inaccurate responses were compared across lags and retention intervals with the model: logit[P(change recollection)] ~ Site × Item type × Retention interval × Lag + Site × Item type × Retention interval × Lag^2 + (1 | Subject) + (1 | Item), which included 21,600 observations ($R^2_M = 0.31$, $R^2_C = 0.48$). Site is included here because the full model improved fit, $\chi^2(12) = 82.13$, $p < .001$ (for Site-specific effects, see Supplementary Materials, Section 2.3). Table 5 (right) reports fixed-effect *ORs*; Table 6 (bottom) summarizes the shapes of the model-estimated lag functions descriptively; Fig. 4 (bottom) shows the model-estimated probabilities.

A significant Item Type × Retention Interval × Lag interaction showed that the negative linearity of the lag functions for A-B, A-C items became shallower across retention intervals and then flat at the longest interval. Conversely, the lag functions for A-B items were essentially flat. This led to increasingly smaller significant differences between accurate and inaccurate change recollection responses across intervals 4, 16, and 48 at lag 4 (*ORs*: 25.64 > 14.96 > 8.73) and lag 16 (*ORs*: 16.35 > 12.63 > 9.76), but not lag 48 (*ORs*: 10.43 ≈ 10.66 ≈ 10.90), all *ps* < 0.001. A significant Retention interval × Lag^2 interaction indicated that the curvature of the lag functions changed across retention intervals. The model-estimated lag functions (Table 6) showed little curvature at retention interval 4 but increasingly negative curvature at intervals 16 and 48. At the longest interval, the functions were descriptively consistent with a non-monotonic pattern, unlike Experiment 1.

Table 7
Wald's χ^2 Tests from Mixed-Effects Logistic Regression of First-Response Recall by Change Classification.

	Experiment 1		Experiment 2	
Predictor	$\chi^2(df)$	p	$\chi^2(df)$	p
Classification	**1491.81 (2)**	**< 0.001**	**1961.31 (2)**	**< 0.001**
Retention interval	3.72 (1)	= 0.054	**6.81 (1)**	**= 0.009**
Lag	**29.28 (1)**	**< 0.001**	**58.50 (1)**	**< 0.001**
Lag^2	0.35 (1)	= 0.553	0.03 (1)	= 0.858
Classification × Retention interval	0.09 (2)	= 0.955	4.87 (2)	= 0.088
Classification × Lag	**8.32 (2)**	**= 0.016**	**10.89 (2)**	**= 0.004**
Retention interval × Lag	**7.06 (1)**	**= 0.008**	**4.87 (1)**	**= 0.027**
Classification × Lag^2	**11.23 (2)**	**= 0.004**	4.90 (2)	= 0.086
Retention interval × Lag^2	**6.01 (1)**	**= 0.014**	0.12 (1)	= 0.733
Classification × Retention interval × Lag	1.13 (2)	= 0.568	1.34 (2)	= 0.512
Classification × Retention interval × Lag^2	0.20 (2)	= 0.906	0.36 (2)	= 0.836

Note. Predictors for Retention interval and Lag were standardized (mean-centered and scaled to unit variance) prior to model estimation to address model convergence issues. Because the categorical predictor change classification (Classification) had more than two levels, Wald's χ^2 tests are reported instead of odds ratios. The omnibus test provides a more interpretable summary of the overall effect than individual contrasts from the model output. The χ^2 statistic reflects how much the inclusion of the predictor improves model fit, providing an index of the overall contribution of that predictor rather than focusing only on pairwise contrasts. Interaction terms represent combined effects of standardized predictors. For categorical contrasts, Change detected and recollected served as the reference level for change classification. The statistically significant χ^2 statistic are displayed in bold.

First-response recall conditionalized on change classifications

To determine if first-response recall scales with change awareness as in Experiment 1, we conditionalized such recall for A-B, A-C items on the three change classifications. If so, recall should be best for detected and recollected changes, and it should decline with diminishing change awareness. The differences in recall associated with change awareness were compared across lags and retention intervals using the following model: logit[P(recall)] ~ Site × Classification × Retention interval × Lag + Site × Classification × Retention interval × Lag^2 + (1 | Subject) + (1 | Item), which included 10,800 observations ($R^2_M = 0.34$, $R^2_C = 0.48$). Site is included here because the full model improved fit, $\chi^2(18) = 39.78$, $p = .002$ (for Site-specific effects, see Supplementary Materials, Section 2.4). Table 7 (right) reports fixed-effect *ORs*; Table 8 (bottom) summarizes the shapes of the model-estimated lag functions descriptively; Fig. 5 (bottom) shows the model-estimated probabilities. A significant effect of Classification indicated that first-response recall was significantly higher when detected changes were recollected than when they were not recollected, $OR = 9.24$, $p < .001$, and it was also significantly higher for both instances when changes were detected than when changes were not detected, smallest $OR = 3.06$, $p < .001$. A significant Classification × Lag interaction showed steeper negative linear slopes when changes were not detected than in the other classifications. A significant Retention Interval × Lag interaction also showed that negative linear slopes became shallower across intervals. The two-way interactions involving the quadratic lag predictor (Classification × Lag^2 and Retention Interval × Lag^2) were not significant, in contrast to Experiment 1 where both were significant.

Second-response recall conditionalized on change classification

As described in the Introduction to Experiment 2, the MMFR task allowed us to assess second-response recall, which provides a stronger test of memory-for-change frameworks. Based on that view, we expected that detecting and recollecting change would also be associated with enhanced recall of the second responses in A-B, A-C pairs. We tested this hypothesis by comparing second-response recall across levels of change awareness with the model: logit[P(recall)] ~ Classification × Retention Interval + Classification × Retention Interval^2 (1 | Subject) + (1 | Item), which included 10,800 observations ($R^2_M = 0.23$, $R^2_C = 0.42$). Unlike the previous analysis, we did not model a Lag fixed effect because second-responses appeared after the lag manipulation. Instead, a quadratic term for retention interval term was included for exploratory purposes to capture any potential curvilinearity in forgetting across intervals.

Fig. 6 shows the model-estimated probabilities of second-response

Table 8
Slope parameters and monotonicity of lag functions in first-response recall by change classification.

Classification	Retention interval	*Lag b (95% CI)*	*Lag b^2 (95% CI)*	Slope min	Slope max	Vertex (x*)	Monotonicity
Experiment 1							
Classification	**Retention interval**	***Lag b (95% CI)***	***Lag b^2 (95% CI)***	**Slope min**	**Slope max**	**Vertex (x*)**	**Monotonicity**
Detected & recollected	4	−0.23 (−0.34, −0.12)	0.05 (−0.10, 0.21)	−0.34	−0.08	61.20	Monotone decreasing
Detected & recollected	16	−0.16 (−0.24, −0.08)	−0.05 (−0.16, 0.06)	−0.06	−0.29	−8.18	Monotone decreasing
Detected & recollected	48	−0.09 (−0.21, 0.02)	−0.15 (−0.32, 0.01)	0.21	−0.51	16.90	Non-monotone
Detected & not recollected	4	−0.23 (−0.54, 0.07)	−0.16 (−0.58, 0.26)	0.09	−0.67	8.99	Non-monotone
Detected & not recollected	16	−0.07 (−0.27, 0.12)	−0.30 (−0.57, −0.03)	0.53	−0.89	20.41	Non-monotone
Detected & not recollected	48	0.09 (−0.12, 0.30)	−0.44 (−0.75, −0.13)	0.97	−1.11	24.57	Non-monotone
Not detected	4	−0.50 (−0.72, −0.29)	0.43 (0.11, 0.75)	−1.37	0.67	33.53	Non-monotone
Not detected	16	−0.38 (−0.54, −0.22)	0.27 (0.05, 0.50)	−0.93	0.37	35.45	Non-monotone
Not detected	48	−0.25 (−0.48, −0.02)	0.12 (−0.21, 0.44)	−0.49	0.07	42.39	Non-monotone
Experiment 2							
Classification	**Retention interval**	***Lag b (95% CI)***	***Lag b^2 (95% CI)***	**Slope min**	**Slope max**	**Vertex (x*)**	**Monotonicity**
Detected & recollected	4	−0.33 (−0.46, −0.20)	−0.01 (−0.19, 0.17)	−0.31	−0.35	−353.58	Monotone decreasing
Detected & recollected	16	−0.27 (−0.36, −0.18)	−0.06 (−0.19, 0.07)	−0.16	−0.43	−21.76	Monotone decreasing
Detected & recollected	48	−0.22 (−0.35, −0.08)	−0.11 (−0.30, 0.08)	−0.00	−0.50	3.63	Monotone decreasing
Detected & not recollected	4	−0.28 (−0.53, −0.03)	−0.11 (−0.46, 0.24)	−0.06	−0.57	−1.23	Monotone decreasing
Detected & not recollected	16	−0.13 (−0.29, 0.03)	−0.09 (−0.31, 0.14)	0.04	−0.36	8.49	Non-monotone
Detected & not recollected	48	0.02 (−0.17, 0.20)	−0.06 (−0.33, 0.20)	0.14	−0.16	24.86	Non-monotone
Not detected	4	−0.64 (−0.87, −0.41)	0.23 (−0.10, 0.56)	−1.10	−0.01	48.38	Monotone decreasing
Not detected	16	−0.50 (−0.67, −0.33)	0.23 (−0.01, 0.48)	−0.96	0.14	42.52	Non-monotone
Not detected	48	−0.36 (−0.61, −0.10)	0.24 (−0.14, 0.61)	−0.83	0.28	36.79	Non-monotone

Note. b = linear slope in log-odds; b^2 = quadratic slope in log-odds. Slope min and slope max represent the estimated slopes at the boundaries of the observed lag range. Functions were classified as monotone decreasing if both were negative. If slopes differed in sign, monotonicity depended on the vertex (x*): when the turning point fell within the observed lag range, the function was classified as non-monotone; when it fell outside the range, the function was classified as monotone.

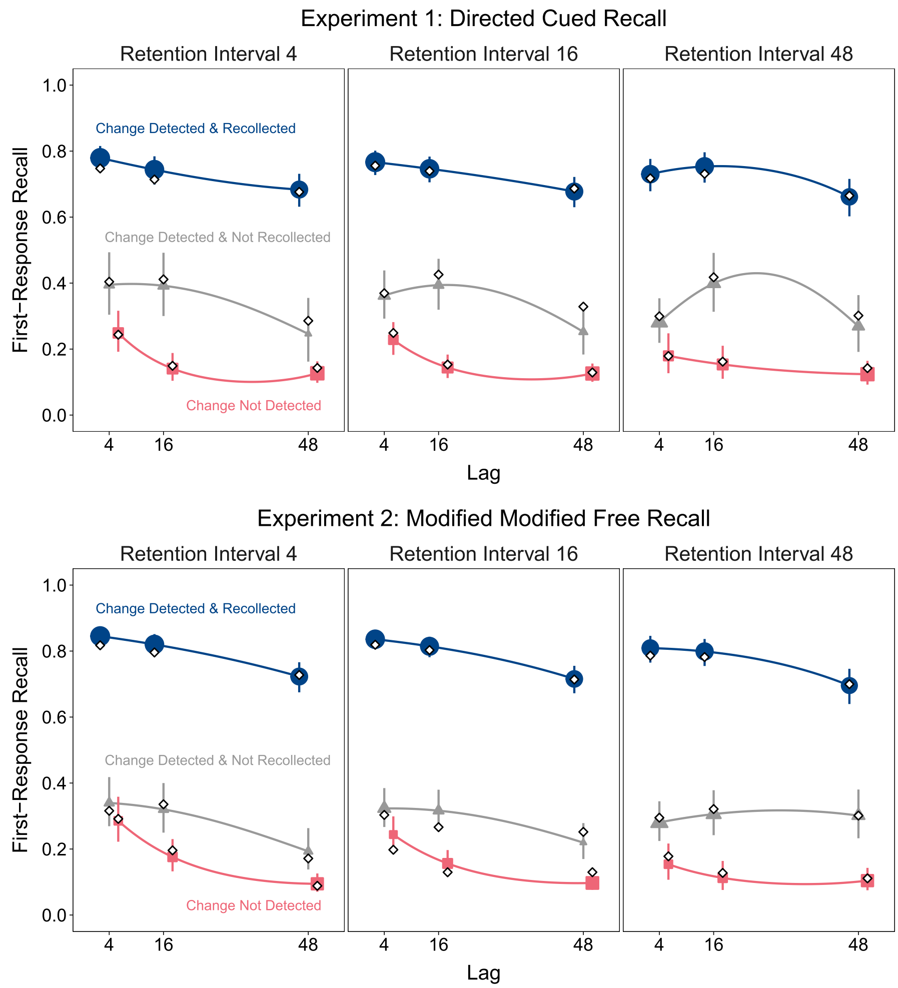

Fig. 5. Probabilities of first-response recall by change classification.
Note. Estimates are model-derived from a mixed-effects analysis; error bars represent 95% confidence intervals. Point sizes indicate observation counts across cells. Observed means appear as white-filled diamonds. Corresponding Wald χ^2 tests and slope estimates of the lag functions are reported in Tables 7 and 8.

recall conditionalized on change awareness across retention intervals. A significant effect of Classification, $\chi^2(2) = 1238.01$, $p < .001$, indicated that second-response recall was significantly higher when changes were detected and recollected than not detected at all, $OR = 3.64$, $p < .001$, and for both of these classifications than when changes were detected and not recollected, smallest $OR = 2.99$, $p < .001$. There was also a significant effect of Retention Interval, $\chi^2(1) = 103.85$, $p < .001$, showing a negative linear slope ($b = -0.36$ [-0.44, -0.28]), and a significant effect of Retention Interval^2, $\chi^2(1) = 9.10$, $p = .003$, showing a negative quadratic slope ($b^2 = -0.19$ [-0.30, -0.07]). No other effects were significant, largest $\chi^2(2) = 1.94$, $p = .38$. These results support the memory-for-change prediction in showing that detecting and recollecting changes were associated with enhanced second-response recall. Importantly, the finding that second-response recall was lowest for detected changes that were not recollected is consistent with findings showing that proactive interference emerges in such instances (Wahlheim et al., 2021).

Order judgment accuracy conditionalized on change classification

Finally, memory-for-change frameworks assume that change detection and recollection are associated with interference reduction because the processes involved engender encoding of order information. This leads to the prediction here that order memory judgments should be most accurate when changes are detected and recollected. We tested this by conditionalizing order memory accuracy across first and second responses for A-B, A-C items on change awareness with the model: logit [P(order accuracy)] ~ Classification + (1 | Subject) + (1 | Item), which included 10,800 observations ($R^2_M = 0.74$, $R^2_C = 0.81$). Consistent with our predictions, order judgments were significantly more accurate when changes were detected and recollected (0.86 [0.83, 0.88]) than when changes were detected and not recollected (< 0.01), $OR = 14{,}035.67$, $p < .001$, and when changes were not detected (0.08 [0.06, 0.10]), $OR = 69.61$, $p < .001$.

Discussion

Under an MMFR test, first-response recall again showed retroactive interference and retroactive facilitation. Consistent with interference accounts, recall for single A-B items decreased with the number of events intervening between study and test, and recall was lower for A-B, A-C pairs than for retention-interval-matched A-B pairs when study-phase retrieval did not occur. There was also clear evidence for retroactive facilitation. First, recall of first responses was higher overall for A-B, A-C items than for retention-matched A-B controls, and this facilitation increased across retention intervals. Conditional analyses further showed that recall scaled with change awareness, being highest when changes were detected and recollected, intermediate when detected but not recollected, and lowest when changes were not detected. This pattern replicated the graded association observed in Experiment 1.

Experiment 2 also allowed us to examine recall of the second (C)

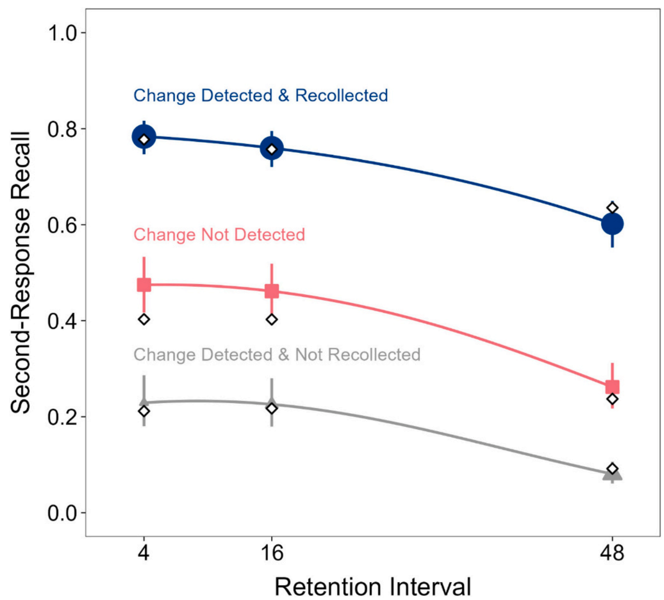

Fig. 6. Probabilities of second-response recall by change classification: experiment 2.

Note. Estimates are model-derived from a mixed-effects analysis; error bars represent 95% confidence intervals. Point sizes indicate observation counts across cells. Observed means appear as white-filled diamonds.

responses for A-B, A-C trials. Critically, second-response recall was highest when changes were detected and recollected. This pattern is inconsistent with a rehearsal-borrowing account, which predicts that detecting change would increase the accessibility of A-B at the expense of encoding A-C. Instead, the enhanced recall of both responses, along with more accurate order judgments, supports the view that facilitation partly reflects encoding of relations between responses, consistent with integrative encoding of an ordered representation (e.g., "A was paired with B, but now with C"). This pattern is also inconsistent with accounts based solely on competition between independent traces, which predict that stronger A-B representations should reduce access to A-C responses. Moreover, these findings address concerns that conditional analyses reflect only baseline memory strength (i.e., that memory is better when it is better). If stronger A-B memories simply increased both detection and later recall, they should also increase competition and reduce A-C recall. Instead, improved recall of both responses following detection and recollection indicates that these trials reflect encoding processes that enhance access to both associations. These findings therefore constrain possible mechanisms without uniquely identifying them.

Finally, as anticipated from Experiment 1, the non-monotonic lag pattern at the longest retention interval was not statistically reliable. One possibility is that this difference reflects how retrieval was directed at test. In Experiment 1, participants were required to retrieve the first response, which may have increased reliance on reinstating contextual information to resolve competition between responses, thereby amplifying lag-dependent effects. In contrast, the MMFR procedure in Experiment 2 permitted retrieval of multiple responses, potentially reducing reliance on context-based selection processes and attenuating sensitivity to non-monotonic lag patterns. In addition, sequential output in MMFR may further reduce detectability if reporting A-C responses first introduces output interference that could disrupt access to A-B responses. We therefore do not interpret these effects in terms of benefits of more difficult retrieval.

General discussion

Across two experiments, we examined retroactive effects of memory across lags and retention intervals in a continuous paired associate A-B, A-C paradigm using cued recall (Experiment 1) and MMFR (Experiment 2) tests. We were interested in examining classic and more recent claims regarding retroactive interference and retroactive facilitation and how each process would vary across lags and retention intervals. Our experiments found evidence consistent with both retroactive interference and retroactive facilitation.

Retroactive interference

Regarding retroactive interference, there was clear evidence of a decline in cued recall with increasing numbers of intervening items between study and test. In addition, the cued recall of the first response (A-B) from related A-B, A-C pairs was reduced relative to retention-interval matched single A-B pairs on trials in which the change of response was not detected. We also observed reductions in participants' ability to detect changed responses as the number of intervening events increased between the first and second responses.

These retroactive interference effects can be explained in the A-B, A-C paradigm by many classic and more recent accounts that assume that the two related study pairs are encoded as independent traces. Within such frameworks, memory performance depends on the relative accessibility of those traces at test. Accessibility may be determined by differences in trace strength that evolve over time (Baddeley, 1976) or by the match between the study context and the retrieval context, which also changes with test delay (Mensink & Raaijmakers, 1988). Similar ideas also underlie more recent approaches emphasizing temporal distinctiveness and retrieved-context dynamics (Brown et al., 2007; Glenberg & Swanson, 1986; Howard & Kahana, 2002). Across these accounts, increasing the retention interval should reduce the accessibility of earlier A-B associations, either because memory accessibility decays or because the match between study and test contexts is reduced. Consistent with these expectations, recall of the single A-B items in the present experiments declined across retention intervals, replicating well-established recency patterns in cued recall (Glenberg, 1976).

Temporal distinctiveness accounts can also explain the poor first-response recall observed on A-B, A-C trials when change was not detected during study. When change is not detected, the A-B and A-C events can be encoded as independent traces, allowing the later A-C association to compete with retrieval of the earlier A-B association. These trials likely include cases in which the A-B association was relatively inaccessible at the time of A-C study, making it less likely to be retrieved and therefore more susceptible to interference from the interpolated learning. Under these conditions, recall of the earlier response should reflect competition between independently encoded traces rather than the benefits of study-phase retrieval or integrative encoding. However, contrary to the sufficiency of these accounts, our experiments were dominated by clear evidence of retroactive facilitation.

Retroactive facilitation

In both experiments, there was clear improvement overall in the recall of the first response from related A-B, A-C pairs relative to retention-interval matched single A-B pairs on trials when the data were unconditionalized. This overall improvement was driven by the heightened recall of the first response from related A-B, A-C pairs when change was detected and later recollected. This pattern regarding first responses was observed when participants were required to perform cued recall of the first response at test in Experiment 1, and cued recall of both responses in the MMFR testing of Experiment 2.

Importantly, the MMFR findings in Experiment 2 showed that facilitated first-response recall for A-B, A-C trials in which changes were detected and recollected was accompanied by heightened second-response recall. This pattern supports accounts proposing that detecting change during study allows the second response to cue retrieval of the earlier response (study-phase retrieval), promoting encoding of an integrated representation. Such representations are assumed to be more durable and to include qualitative information about the change (e.g., "A was paired with B but now with C"). Consistent with this view, trials in which change was detected and recollected showed: enhanced recall of A-B and A-C responses, greater accuracy in ordering responses as first and second, and reduced forgetting of first responses across retention intervals relative to control pairs and trials without detected or recollected change.

The heightened recall of second responses when change in the first response is detected also helps assuage concerns regarding the circularity of reasoning regarding our conditional analyses. If we limit our interpretation to the recall of first responses, there is a danger that our findings could be reduced to merely demonstrating evidence that word pairs that are more strongly encoded (those that are most likely for change to be detected) are associated with better memory performance (heightened cued recall of first responses). However, extending our interpretation to the recall of second responses helps circumvent this concern: those word pairs that are more strongly encoded (those that are most likely for change to be detected) are also associated with better memory performance for their competitors (heightened cued recall of second responses). If interference theory was solely operating, then stronger A-B traces (as identified through change detection) should *reduce* accessibility to A-C when change was detected relative to when not. By contrast, our observation that stronger A-B traces are associated with more memorable A-C traces strongly suggests that our results are more than just circular reasoning, and instead provide the evidence for a more meaningful interpretation in memory.

Consistent with accounts emphasizing study-phase retrieval and integrative encoding, retroactive facilitation has been observed in situations that increase the likelihood that earlier associations are retrieved during interpolated learning, including semantic relatedness between responses (Antony et al., 2022; Barnes & Underwood, 1959), extended study of interpolated pairs (Garlitch & Wahlheim, 2020; Negley et al., 2018), retrieval practice of the initial associations (Bäuml & Trißl, 2022; Wahlheim et al., 2023), and encoding strategies that promote integration of the two events (Burton et al., 2017). The present findings align with this work by showing that facilitation was most evident when change was detected and later recollected. Moreover, the increase in facilitation across retention intervals replicates findings that facilitation becomes more apparent at longer test delays (Antony et al., 2022; Cox et al., 2026), which has been interpreted as reflecting the durability of integrated representations of A-B and A-C pairs.

Elaborating on this explanation, we invoke the memory-for-change account (Wahlheim & Jacoby, 2013), which proposes that when A-B memories are retrieved and compared with A-C pairs during study, the resulting representation can include information about the change itself. Such information has been described as a memory attribute (Underwood, 1969); here, we refer to it as change awareness. This attribute provides diagnostic information about response order that can guide reporting decisions when retrieval is directed toward a specific temporal context and inform judgments about order when responses are freely reported. By this account, integrated memories are durable not only because they are quantitatively stronger, but also because they include additional qualitative information that constrains reporting decisions. This perspective aligns with remindings accounts of order memory, which propose that study-phase retrieval allows later events to be encoded as reminders of earlier ones, enabling inference of their relative order at test (Tzeng & Cotton, 1980; Winograd & Soloway, 1985). Consistent with this view, Experiment 2 showed that when change awareness persisted from study to test, both second-response recall and order judgments were substantially more accurate.

A further development of the memory-for-change perspective may help explain how the lag between A-B and A-C pairs relates to facilitation across retention intervals. One possibility is to incorporate retroactive interference ideas from dual-mechanism accounts of lag-repetition effects. Non-monotonic lag repetition functions have been explained as arising from interactions between contextual variability and study-phase retrieval (Benjamin & Tullis, 2010; Glenberg, 1976; Greene, 1989; see also Raaijmakers, 2003). According to this view, increasing lag reduces the likelihood that the later presentation will retrieve the earlier one, but when retrieval does occur it tends to be more difficult and therefore more beneficial for memory. As a result, retrieval benefits are often greatest at intermediate lags, where retrieval remains possible but requires greater effort. Short lags provide little retrieval challenge, whereas very long lags reduce the probability that retrieval will occur. Because retrieval success also depends on the match between study and test contexts, the lag that produces the strongest benefits may vary with the retention interval. The present experiments provided mixed evidence for such lag-dependent curvature, suggesting that more systematic investigation will be necessary to clarify when these patterns emerge and how they operate in paradigms involving temporally distributed feature changes.

A modified memory-for-change account could incorporate ideas from dual-mechanism lag-repetition accounts when applied to paradigms involving feature changes, such as the A-B, A-C paradigm. Such an extension could help explain how lag functions arise in both first-response recall and change recollection, which index different aspects of the accessibility of memories that include change information. Across the present experiments, these measures showed partially related patterns but differed in the statistical support for non-monotonic lag functions. This pattern suggests that memories preserving information about change contributed to the observed lag functions, while also indicating that the expression of these functions depends on how retrieval is directed and how change recollection is assessed. From a dual-mechanism perspective, intermediate lags may sometimes produce integrated memories in which the A-B response, the A-C response, and information about their relationship are encoded together during study-phase retrieval and comparison. When the retention interval allows reinstatement of contexts associated with the earlier event, awareness that the cue was associated with a change may provide a retrieval route to memories that include this change information by matching the operations engaged during study-phase retrieval (see Tulving & Thomson, 1973).

Support for the proposal that lag and retention interval determine how change attributes are retrieved comes from experiments examining proactive memory effects while varying the lag between A-B and A-C pairs (Jacoby et al., 2015). In that study, A-B and A-C pairs appeared across two lists (longer lags) or within the same list (shorter lags). Participants detected fewer changes during A-C study when A-B pairs occurred at longer lags, indicating reduced study-phase retrieval under those conditions. However, change recollection at test showed the opposite pattern: Participants recollected more changes for pairs separated by longer lags. From a dual-mechanism perspective, this pattern suggests that longer lags may produce more distinctive contextual encodings of change information that are better reinstated at test, increasing the likelihood that change attributes are retrieved when the retention interval aligns with those contexts. Further empirical work examining broader combinations of lag and retention interval, together with formal modeling of contextual mechanisms, will be necessary to test this proposal.

One promising direction for future work would be to adapt variants of retrieved context models (Howard & Kahana, 2002) to account for the present findings. Current implementations primarily emphasize retrieval processes at test, where the context associated with a recalled item cues other studied items that share related contextual states. During study, by contrast, retrieved context is typically assumed to update the evolving study context rather than to reinstate and re-encode earlier items in association with later ones (Halpern et al., 2025). However, in the present paradigm, encountering an A-C pair may cue retrieval of the earlier A-B association (Ward & Tan, 2023), allowing the two events to be encoded in relation to one another along with information about how they differ. If such retrievals occur, the resulting representations could include contextual information indicating that the cue was associated with a change. Spaced-repetition variants of the retrieved-context model (Lohnas, 2025; Siegel & Kahana, 2014) provide a potential starting point because they treat repeated cues as opportunities for study-phase retrieval. Extending those models to allow retrieved A-B pairs to be encoded with A-C pairs, along with their temporal contexts and the differences between them, could provide a formal framework for examining how study-phase retrieval and integrative encoding contribute to retroactive facilitation across lags and retention intervals and to the various associations between recall and change awareness observed here.

Limitations and opportunities

A central inference in this study is that awareness of change during A-C study often reflected retrieval of the earlier A-B response. This inference is plausible but indirect, as some detected changes may have been supported by cue familiarity or response novelty without retrieval of the earlier response itself. Prior work indicates that change detection is often, but not always, accompanied by recall of the earlier response, whereas change recollection depends more strongly on study-phase retrieval of A-B pairs (Wahlheim et al., 2019). Consequently, we treat detected-and-recollected changes as more likely than detected-and-not-recollected changes to index successful study-phase retrieval. However, detected changes that were not later recollected may also include instances in which retrieval occurred but was weak or inaccurate, providing sufficient information to signal change without supporting durable encoding of the earlier response or their relation. Because participants were not asked to report study-phase retrievals during study, these interpretations remain inferential rather than directly observed.

We examined a range of lags and retention intervals that could be accommodated within a single session, which was sufficient to reveal theoretically relevant shifts in lag functions at the longest retention interval. However, this design does not determine where non-monotonic lag functions emerge or how they evolve across a broader range of lag × retention interval combinations, which may partly explain the inconsistent evidence for such patterns across experiments and measures. Because non-monotonic spacing functions depend on the relation between lag and retention interval, identifying conditions that maximize retroactive facilitation will require sampling a wider parameter space (Cepeda et al., 2008). Notably, the observed lag patterns resemble those reported in the lag-repetition literature, raising the possibility that related mechanisms contribute across paradigms (Jacoby et al., 2015; Wahlheim et al., 2014). However, the present design did not allow direct comparison of repetition (A-B, A-B) and change (A-B, A-C) conditions, which may differ in how study-phase retrieval affects memory. Whereas retrieval should benefit memory for the original response in both cases, only changed events introduce potential tradeoffs between facilitation via integration and interference via response competition. Future work directly comparing repetitions and changes within a common framework will be necessary to clarify the extent to which shared mechanisms underlie these effects.

A related limitation concerns whether the present findings can adjudicate between divergent theoretical accounts of change detection and recollection. We have interpreted the present results within a memory-for-change framework, in which detecting and recollecting a change indexes study-phase retrieval and integrative encoding. However, trace-competition and context-based accounts of retroactive memory effects (e.g., Baddeley, 1976; Mensink & Raaijmakers, 1988) did not include change detection and recollection as theoretical constructs. Consequently, there is not an alternative account of change detection and recollection against which a memory-for-change interpretation could be directly tested. The present findings are therefore best understood as showing that change awareness is a useful empirical marker of conditions that produce retroactive facilitation. Despite this limitation, the value of this account is evident in the finding that detecting and recollecting a change was associated with enhanced recall of both the first and second responses together with more accurate memory for their order. This is not easily explained by accounts in which similar events are represented as independent traces. Importantly, the memory-for-change account may help explain the variability observed across lags and retention intervals. The benefits of retrieval practice and the quality of integrative encoding likely depend on how the earlier association is retrieved during A-C study. Therefore, the representational quality should vary with lag and retention interval, producing variability in whether changes are recollected and the degree of associated memory benefits.

Conclusion

The present study characterized retroactive facilitation across lags and retention intervals in a continuous A-B, A-C paradigm and examined how these effects relate to awareness of change during study and test. Retroactive facilitation increased across retention intervals and was associated with detection and recollection of changes, consistent with the view that when A-C pairs cue retrieval of earlier A-B associations, retrieval and comparison processes can improve memory for earlier responses and sometimes produce integrated representations. Critically, these benefits extended to both responses and their order, indicating that facilitation reflects relational encoding rather than baseline memory strength. Evidence regarding the shape of lag functions was less consistent: Experiment 1 showed a non-monotonic pattern at the longest retention interval, but this did not replicate in Experiment 2. Accordingly, the present results highlight the role of study-phase retrieval and integrative encoding indexed by change awareness, while suggesting that detectability of lag-function curvature depends on how retrieval is directed at test and on the range of lag and retention interval combinations examined. This variability highlights the need for future work that systematically samples lag and retention interval combinations and develops formal models explaining how study-phase retrieval and contextual dynamics determine when later learning impairs or enhances memory for earlier associations.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the author(s) used ChatGPT (OpenAI, GPT-5, 2025) to assist in developing and refining analysis and visualization code, and to suggest stylistic edits to improve clarity and concision. The author(s) also used Claude Sonnet 5 and Claude Opus 4.8 (Anthropic, 2026) to assist in writing and debugging R code for statistical figures and to help phrase and organize sections of the manuscript based on content, analyses, and decisions provided by the author(s). After using these tools/services, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Author note

The materials, data, and analysis scripts are available on the Open Science Framework: https://osf.io/cvmqx/. We express appreciation to Sydney T. Smith for her assistance with experimental design, programming, and data collection for Experiment 1.

CRediT authorship contribution statement

Christopher N. Wahlheim: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Geoff Ward:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Samuel W. Cox:** Writing – review & editing, Visualization, Validation, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jml.2026.104796.

Data availability

I have shared the link to my data/code.

References

Antony, J. W., Romero, A., Vierra, A. H., Luenser, R. S., Hawkins, R. D., & Bennion, K. A. (2022). Semantic relatedness retroactively boosts memory and promotes memory interdependence across episodes. *eLife, 11*, Article e72519. https://doi.org/10.7554/eLife.72519

Baddeley, A. D. (1976). *The psychology of memory* (1st ed.). Basic Books.

Barnes, J. M., & Underwood, B. J. (1959). "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology, 58*(2), 97–105. https://doi.org/10.1037/h0047507

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bäuml, K.-H. T., & Trißl, L. (2022). Selective memory retrieval can revive forgotten memories. *Proceedings of the National Academy of Sciences, 119*(8), Article e2114377119. https://doi.org/10.1073/pnas.2114377119

Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective? *Cognitive Psychology, 61*(3), 228–247. https://doi.org/10.1016/j.cogpsych.2010.05.004

Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior, 9*(3), 255–268. https://doi.org/10.1016/S0022-5371(70)80059-7

Briggs, G. E. (1954). Acquisition, extinction, and recovery functions in retroactive inhibition. *Journal of Experimental Psychology, 47*(5), 285–293. https://doi.org/10.1037/h0060251

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114*(3), 539–576. https://doi.org/10.1037/0033-295X.114.3.539

Bruce, D., & Weaver, G. E. (1973). Retroactive facilitation in short-term retention of minimally learned paired associates. *Journal of Experimental Psychology, 100*(1), 9–17.

Burton, R. L., Lek, I., & Caplan, J. B. (2017). Associative independence revisited: Competition between conflicting associations can be resolved or even reversed in one trial. *Quarterly Journal of Experimental Psychology, 70*(4), 832–857. https://doi.org/10.1080/17470218.2016.1171886

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science, 19*(11), 1095–1102. https://doi.org/10.1111/j.1467-9280.2008.02209.x

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics - Simulation and Computation, 39*(4), 860–864. https://doi.org/10.1080/03610911003650383

Cox, S. W., Kelley, C. M., Antony, J. W., & Wahlheim, C. N. (2026). *Test delay and change awareness moderate retroactive and proactive memory effects*. . Unpublished.

E-Prime. (2012). Psychology software tools (Version 2.0) [Computer software]. https://www.pstnet.com.

Fox, J., & Weisberg, S. (2019). *An {R} companion to applied regression, Third Edition*. Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Garlitch, S. M., & Wahlheim, C. N. (2020). The role of reminding in retroactive effects of memory for older and younger adults. *Psychology and Aging, 35*(5), 697–709. https://doi.org/10.1037/pag0000427

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior, 15*(1), 1–16. https://doi.org/10.1016/S0022-5371(76)90002-5

Glenberg, A. M., & Swanson, N. G. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(1), 3–15.

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(3), 371–377.

Halpern, D. J., Lega, B. C., Gross, R. E., Wu, C., Sperling, M. R., Aronson, J. P., … Kahana, M. J. (2025). Study-phase reinstatement predicts subsequent recall. *Nature Neuroscience, 28*(4), 883–890. https://doi.org/10.1038/s41593-025-01884-8

Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Memory & Cognition, 38*(1), 102–115. https://doi.org/10.3758/MC.38.1.102

Horner, A. J., & Burgess, N. (2013). The associative structure of memory for multi-element events. *Journal of Experimental Psychology: General, 142*(4), 1370–1383. https://doi.org/10.1037/a0033626

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Jacoby, L. L., Wahlheim, C. N., & Kelley, C. M. (2015). Memory consequences of looking back to notice change: Retroactive and proactive facilitation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(5), 1282–1297. https://doi.org/10.1037/xlm0000123

Koppenaal, R. J. (1963). Time changes in the strengths of A–B, A–C lists; Spontaneous recovery? *Journal of Verbal Learning and Verbal Behavior, 2*(4), 310–319. https://doi.org/10.1016/S0022-5371(63)80099-7

Lenth, R. V. (2024). emmeans: Estimated marginal means, aka least-squares means (Version 1.10.2) [Computer software]. https://CRAN.R-project.org/package=emmeans.

Lohnas, L. J. (2025). Temporal associations supporting repetitions in free recall. *Psychonomic Bulletin & Review, 32*(5), 2211–2219. https://doi.org/10.3758/s13423-025-02673-x

Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2015). Expanding the scope of memory search: Modeling intralist and interlist effects in free recall. *Psychological Review, 122*(2), 337–363. https://doi.org/10.1037/a0039036

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *The Journal of Open Source Software, 6*(60), 3139. https://doi.org/10.21105/joss.03139

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28*(2), 203–208. https://doi.org/10.3758/BF03204766

MacLeod, C. M. (2024). Interference theory: History and current status. In M. J. Kahana, & A. D. Wagner (Eds.), *The Oxford handbook of human memory, two Volume Pack* (1st ed., pp. 1173–1208). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190917982.013.40.

Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *The American Journal of Psychology, 53*(2), 173–203.

Mensink, G.-J., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review, 95*(4), 434–455. https://doi.org/10.1037/0033-295X.95.4.434

Negley, J. H., Kelley, C. M., & Jacoby, L. L. (2018). The importance of time to think back: The role of reminding in retroactive effects of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(9), 1352–1364 (2018-38490-001) https://doi.org/10.1037/xlm0000512.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36*(3), 402–407. https://doi.org/10.3758/BF03195588

R Core Team. (2025). *R software*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Raaijmakers, J. G. W. (2003). Spacing and repetition effects in human memory: Application of the SAM model. *Cognitive Science, 27*(3), 431–452. https://doi.org/10.1207/s15516709cog2703_5

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88*(2), 93–134.

Robbins, D., & Bray, J. F. (1974). The spacing effect and the A-B, A-C paradigm: Evidence for retroactive facilitation. *Journal of Experimental Psychology, 103*(3), 420–425. https://doi.org/10.1037/h0037203

Siegel, L. L., & Kahana, M. J. (2014). A retrieved context account of spacing and repetition effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(3), 755–764. https://doi.org/10.1037/a0035585

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373. https://doi.org/10.1037/h0020071

Tzeng, O. J., & Cotton, B. (1980). A study-phase retrieval model of temporal coding. *Journal of Experimental Psychology: Human Learning and Memory, 6*(6), 705–716. https://doi.org/10.1037/0278-7393.6.6.705

Underwood, B. J. (1948). "Spontaneous recovery" of verbal associations. *Journal of Experimental Psychology, 38*(4), 429–439. https://doi.org/10.1037/h0059565

Underwood, B. J. (1969). Attributes of memory. *Psychological Review, 76*(6), 559–573. https://doi.org/10.1037/h0028143

Wahlheim, C. N., Garlitch, S. M., & Kemp, P. L. (2021). Context differentiation and remindings in episodic memory updating. *The Psychology of Learning and Motivation, 75*, 245–277.

Wahlheim, C. N., & Jacoby, L. L. (2013). Remembering change: The critical role of recursive remindings in proactive effects of memory. *Memory & Cognition, 41*(1), 1–15. https://doi.org/10.3758/s13421-012-0246-9

Wahlheim, C. N., Maddox, G. B., & Jacoby, L. L. (2014). The role of reminding in the effects of spaced repetitions on cued recall: Sufficient but not necessary. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(1), 94–105. https://doi.org/10.1037/a0034055

Wahlheim, C. N., Smith, S. T., Garlitch, S. M., & Wiley, R. W. (2023). Interpolated retrieval retroactively increases recall and promotes cross-episode memory interdependence. *Learning & Memory, 30*(8), 151–163. https://doi.org/10.1101/lm.053782.123

Wahlheim, C. N., Smith, W. G., & Delaney, P. F. (2019). Reminders can enhance or impair episodic memory updating: A memory-for-change perspective. *Memory, 27* (6), 849–867. https://doi.org/10.1080/09658211.2019.1582677

Wahlheim, C. N., & Zacks, J. M. (2019). Memory guides the processing of event changes for older and younger adults. *Journal of Experimental Psychology: General, 148*(1), 30–50. https://doi.org/10.1037/xge0000458

Wahlheim, C. N., & Zacks, J. M. (2025). Memory updating and the structure of event representations. *Trends in Cognitive Sciences, 29*(4), 380–392. https://doi.org/10.1016/j.tics.2024.11.008

Ward, G., & Tan, L. (2023). The role of rehearsal and reminding in the recall of categorized word lists. *Cognitive Psychology, 143*, Article 101563. https://doi.org/10.1016/j.cogpsych.2023.101563

Winograd, E., & Soloway, R. M. (1985). Reminding as a basis for temporal judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(2), 262–271.