Memory & Cognition
https://doi.org/10.3758/s13421-025-01834-6

A role for context-cued study-phase retrievals in episodic memory updating

Jennifer L. Fiedler[1] · Kelly S. Giovanello[1] · Robert W. Wiley[2] · Christopher N. Wahlheim[2]

Received: 16 June 2025 / Accepted: 21 November 2025
© The Psychonomic Society, Inc. 2026

Abstract
Navigating changes is fundamental to everyday life and requires updating existing memories to incorporate new details. This study examined mechanisms underlying how reinstating an earlier event's context during a later event influences memory for both events. Two theories predict opposite outcomes. Interference theory holds that reinstating context from an existing memory while experiencing a new, overlapping event produces response competition and impairs memory for both. In contrast, integration theory predicts that context reinstatement cues retrieval of earlier memories, enabling associative encoding of past and present events that enhances memory. Prior work favors the latter, showing that reinstatement improves memory. Three experiments extended this work by directly testing roles for study-phase retrievals and change awareness during study and test. Word pairs with shared cues but changed responses (A-B, A-D) were presented with background contexts that either repeated or changed. Repeating contexts increased detection of changes and recall of earlier responses during study, both indexes of study-phase retrievals, as well as later cued recall of earlier (B) and changed (D) responses. The recall benefit was proportional to the extent of study-phase retrievals, implicating retrieval practice. Moreover, the effect was enhanced when participants remembered that changes had occurred, highlighting the role of recollecting integrated representations that included change attributes. These findings align with integration theory, suggesting that memory updating is most effective when current events cue retrieval of prior memories and engender associative encoding of past and present events, establishing elaborate representations that support subsequent recall.

Keywords Integration · Interference · Study-phase retrieval · Memory updating · Retroactive effects of memory

Every day, we experience new events that are related to but differ from past experiences. Consider the following scenario. Ana attended the same professional conference two years in a row. In 2023, she attended a talk about implicit memory by Dr. Smith. In 2024, Ana attended another talk by Dr. Smith, this time about explicit memory. Due to their overlap, Ana's subsequent memory of the recent talk could interfere with her recall of the earlier talk. This is an example of retroactive interference—a primary cause of forgetting (for reviews, see Anderson & Neely, 1996; MacLeod, 2024). To reduce interference, the contexts of the talks could be differentiated from each other, thus preventing confusion about what concepts were discussed on each occasion. Counterintuitive to this notion, interference may also be reduced by associating events, their contexts, and details into a unified memory—a process referred to as integrative encoding. When successful, this process can update episodic memories, leading to retroactive facilitation, as shown by better memory after changes. Here, we replicate earlier findings showing that reinstating contextual details of an earlier event during a current event improves episodic memory updating, which we operationally define as correct recall of conflicting details from both events. Critically, we directly examine how prior-event retrievals yield awareness of changes that support such updating, thus leading to retroactive facilitation.

One method for assessing such retroactive effects of memory is the A-B, A-D paired-associate paradigm. In this paradigm, participants learn cue-response pairs (A-B) before learning changed pairs with shared cues and different

✉ Jennifer L. Fiedler
jfiedler@unc.edu

1 Department of Psychology & Neuroscience, University of North Carolina at Chapel Hill, 235 E. Cameron Avenue, Chapel Hill, NC 27599, USA

2 Department of Psychology, University of North Carolina at Greensboro, Greensboro, NC, USA

Published online: 23 January 2026

responses (A-D). Retroactive effects are assessed by measuring recall of first (B) and second (D) responses (for a summary, see Anderson & Neely, 1996). To relate this to the earlier example, A-B and A-D pairs are analogous to the first talk (e.g., "Smith–implicit") and second talk (e.g., "Smith–explicit"), respectively. Retroactive interference would present as impaired recall of the B response relative to a control condition in which changes did not occur, and/or increased intrusions from D responses. Retroactive facilitation would present as improved recall of the B response, which could be accompanied by fewer intrusions. Observing reduced interference together with facilitation resulting from a manipulation suggests that the experimental conditions improved episodic memory updating.

Retroactive interference and facilitation have both been observed in A-B, A-D paradigms. Interference effects were once theorized to arise from the unlearning of original associations when learning changed associations (Melton & Irwin, 1940) and response competition at retrieval when distinct events share a cue (Postman & Underwood, 1973). Support for these accounts comes from studies showing that changing responses paired with the same cue from one list to the next impaired subsequent memory for earlier responses, and evidence for facilitation effects—enhanced memory for earlier responses—comes from other A-B, A-D studies (for a review, see Anderson & Neely, 1996). It was originally theorized that facilitation effects arise from an associative mediation process activated when responses are semantically similar (Barnes & Underwood, 1959). Later, others suggested that retroactive facilitation may arise when A-D pairs cue retrieval of A-B pairs (Bruce & Weaver, 1973), which may occur more for similar pairs. To extend this reasoning, interference reduction may depend on the potential for cross-episode associations to be encoded, though such integration was not articulated early on.

Other work has established a role for contextual associations, showing that creating differentiated contextual representations can reduce interference (for a review, see Smith & Vela, 2001). For example, reductions in retroactive interference in A-B, A-D paradigms was observed when the learning contexts were varied between lists by changing the environments (e.g., Bilodeau & Schlosberg, 1951) and increasing time between A-B and A-D learning (e.g., Underwood & Ekstrand, 1966). Creating dissimilar list contexts in these ways presumably counteracted interference by preventing response competition. Findings such as these support classic interference theory and yet they clearly conflict with results suggesting that uniting contexts can enhance memory in A-B, A-D paradigms (e.g., Barnes & Underwood, 1959; Bruce & Weaver, 1973). One goal here was to further test these mechanisms by manipulating whether perceptual contexts associated with A-B and A-D pairs are reinstated. To foreshadow, we find support for integration as a key mechanism. Another goal was to examine potential roles for study-phase retrieval and awareness of changes in such effects. We consider relevant background studies next before describing our approach.

As mentioned above, research showing retroactive facilitation in A-B, A-D tasks suggests that reinstating context should promote memory updating. The present study was inspired by work showing retroactive facilitation across various lags between A-B and A-D pairs that also proposed a role for retrieval practice of A-B pairs in such effects (Bruce & Weaver, 1973). They suggested that overlapping features of A-D pairs cue retrievals of A-B pairs. Such retrievals have been described as "study-phase retrievals," and are assumed to play a key role in memory enhancement across a variety of tasks (e.g., Hintzman et al., 1975; Tzeng & Cotton, 1980). Importantly, these retrievals can occur spontaneously and covertly (e.g., Hintzman, 2011), which may occur when more features are shared between A-B and A-D pairs, such as semantic and contextual information. However, study-phase retrievals may also be engaged strategically using controlled processes, such as when task instructions require overt responses (Jacoby, 1974). Controlled study-phase retrievals may have played a critical role in the retroactive facilitation observed when participants were forewarned that items would change (see Robbins & Bray, 1974a, b). This advisement of change was likely to have motivated participants to look for changes and engage encoding processes that compared related events.

The role of change awareness in retroactive facilitation was directly examined using an A-B, A-D paradigm that manipulated how participants engaged strategically controlled, study-phase retrievals (Jacoby et al., 2015, Experiment 1). Participants learned two lists with A-B pairs in both lists and A-D pairs only in List 2. The manipulation of change awareness occurred before List 2. One group was told to look for changes from List 1 or 2 (*N*-Back) while the other group was told to look for changes only within List 2 (Within-List Back). The *N*-Back group should have been more aware of between-list changes. On a final cued recall test, B-response recall (from List 1) showed retroactive facilitation in the *N*-Back group but not the Within-List Back group. Memory for changes was also higher in the N-Back group, suggesting that instructions to look for changes promoted study-phase retrievals that engendered integrative encoding of changes. Related studies showed similar results of increasing change awareness, for example by increasing study time for A-D pairs in List 2 (Garlitch & Wahlheim, 2020; Negley et al., 2018) and interpolating A-B retrieval practice with feedback prior to A-D learning (Wahlheim et al., 2023). Importantly, such manipulations showed increased B-response recall in association with increased memory for change, thus implicating roles for study-phase retrieval and integration in facilitation effects.

Of particular relevance here, reinstatement of perceptual contextual features has also been used to promote retroactive facilitation in A-B, A-D paradigms. To show such effects, the key study upon which the current experiments are based manipulated whether the same or different background scene appeared with corresponding A-B and A-D word pairs (Cox et al., 2021). Both study phases and the final cued recall test were separated by 24-hr intervals. In the first two experiments, participants learned pairs to criterion across each study phase, then took a cued recall test without background scenes present. The test directed retrieval first to the B then to the D response (Experiment 1) or used a modified modified free recall (MMFR) procedure in which participants reported responses from both lists in the order that they came to mind (Experiment 2). Retroactive facilitation was observed when scenes repeated for A-B and A-D pairs, whereas retroactive interference was observed when scenes changed. Evidence for context-cued integration was indicated by statistical dependences showing higher B-response recall when D responses were also recalled, but only when the A-B scenes were reinstated during A-D learning. These findings were attributed to an integrative encoding mechanism that was amplified when repeated scenes served as potent retrieval cues. However, that study did not verify differences in study-phase retrieval and change awareness that would further implicate integrative encoding.

According to the Memory-for-Change (MFC) framework proposed by Larry Jacoby and colleagues (Jacoby & Wahlheim, 2013; Wahlheim & Jacoby, 2013), such processes should play critical roles in retroactive facilitation. Specifically, the account proposes that whether changed responses improve or impair memory depends on whether study-phase retrievals engender awareness of changes that become integrated with responses. Importantly, the associative encoding that characterizes integration should lead to memory representations that carry the subjective experience of change as a contextual attribute, akin to metadata, that later aids in identifying the sources of generated responses. This view was inspired by findings from the interference and temporal memory literatures showing that repeated stimulus features promote retroactive facilitation (Barnes & Underwood, 1959; Bruce & Weaver, 1973) and memory for order (Tzeng & Cotton, 1980; Winograd & Soloway, 1985), presumably by reminding people of past events, thus enabling integrative encoding with current events. This idea has also been supported by neural investigations showing that study-phase retrievals and context reinstatement are associated with reductions in retroactive interference (Koen & Rugg, 2016; Kuhl et al., 2010). The assertion that we test here is that reinstating background scenes, a type of perceptual context used by Cox et al. (2021), should evoke study-phase retrievals that promote detection of changes and integration.

The present study

We extended upon Cox et al. (2021) by examining perceptual context reinstatement effects in A-B, A-D paradigms with direct measures of change detection and study-phase retrieval during A-D study, as well as memory for changes on the final test. Retroactive memory effects were assessed by comparing B-response recall in A-B, A-D and A-B, C-D conditions on cued recall tests, which first probed the B response. This approach is justified because Cox et al. showed that retroactive effects converge across directed retrieval and MMFR tests. Across three experiments, we evaluated how retrieval operations during A-D study affected subsequent memory for B and D responses and awareness of their changes, as required to test claims from the MFC framework. Experiments 1 and 2 overtly measured change detection and study-phase retrieval to confirm the roles of A-B retrieval and A-D comparison processes in context reinstatement effects. Experiment 3 tested whether the recall patterns from the prior two experiments would replicate without overt change detection or study-phase retrieval, and whether statistical dependencies between B-response recall and memory for changes would be greater when context was reinstated, as in Cox et al.

This approach allowed us to test competing theoretical predictions of interference and integration theories regarding context reinstatement effects on episodic memory updating. To reiterate, interference theory predicts that context reinstatement during A-D study will impair memory updating, especially when changes are detected and study-phase retrievals occur. Conversely, integration theory predicts that by promoting context reinstatement and A-B retrievals, episodic memories will be more effectively updated. Detecting change is crucial to updating because it affords the conscious awareness necessary to notice and encode differences. Based on Cox et al. (2021), we expected that reinstating contexts will improve change detection, retrieval of B responses during A-D study, as well as subsequent recall of B and D responses and memory for changes on the cued recall test. We also expected to observe a dependency in recall of B and D responses, showing greater B-response recall when D responses are also recalled and changes are remembered. This will be especially prevalent when changes were detected and B responses were recalled while studying A-D pairs during List 2, because such instances are necessary to engender the integrative encoding that yields later dependencies. Collectively, these findings would be compatible with predictions from the MFC framework (Wahlheim & Jacoby, 2013).

General statistical methods

All analyses were conducted with R software (R Core Team, 2025). We fitted logistic mixed-effects models with the "glmer" function from the *lme4* package (Bates et al., 2015), including subjects and items as random intercept effects. We performed hypothesis testing with Type 2 Wald's Chi Square tests using the Anova function from the car package (Fox & Weisberg, 2019) and pairwise comparisons using the Tukey method from the "emmeans" function in the *emmeans* package (Lenth, 2024). We report odds ratios (*OR*s) effect sizes along with their 95% confidence intervals. We coded cued recall responses using a strict criterion that required exact matches to the original spelling. Misspelled words comprised <2% of all trials.

Experiment 1

Experiment 1 examined the effect of context reinstatement on episodic memory updating while overtly measuring the detection of changed A-D pairs. Participants studied two lists with word pairs that shared cues and changed responses across lists (A-B, A-D) or were unique across lists (A-B, C-D). We used a change detection measure to assess the extent to which repeating a background scene across lists encouraged participants to think back to List 1 during List 2. In a cued recall test, participants attempted to recall the B response, indicated whether the response changed from List 1 to List 2, and if so, attempted to recall the D response. We predicted that repeating scenes would increase change detection, and thus, study-phase retrievals, leading to improved subsequent memory for both responses and memory for the fact that they changed.

Method

All research reported below was approved by the Institutional Review board at the University of North Carolina at Chapel Hill (UNC-CH).

Participants

The participants were 74 undergraduate students from the University of North Carolina at Chapel Hill. Participants were recruited from the Psychology and Neuroscience Department subject pool and compensated with course credit. One participant was excluded from analysis due to program failure and another was excluded due experimenter error. The analyses reported below include the data from the remaining 72 participants (50 women), ages 18–30 years (M = 19.80, SD = 2.10). The stopping rule was based on available resources and sample size comparisons with prior studies using A-B, A-D paradigms. We recruited at least 50% more participants than in prior studies (e.g., Cox et al., 2021; Garlitch & Wahlheim, 2020; Negley et al., 2018) to improve sensitivity to modest effect sizes. A sensitivity analysis using *simr* (Green & MacLeod, 2016) indicated that N = 72 provides 80% power to detect the moderate observed effect (OR = 1.61) of the context reinstatement manipulation on D-response intrusions below.

Design

The experiment used a 2 (item type: A-B, C-D vs. A-B, A-D) × 2 (context type: match vs. mismatch) within-subjects design.

Materials

The materials consisted of 84 critical word triplets plus eight buffer triplets, each including a cue (e.g., cloud) and two responses (e.g., white, nine). For the critical items, cue lengths ranged from 3–8 letters (M = 5.39, SD = 1.30) and response lengths also ranged from three to eight letters (M = 5.29, SD = 1.21). We created the triplets such that cues were semantically associated with both responses to ensure that recall performance would be above floor. We verified these associations by extracting the forward association strength values from the University of South Florida Free Association norms (Nelson et al., 2004), which are the normative probabilities that participants generated responses when given cues. For the critical items, the forward association strengths ranged from .01–.10 (M = .04, SD = .02). In contrast, responses within triplets (e.g., white and nine) were not normatively associated with each other.

The materials also included 84 critical scene images (42 indoor / 42 outdoor) plus eight buffer scenes. Scenes were color photographs from the Change Blindness Database (Sareen et al., 2016). For counterbalancing, the 84 scenes were rotated across 24 formats (two lists per format block), such that indoor and outdoor versions of each arbitrarily numbered scene were balanced across context type (match vs. mismatch) and distributed across item type (A-B, C-D vs. A-B, A-D). In context-match conditions, the same scene repeated across lists; in context-mismatch conditions, indoor and outdoor versions were cross-assigned (e.g., Indoor Scene 1 in List 1 / Outdoor Scene 1 in List 2). Counterbalancing was complete for context type but not fully crossed with item type; each scene appeared equally often as a match and a mismatch, but no scene appeared in all four item type × context type combinations to avoid an unmanageable number of formats while still distributing scenes across item types.

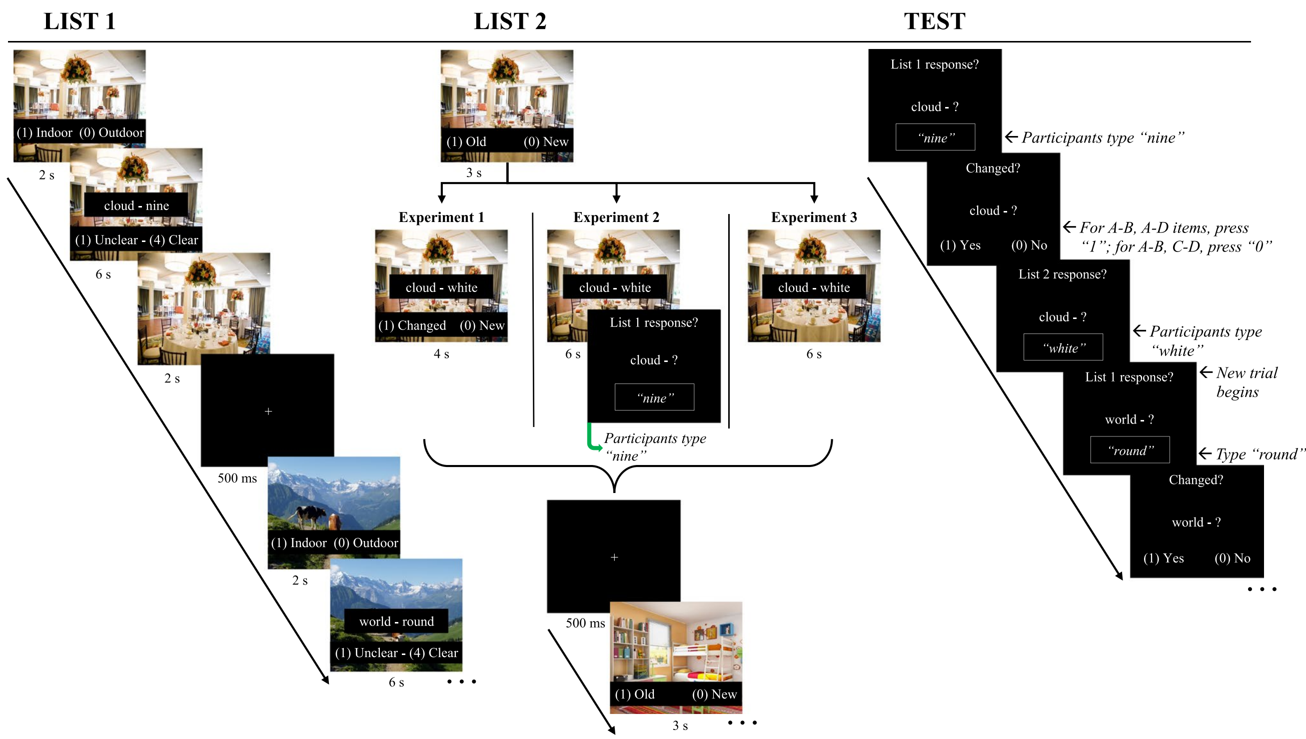

Fig. 1 Schematic of the experimental procedures. On List 1 trials, participants first indicated whether a scene belonged indoors or outdoors. Next, they studied an A-B pair (e.g., cloud–nine) in association with the scene. On List 2 trials, participants first indicated whether a scene was from List 1 (Old/New judgments), then viewed a word pair centered in the scene. Word pairs were either changed (e.g., "cloud–white") or new ("read–page"), and scenes were either the same as in List 1 (match) or new in List 2 (mismatch). Next, participants made change detection judgments (Experiment 1), studied the pair then recalled B responses after each changed pair (Experiment 2), or only studied the pair (Experiment 3). On test trials, prompts appeared against black backgrounds. Participants first attempted to recall the B response, then indicated whether the response changed from List 1 to List 2, and if they responded "yes," then attempted to recall the D response. (Color figure online)

The experiment comprised three phases: List 1, List 2, and test. List 1 and List 2 each included four buffer pairs (one from each combination of item and context type) at the beginning and end of each list to control for primacy and recency effects. Additionally, List 1 and List 2 each included 56 critical pairs divided evenly among the four combinations of item and context types. Eighty-four total items were required because the A-B, A-D conditions comprised one set of 28 items that share cues but included different responses across lists, whereas the A-B, C-D conditions comprised two unique sets of 28 cue-response pairs in each list (56 items total). The cued recall test began with eight practice trials derived from the buffers (two from each combination of item and context type) then presented cues from all 56 critical items in List 1.

Procedure

E-Prime 2 (Psychology Software Tools, Pittsburgh, PA, USA) controlled the stimulus presentation. Figure 1 displays a schematic of the procedure. Variations in presentation durations across phases were determined through piloting. We chose durations that allowed participants to comfortably execute each task. In all phases, trials appeared in pseudorandom orders with the stipulation that no more than three items from the same condition appeared consecutively. The average serial position was matched across all within-subject conditions.

Before List 1 began, participants were told that they would view indoor and outdoor scenes, then word pairs would appear superimposed onto the scenes. They were told that their tasks would be to indicate if the scenes belonged indoors or outdoors and to rate how well they could imagine the words interacting with the scenes. On each List 1 trial, a scene first appeared for 2 s, and participants pressed a key to indicate whether it belonged indoor (1) or outdoor (0).[1] After 2 s, the scene remained, a word pair appeared in the center for 6 s, and participants pressed a key to rate

[1] Scene classification accuracy was compared in each experiment using models with item and context type as fixed effects. Accuracy was extremely high in all experiments: The model estimated probabilities for all cells across combinations of item and context type were ≥.98.

the clarity with which they could imagine the words interacting with the scene on a four-point scale (1 = unclear - 4 = clear). Participants were also instructed to study the words for a later test. After 6 s, the words disappeared, and the scene remained for 2 s so that participants could continue encoding their image. A black screen ISI intervened between trials for 500 ms.

Before List 2 began, participants were told that they would view indoor and outdoor scenes, but that their tasks would be to indicate if they recognized the scenes from List 1 and to indicate if they detected changed word pairs. On each List 2 trial, a scene first appeared for 3 s, and participants pressed a key to indicate whether it was "old" from List 1 (1) or "new" to List 2 (0). After 3 s, the scene remained, a word pair appeared in the center for 4 s, and participants pressed a key to indicate whether the pair comprised a repeated cue and changed response (1) or was entirely new (0), relative to List 1. A black screen ISI intervened between trials for 500 ms.

Before the cued recall test, participants were told that they would see cue words and be asked to recall responses from each list and whether those responses changed. On each cued recall trial, a cue from a List 1 pair appeared with a question mark in lowercase white text on a black screen (e.g., cloud - ?). Participants first attempted to type the response word from the List 1 A-B pair and then pressed a key to indicate whether the response word changed across lists (1 = yes, 0 = no). When participants responded "yes," another prompt appeared asking for recall of the response word from the A-D pair in List 2. After participants attempted to type that response, the program advanced to the next trial. When participants responded "no," the program advanced to the next trial. Responses were self-paced, and participants were allowed to pass when they could not recall the correct response. A black screen ISI intervened between trials for 500 ms.

Results and discussion

List 2 study phase: Scene recognition

We compared the probability of "old" responses on recognition trials in List 2 (Fig. 2, left) to matched (old) and mismatched (new) scenes to assess task adherence and the extent to which context reinstatement directed retrieval to List 1. A model with fixed effects of context and item type indicated a significant effect of context type, $\chi^2(1) = 1189.45$, $p < .001$, $OR = 151.12$ (CI [113.63, 200.98]), showing that that participants responded "old" more to repeated scenes. This verifies that scene reinstatement cued retrieval of List 1 contexts for most items. No other effects were significant, largest $\chi^2(1) = 2.37$, $p = .12$, $OR = 1.45$ (CI [0.90, 2.32]).

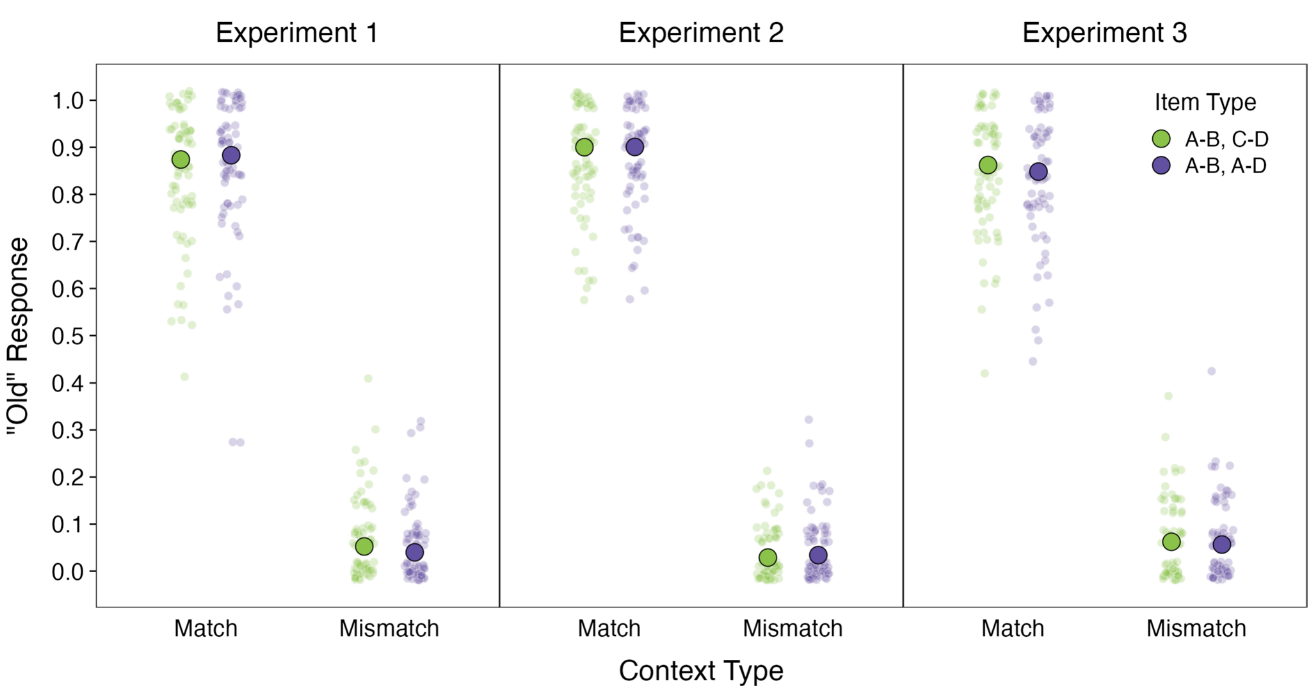

Fig. 2 Scene recognition accuracy in List 2. Smaller points are individual participant probabilities, and larger points are probabilities for each condition in aggregate estimated from mixed-effects models. Error bars are omitted from this figure because they were fully obscured by the diameters of model-estimate points. (Color figure online)

List 2 study phase: Change detection

We compared the probability of "changed" responses in List 2 (Fig. 3A) to assess correct detection (A-B, A-D conditions) and incorrect false alarms (A-B, C-D conditions). A model with fixed effects of context and item type indicated a significant effect of item type, showing higher correct detection than false alarms, $\chi^2(1) = 873.72$, $p < .001$, $OR = 22.72$ (*CI* [17.86, 28.91]). There was also a significant effect of context type, $\chi^2(1) = 58.62$, $p < .001$, $OR = 1.67$ (*CI* [1.42, 1.95]), that was qualified by a significant interaction, $\chi^2(1) = 46.65$, $p < .001$, $OR = 3.02$ (*CI* [2.20, 4.14]). Pairwise comparisons showed significantly higher correct detection in the match than the mismatch condition, z ratio = 10.25, $p < 001$, OR = 2.89 (*CI* [2.36, 3.55]), and no significant difference between context conditions for false alarms, z ratio = 0.34, $p = .73$, $OR = 1.04$ (*CI* [0.82, 1.33]). These results suggest that context reinstatement directed retrieval to List 1, which enabled the detection of changes in List 2.

Test phase: Cued recall and memory for change

Figure 4 (left panels) displays the response probabilities for each cued recall measure. For correct recall of B responses and "changed" classifications (Fig. 4A and C), we compared probabilities across conditions using models with context and item type as fixed effects. For intrusions and recall of D responses (Fig. 4B and D), we only compared context types for A-B, A-D items because A-B, C-D items had List 2 cues that did not match the List 1 cues.

The model for B-response recall (Fig. 4A, left) indicated significant effects of context, $\chi^2(1) = 41.22$, $p < .001$, OR = 2.13 (*CI* [1.74, 2.60]), and item type, $\chi^2(1) = 16.98$, $p <$ .001, $OR = 1.33$ (*CI* [1.16, 1.54]), as well as a significant interaction, $\chi^2(1) = 16.34$, $p < .001$, $OR = 1.80$ (*CI* [1.35, 2.39]). Pairwise comparisons indicated significantly higher recall for A-B, A-D than A-B, C-D items in the match condition, z ratio = 5.78, $p < .001$, $OR = 1.79$ (*CI* [1.47, 2.18]) and no significant difference between item types in the mismatch condition, z ratio = 0.07, $p = .95$, $OR = 1.01$ (*CI* [0.82, 1.24]). Additionally, recall was significantly higher for A-B, A-D items in the match than the mismatch condition, z ratio = 7.41, $p < .001$, $OR = 2.13$ (*CI* [1.74, 2.60]), but not significantly different between contexts for A-B, C-D items, z ratio = 1.63, $p = .10$, $OR = 1.18$ (*CI* [0.97, 1.45]). Finally, recall was significantly higher for A-B, A-D items in the match condition than A-B, C-D items in the mismatch condition, z ratio = 7.34, $p < .001$, $OR = 2.12$ (*CI* [1.63, 2.75]). These results suggest that context reinstatement enhanced recall enough to produce retroactive facilitation for changed items, just as before (cf. Cox et al., 2021).

The model for D-response intrusions (Fig. 4B, left), including only context type as a fixed effect, indicated a significantly lower probability in the match than the mismatch condition, $\chi^2(1) = 4.42$, $p = .04$, $OR = 1.45$ (*CI* [1.03, 2.06]). This suggests that repeating scenes promoted study-phase retrievals that supported later rejection of those response as intrusions.

The model for "changed" classifications (Fig. 4C, left), which were correct remembrances for A-B, A-D items

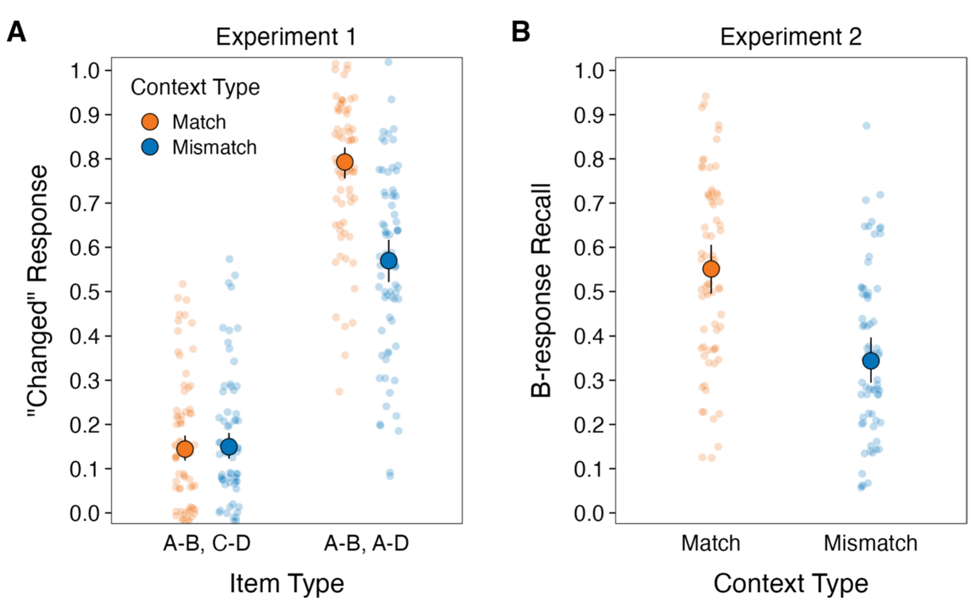

Fig. 3. **A** Change detection (Experiment 1). **B** List 2 B-response recall (Experiment 2). Smaller points are individual participant probabilities, and larger points are probabilities for each condition in aggregate estimated from mixed-effects models. Error bars are 95% confidence intervals. (Color figure online)

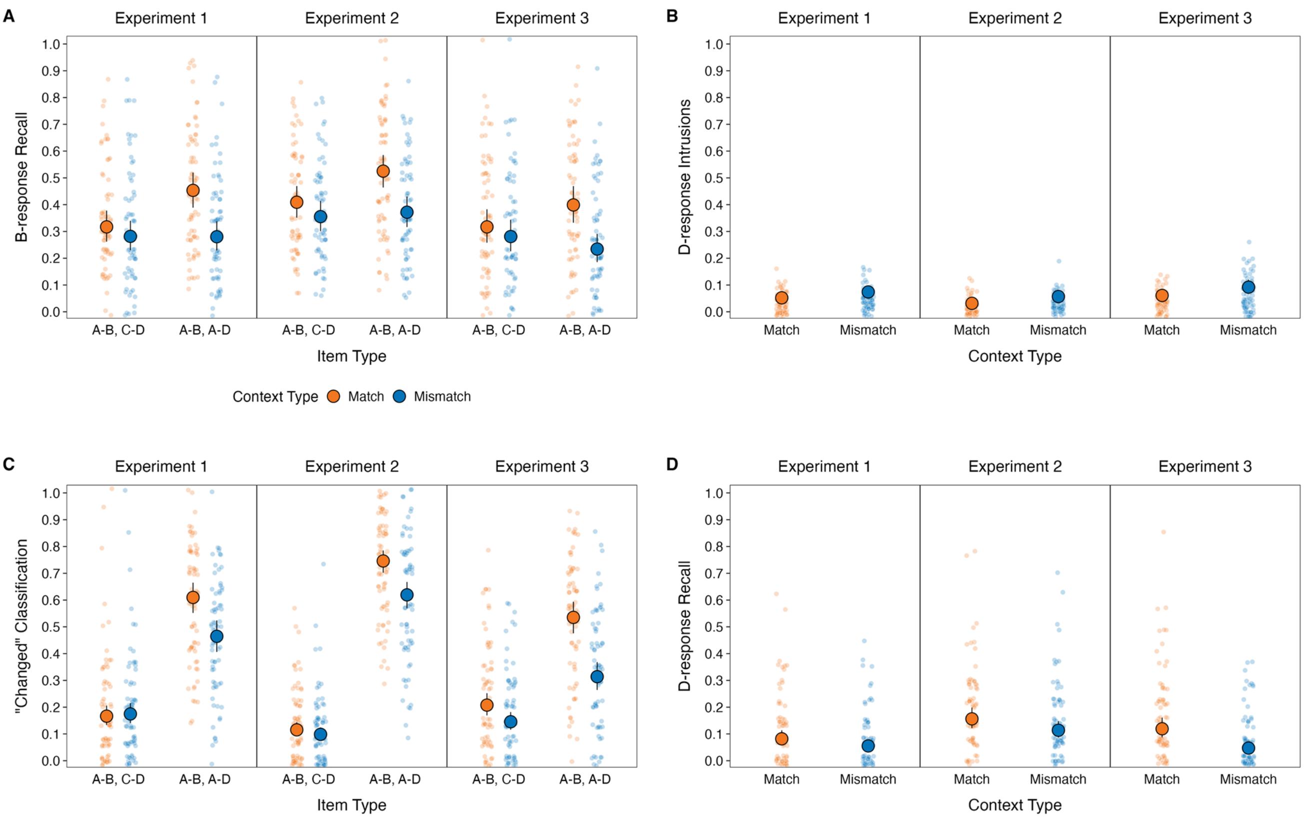

Fig. 4 Cued recall test response probabilities. Smaller points are individual participant probabilities, and larger points are probabilities for each condition in aggregate estimated from mixed-effects models. Error bars are 95% confidence intervals. Error bars that appear to be missing are fully obscured by the diameters of point estimates. (Color figure online)

and incorrect false alarms for A-B, C-D items, indicated a significant effect of item type, $\chi^2(1) = 476.60$, $p < .001$, $OR = 5.66$ (CI [4.85, 6.61]), showing higher probabilities for correct than incorrect classifications. There was also a significant effect of context type, $\chi^2(1) = 20.02$, $p < .001$, $OR = 1.30$ (CI [1.12, 1.51]), that was qualified by a significant interaction, $\chi^2(1) = 17.82$, $p < .001$, $OR = 1.91$ (CI [1.41, 2.58]). Pairwise comparisons indicated significantly higher correct classifications in the match than the mismatch condition, z ratio $= 6.13$, $p < .001$, $OR = 1.80$ (CI [1.49, 2.17]), and no significant difference in incorrect classifications between contexts, z ratio $= 0.49$, $p = .63$, $OR = 1.06$ (CI [0.84, 1.34]). These results suggest that by improving the detection of changes in List 2, context reinstatement also promoted better subsequent memory for changes at test.

The model for D-response recall after participants reported remembering that responses had changed (Fig. 4D, left), which included only a fixed effect of context type, indicated a significant effect, $\chi^2(1) = 7.09$, $p < .001$, $OR = 1.52$ (CI [1.12, 2.07]), showing higher recall in the match than the mismatch condition. Taken together with the other cued recall measures, these findings suggest that reinstating background scenes promoted change detection by directing retrieval to List 1, which improved recall of all responses (i.e., enhanced episodic memory updating), suggesting that context reinstatement promoted integrative encoding.

Test phase: Cued recall conditionalized on memory for change

To examine how context reinstatement moderated the associations between memory for changes and B-response recall, we examined B-response recall conditionalized on detection of changed responses in List 2 and "changed" classifications at test for A-B, A-D items (Fig. 5). Prior work showed that detecting change during study and remembering changes at test were associated with enhanced recall of B responses (Negley et al., 2018). This suggests that the retrieval of B and its co-activation with A-D leads to memory representations of B responses that includes attributes of the change experience, which should confer mnemonic advantages. If this occurred here, then B-response recall would be enhanced when participants remembered changes, especially if they could also recall the D response. Taken with the increased D-response recall when contexts were reinstated, this would

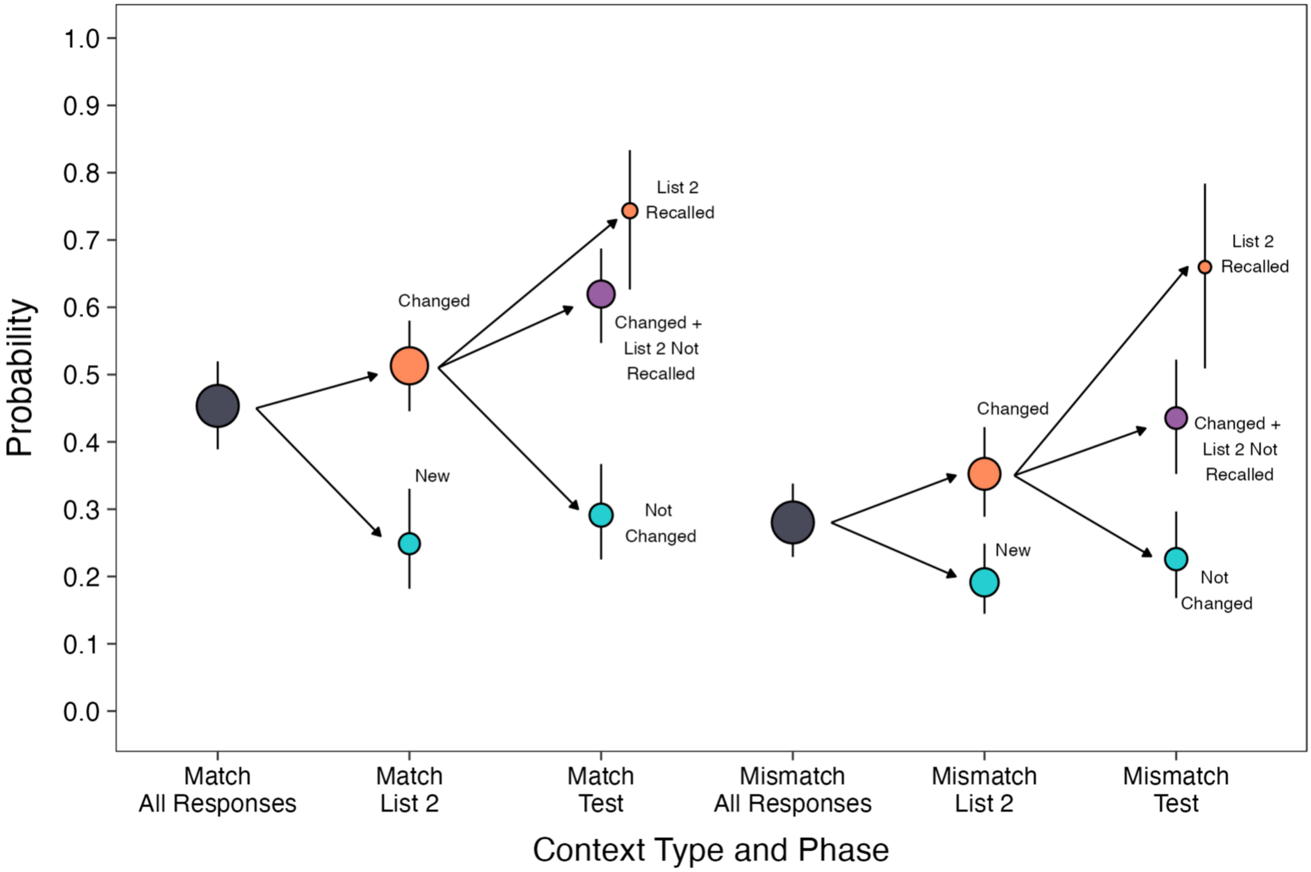

Fig. 5 Conditional recall of B responses from List 1 for A-B, A-D items: Experiment 1. Points are estimated probabilities from a mixed-effects model. Point areas reflect the relative observation counts. The aggregate values from Fig. 4 appear here in dark gray (Match and Mismatch All Responses). Recall conditioned on List 2 change detection responses appear in orange and teal (Match and Mismatch List 2). Recall for detected changes conditioned on "changed" classifications at test appear in orange, purple, and teal (Match and Mismatch Test). Error bars are 95% confidence intervals. (Color figure online)

suggest that context reinstatement promotes study-phase retrievals, change awareness, and effective memory updating. In the following conditional analyses, we do not discuss context effects redundant with those above.

B-response recall conditionalized on change detection in List 2 suggested that such detection assayed study-phase retrievals. A model with change detection and context type as fixed effects indicated a significant effect of change detection, showing higher recall when change was detected than when it was not detected, $\chi^2(1) = 60.74$, $p < .001$, OR = 2.71 (CI [2.11, 3.47]), and no interaction, $\chi^2(1) = 1.71$, p = .19, OR = 1.38 (CI [0.85, 2.25]), suggesting that change detection was often accompanied by retrieval practice of B responses.

More critical to the proposal that study-phase retrievals and change awareness contribute to memory updating, we assessed B-response recall for detected changes, conditionalized on memory for change at test. According to the MFC framework, detected changes should provide more opportunities for integrative encoding, and measuring memory for changes at test should provide evidence for the extent to which integrative encoding engenders encoding of change attributes that are later recollected (Wahlheim & Jacoby, 2013). A classification variable included three types of change remembrance reflecting whether or not changes were remembered, and if so, whether D responses were also recalled. A model with context type and classification as fixed effects indicated a significant effect of classification, $\chi^2(2) = 99.52$, $p < .001$, and no significant interaction, $\chi^2(2) = 2.23$, $p = .33$. Pairwise comparisons indicated significantly higher recall when changes were remembered with D recall as compared to without D recall, and when changes were remembered without D recall than when changes were not remembered, smallest z ratio = 3.36, $p = .002$, $OR = 2.11$ (CI [1.25, 3.57]). Together, these results replicate prior findings of retroactive facilitation accompanied by dependencies between B and D recall arising when detected changes were later recollected (Negley et al., 2018), and support the predictions from the MFC framework.

Summary

Experiment 1 showed that reinstating background scenes for changed word pairs enhanced change detection, B-response recall, remembrance of changes, and D-response recall. Conditional analyses showed that detecting change was associated with enhanced B-response recall, suggesting that such detection reflected study-phase retrievals. These analyses also showed that when changes were detected, subsequent recall of B responses was higher when participants both remembered that responses changed and recalled D responses. Collectively, these results suggest that context reinstatement promoted integrative encoding between A-B and A-D pairs, which counteracted retroactive interference. According to the MFC framework, this suggests that integrative encoding established associations between changed responses and the experience of changes, thereby marking the order of responses.

Experiment 2

In Experiment 1, participants detected more changes when background scenes repeated across lists. Based on prior work showing that manipulations affect change detection and study-phase retrievals similarly (e.g., Wahlheim & Jacoby, 2013), we assumed that this increase in change detection reflected an increase in study-phase retrievals of B responses during List 2. However, change detection can occur without study-phase retrievals, as when participants sense a combination of cue familiarity and response novelty without perceiving what exactly changed. To verify that the context reinstatement benefits to change detection (and memory updating more generally) in Experiment 1 largely reflected effects on study-phase retrievals, Experiment 2 replaced the change detection measure in List 2 with an overt measure of B-response recall.

We predicted that repeating background scenes would increase study-phase retrievals and replicate the patterns of memory improvement for all other downstream measures. Importantly, the study-phase retrieval measure increased the precision of the conditional analyses by allowing us to examine the downstream consequences of verified instances when B responses were coactivated with A-D pairs. We expected that final B-response recall would be highly successful following earlier study-phase retrievals due to the established benefits of retrieval practice in similar tasks (Wahlheim et al., 2023). Moreover, we expected those effects to be enhanced when participants also remembered changes, because doing so would suggest that integration had occurred, which should create a qualitatively superior mnemonic representation. This conditional effect may be more pronounced when D responses are recalled (Negley et al., 2018); however, this additional benefit has not consistently occurred (Garlitch & Wahlheim, 2020).

Method

Participants

The participants were 74 members of the UNC Chapel Hill community recruited from the university's Psychology and Neuroscience Department participant pool ($n = 73$) and flyers posted around the campus and local community ($n = 1$). Participants received partial course credit or $10 as compensation, depending on the method of recruitment. Two participants were excluded from analysis due to experimenter error. The final sample size was 72 (53 women), ages 18–23 years ($M = 19.30$, $SD = 1.20$). The stopping rule was the same as in Experiment 1.

Design and materials

The design and materials were identical to Experiment 1.

Procedure

The List 1 and test procedures were the same as in Experiment 1, but the List 2 procedure was modified to directly measure study-phase retrievals. Before List 2, participants were again told that their first task would be to indicate whether scenes repeated from List 1. They were told that their next task would be to learn the word pairs in association with scenes and to consider how those pairs related to List 1 pairs. They were told that their final task would be to recall response words from List 1 after changed pairs appeared in List 2. On each List 2 trial, after making a recognition judgment, participants studied a centered word pair for 6 s (instead of 4 s as in Experiment 1). We increased the study duration here because the low rates of D-response recall and intrusions in Experiment 1 suggested that more time was needed to fully encode List 2 pairs. For A-D pairs, the scene and word pair were replaced by the cue that just appeared paired with a question mark, prompting participants to type the B response from List 1 (see Fig. 1). For C-D pairs, the trial advanced after study; participants did not attempt B-response recall.

Results and discussion

List 2 study phase: Scene recognition

We compared the probability of "old" responses on recognition trials in List 2 (Fig. 2, middle) using a model with fixed effects of context and item type. Replicating Experiment 1, a significant effect of context type, $\chi^2(1) = 1075.13$, p

$< .001$, $OR = 279.95$ (CI [199.88, 392.11]), showed more "old" responses for matched than mismatched scenes, thus verifying that repeating scenes cued retrieval of List 1 contexts. No other effects were significant, largest $\chi^2(1) = 0.35$, $p = .55$, $OR = 1.18$ (CI [0.68, 2.03]).

List 2 study phase: Cued recall

We compared B-response recall in List 2 for A-B, A-D items (Fig. 3B) to assess differences in study-phase retrievals. A model with a fixed effect of context type indicated a significant effect, $\chi^2(1) = 75.08$, $p < .001$, $OR = 2.34$ (CI [1.93, 2.84]), showing higher recall for matched than mismatched scenes. This difference confirms the assumption we made in Experiment 1 that reinstating scenes from List 1 promoted study-phase retrievals in List 2.

Test phase: Cued recall and memory for change

Figure 4 (middle panels) displays the response probabilities for each cued recall measure, analyzed using the modeling approach from Experiment 1. The model for B-response recall (Fig. 4A, middle) indicated significant effects of context, $\chi^2(1) = 37.83$, $p < .001$, $OR = 1.53$ (CI [1.34, 1.76]), and item type, $\chi^2(1) = 15.53$, $p < .001$, $OR = 1.60$ (CI [1.32, 1.93]), as well as a significant interaction, $\chi^2(1) = 8.19$, $p = .004$, $OR = 1.49$ (CI [1.13, 1.95]). Pairwise comparisons indicated significantly higher recall for A-B, A-D than A-B, C-D items in the match condition, z ratio = 4.82, $p < .001$, $OR = 1.60$ (CI [1.32, 1.93]) and no significant difference between item types in the mismatch condition, z ratio = 0.72, $p = .47$, $OR = 1.07$ (CI [0.89, 1.30]). Additionally, recall for A-B, A-D items was significantly higher in the match than the mismatch condition, z ratio = 6.38, $p < .001$, $OR = 1.87$ (CI [1.54, 2.26]) and the same was true for A-B, C-D items, z ratio = 2.32, $p = .02$, $OR = 1.26$ (CI [1.04, 1.52]). Finally, recall was significantly higher for A-B, A-D items in the match condition than A-B, C-D items in the mismatch condition, z ratio = 7.10, $p < .001$, $OR = 2.01$ (CI [1.56, 2.58]). These results show that context reinstatement again led to retroactive facilitation.

The model for D-response intrusions (Fig. 4B, middle), including only context type as a fixed effect, indicated a significantly lower probability in the match than the mismatch condition, $\chi^2(1) = 7.88$, $p < .001$, $OR = 1.85$ (CI [1.20, 2.85]). This replicates Experiment 1, and more convincingly suggests that repeating scenes enabled study-phase retrievals, which supported the later rejection of intrusions.

The model for "changed" classifications (Fig. 4C, middle) indicated a significant effect of item type, $\chi^2(1) = 997.15$, $p < .001$, $OR = 18.30$ (CI [15.28, 21.92]), showing higher probabilities for correct than incorrect classifications. There was also a significant effect of context type, $\chi^2(1) = 29.97$, $p < .001$, $OR = 1.47$ (CI [1.24, 1.74]), that was qualified by a significant interaction, $\chi^2(1) = 5.82$, $p = .02$, $OR = 1.51$ (CI [1.08, 2.12]). Pairwise comparisons indicated significantly higher correct classifications in the match than the mismatch condition, z ratio = 5.85, $p < .001$, $OR = 1.80$ (CI [1.48, 2.20]), and no significant difference in incorrect classifications between contexts, z ratio = 1.27, $p = .20$, $OR = 1.19$ (CI [0.91, 1.57]). These findings indicate that context reinstatement promoted study-phase retrievals that increased awareness of changes and made them more memorable at test.

The model for D-response recall after participants reported remembering that responses had changed (Fig. 4D, middle), which included only a fixed effect of context type, indicated a significant effect, $\chi^2(1) = 8.74$, $p = .003$, $OR = 1.44$ (CI [1.13, 1.84]), showing higher recall in the match than the mismatch condition, as in Experiment 1. This result suggests that reinstating contexts promoted study-phase retrievals that increased opportunities for integrative encoding.

Test phase: Cued recall conditionalized on memory for change

Similar to the approach from Experiment 1, we conditionalized B-response recall on study-phase retrieval in List 2 and "changed" classification at test for A-B, A-D items (Fig. 6). The MFC framework assumes that study-phase retrievals in List 2 enhance B-response recall via retrieval practice effects and by enabling co-activation of the B and D response that engenders encoding of their relationship, including that they changed (Wahlheim & Jacoby, 2013). In the following conditional analyses, we do not discuss context effects redundant with those above.

To determine if there was an overall study-phase retrieval benefit, we first compared B-response recall at test conditionalized on B-response recall during List 2. A model with B-response recall and context type as fixed effects unsurprisingly indicated a significant effect, $\chi^2(1) = 607.26$, $p < .001$, $OR = 171.12$ (CI [113.68, 257.60]). B responses recalled during List 2 were much more likely to be recalled at test than those that were not initially recalled. There was no significant interaction with context type, $\chi^2(1) = 0.62$, $p = .43$, $OR = 1.32$ (CI [0.66, 2.63]).

To determine if encoding information about changes when practicing B-response recall during study further enhanced subsequent B-response recall at test, we compared such recall across the three types of change remembrance (i.e., classifications) only for items that were recalled in List 2. A model with context type and classification as fixed effects showed a significant effect of classification, $\chi^2(2) = 49.15$, $p < .001$, and no significant interaction with context type, $\chi^2(2) = 0.52$, $p = .77$. This result indicated that

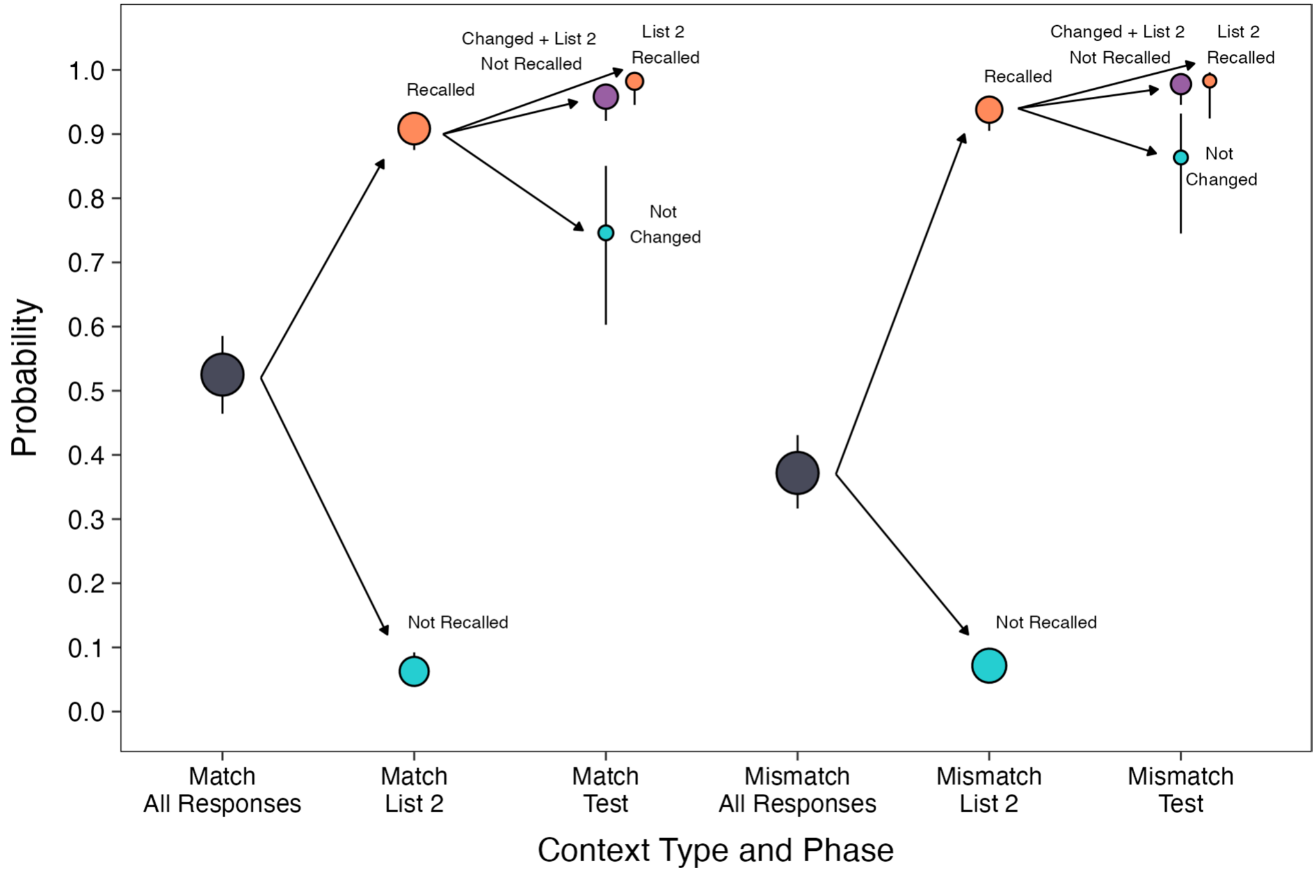

Fig. 6 Conditional recall of B responses from List 1 for A-B, A-D items: Experiment 2. Points are estimated probabilities from a mixed-effects model. Point areas reflect the relative observation counts. The aggregate values from Fig. 4 appear here in dark gray (Match and Mismatch All Responses). Recall conditioned on List 2 B-response recollection appear in orange and teal (Match and Mismatch List 2). Recall for recollected B responses conditioned on "changed" classifications at test appear in orange, purple, and teal (Match and Mismatch Test). Error bars are 95% confidence intervals. (Color figure online)

B-response recall was significantly higher when changes were remembered than when changes were not remembered, smallest z ratio = 4.89, $p < .001$, $OR = 13.11$ (CI [3.82, 44.95]), suggesting that encoding change attributes resulted in more elaborated representations. However, when changes were remembered, B-response recall did not differ depending on whether D responses were also recalled, z ratio = 1.10, $p = .52$, $OR = 1.77$ (CI [0.52, 6.05]). Collectively, these findings converge with Experiment 1 to suggest that study-phase retrievals promote subsequent recall via retrieval practice and that integrated memories including attributes of change are superior. However, we did not find the strongest possible evidence that integration promoted recollection-based retrieval of *how* responses changed.

Summary

Experiment 2 provided additional support for the MFC framework prediction that context reinstatement will increase study-phase retrievals and enhance recall of B and D responses. However, we did not show the recall dependency from Experiment 1 in which B-response recall on the final test was highest when D responses were also recalled. This discrepancy likely reflects differences in which processes comprise each conditional cell. Here, conditional B-response recall on the final test was examined only for items that evoked B-response recall during List 2 study. This subsetting left little room for memory enhancement associated with D-response recall over items remembered as changed without D-response recall, given their near ceiling performance (Fig. 6). Nevertheless, superior recall associated with the general preservation of change attributes indicates that earlier-retrieved B-responses were compared with D-responses in List 2, and that aspect at least was encoded. Thus, these results suggest that context reinstatement facilitated study-phase retrievals and integrative encoding, leading to qualitatively different memories that included changes with or without D responses.

Experiment 3

Experiments 1 and 2 included measures that assessed study-phase retrievals and their downstream consequences for episodic memory updating. One caveat of asking participants to judge whether pairs changed (Experiment 1) or type responses from the prior list (Experiment 2) is that we also diverted participants' attention away from encoding the List 2 pairs. This is similar to how participants engage in rehearsal borrowing in item-method directed forgetting (Bjork, 1970) and should diminish recall of the second responses. Moreover, including overt measures of study-phase retrieval may have also induced an increase in those retrievals. We conducted Experiment 3 to mitigate these caveats and determine whether context reinstatement would promote retroactive facilitation effects without overt retrievals. Specifically, in List 2, we only asked participants to make recognition decisions about scenes before the word pairs appeared to verify that the scenes induced context reinstatement. After those judgments, participants studied word pairs without any additional task. Results replicating our first two experiments would provide convincing evidence that the patterns implicating integrative encoding in earlier work (Cox et al., 2021) reflected the encoding of context-cued study-phase retrievals, while still allowing us to assess dependencies in retrieval at test along with awareness of changes. Given that overt List 2 retrieval measures have been shown to produce little if any reactivity (Wahlheim & Jacoby, 2013), we predicted that we would largely replicate the cued recall test response patterns from Experiments 1 and 2.

Method

Participants

The participants were 74 members of the UNC Chapel Hill community recruited from the university's Psychology and Neuroscience Department participant pool ($n = 64$) or flyers posted around the campus and local community ($n = 10$). Participants received course credit or $10, depending on the method of recruitment. Two participants were excluded from analysis due to computer errors that led to incomplete data files. The final sample included 72 participants (48 women) ages 18–29 years ($M = 19.60$, $SD = 1.70$).

Design, materials, and procedure

The design, materials, and procedure were the same as in Experiments 1 and 2, with one exception. We removed the measure during List 2 which previously asked participants to detect change (Experiment 1) or recall the B response (Experiment 2). Before List 2, participants were again told that their first task would be to indicate whether scenes repeated from List 1. They were also told that some pairs would change whereas others would be new and that their task would be to learn the pairs in association with the scenes and consider how they related to the pairs in List 1. On each List 2 trial, participants first viewed a scene and made a recognition decision before 3 s elapsed, then studied a word pair that appeared centered for 6 s.

Results and discussion

List 2 study phase: Scene recognition

We compared the probability of "old" responses on recognition trials in List 2 as in the prior experiments (Fig. 2, right panel). A significant effect of context type, $\chi^2(1) = 1287.13$, $p < .001$, $OR = 93.89$ (CI [73.26, 120.34]), showed more "old" responses for matched than mismatched scenes, verifying that repeating scenes cued List 1 retrieval. No other effects were significant, largest $\chi^2(1) = 1.11$, $p = .29$, $OR = 1.12$ (CI [0.90, 1.39]).

Test phase: Cued recall and memory for change

Figure 4 (right panels) displays the response probabilities for each cued recall measure, analyzed using the modeling approach from Experiments 1 and 2. The model for B-response recall (Fig. 4A, right) indicated a significant effect of context type, $\chi^2(1) = 40.25$, $p < .001$, $OR = 2.17$ (CI [1.77, 2.67]), no significant effect of item type, $\chi^2(1) = 0.98$, $p = .32$, $OR = 1.28$ (CI [1.04, 1.58]), and a significant interaction, $\chi^2(1) = 16.62$, $p < .001$, $OR = 1.83$ (CI [1.37, 2.44]). Pairwise comparisons indicated significantly higher recall for A-B, A-D than A-B, C-D items in the match condition, z ratio = 3.52, $p < .001$, $OR = 1.43$ (CI [1.17, 1.74]), and significantly lower recall for A-B, A-D than A-B, C-D items in the mismatch conditions, z ratio = 2.28, $p = .02$, $OR = 1.28$ (CI [1.04, 1.58]). Additionally, recall for A-B, A-D items was significantly higher in the match than the mismatch condition, z ratio = 7.36, $p < .001$, $OR = 2.17$ (CI [1.77, 2.67]), but recall for A-B, C-D items was not significantly different between context types, z ratio = 1.66, $p = .10$, $OR = 1.19$ (CI [0.97, 1.46]). Finally, recall was significantly higher for A-B, A-D items in the match condition than A-B, C-D in the mismatch condition, z ratio = 5.15, $p < .001$, $OR = 1.70$ (CI [1.30, 2.21]). Like Experiments 1 and 2, these results show that context reinstatement led to retroactive facilitation. However, in contrast to those experiments, disrupting context reinstatement by changing scenes across lists impaired recall enough to produce retroactive interference. Taken with the prior experiments, the finding of retroactive interference here suggests that the absence of such interference in Experiments 1 and 2 resulted from the List 2 measures encouraging a more global study-phase retrieval task set .

The model for D-response intrusions (Fig. 4B, right), including only context type as a fixed effect, indicated a significantly lower probability in the match than the mismatch condition, $\chi^2(1) = 7.86$, $p < .001$, $OR = 1.56$ (CI [1.14, 2.13]). Like Experiments 1 and 2, this suggests that context reinstatement promoted study-phase retrievals that aided intrusion rejections.

The model for "changed" classifications (Fig. 4C, right) indicated a significant effect of item type, $\chi^2(1) = 263.18$, $p < .001$, $OR = 3.44$ (CI [2.95, 4.00]), showing higher probabilities for correct than incorrect classifications. There was also a significant effect of context type, $\chi^2(1) = 90.35$, $p < .001$, $OR = 1.97$ (CI [1.69, 2.29]), that was qualified by a significant interaction, $\chi^2(1) = 10.48$, $p = .001$, $OR = 1.64$ (CI [1.21, 2.21]). Pairwise comparisons indicated significantly higher classifications in the match than mismatch condition. This difference was greater for A-B, A-D items, z ratio = 9.35, $p < .001$, $OR = 2.52$ (CI [2.07, 3.05]), than for A-B, C-D items, z ratio = 3.71, $p < .001$, $OR = 1.54$ (CI [1.22, 1.93]). This pattern replicates the prior experiments, except for the difference in the A-B, C-D condition.

The model for D-response recall after participants reported remembering that responses had changed (Fig. 4D, right), which included only a fixed effect of context type, indicated a significant effect, $\chi^2(1) = 40.84$, $p < .001$, $OR = 2.72$ (CI [2.00, 3.70]), showing higher recall in the match than the mismatch condition, as before. This result suggests that reinstating context again promoted study-phase retrievals that increased opportunities for integrative encoding.

Test phase: Cued recall conditionalized on memory for change

To assess whether memory for changes at test were associated with improved B-response recall even without overtly measuring study-phase retrievals, we compared recall across the three types of change remembrance in the classification variable at test in the A-B, A-D conditions (see Fig. 7). A model with context type and classification as fixed effects showed significant effects of context type, $\chi^2(1) = 21.72$, $p < .001$, $OR = 2.00$ (CI [1.59, 2.52]), and classification, $\chi^2(2) = 95.84$, $p < .001$, and a significant interaction $\chi^2(2) = 21.42$, $p < .001$. Pairwise comparisons in the match condition showed that recall was significantly different between each classification level, smallest z ratio = 3.34, $p = .002$, $OR = 2.00$ (CI [1.23, 3.25]). Conversely, in the mismatch condition, recall was significantly higher when changes were remembered and D responses were not recalled than when changes were not remembered, z ratio = 2.66, $p = .02$, $OR = 1.42$ (CI [1.04, 1.93]). No other comparisons were significantly different, largest z ratio = 1.93, $p = .13$, $OR = 1.64$ (CI [0.90, 3.00]). Finally, the magnitude of the recall benefits associated with remembering changes, whether or not D responses were recalled, were greater in the match than the mismatch condition, as shown by higher recall associated with such classifications in the match than the mismatch condition, smallest z ratio = 4.09, $p < .001$, $OR = 3.53$ (CI [1.93, 6.46]), and no significant difference between context types when changes were not remembered, z ratio = 1.18, $p = .24$, $OR = 1.12$ (CI [0.93, 1.34]). These results again implicate a role for integrative encoding of changes in retroactive facilitation and suggest that such encoding is supported by matching contexts. Notably, this mimics the conditional results from Cox et al. (2021)—when neither change awareness nor study-phase retrievals were assessed in List 2—showing dependence between B and D response recall only when scenes matched between study lists.

Summary

Experiment 3 showed that the memory benefits conferred by context reinstatement in the prior experiments occurred when participants were simply instructed to intentionally encode. However, the overall test response patterns here differed in two ways. First, retroactive interference was observed when background scenes were mismatched across lists, likely because the absence of List 2 measures led to fewer study-phase retrievals overall. This replicates the pattern reported by Cox et al. (2021), who also did not overtly measure study-phase retrievals. Second, when scenes matched across lists, "changed" classification false alarm rates for A-B, C-D items were slightly but significantly higher than when scenes mismatched, for reasons unbeknownst to us. Finally, the conditional B-response recall pattern for A-B, A-D items which accounted for change remembrances at test showed that the benefit associated with remembering change and recalling the D response was greater in the match than mismatch condition. A similar retrieval dependency was observed by Cox et al. (2021) and suggests that integrative encoding was better supported when the experimental conditions scaffolded List 1 context reinstatement.

General discussion

We examined the roles of study-phase retrievals and awareness of changes at study and test in retroactive facilitation induced by reinstating perceptual background contexts. Across three A-B, A-D paradigms, we overtly measured change detection (Experiment 1) and B-response recall (Experiment 2) during new learning and verified that memory consequences of context reinstatement could be obtained

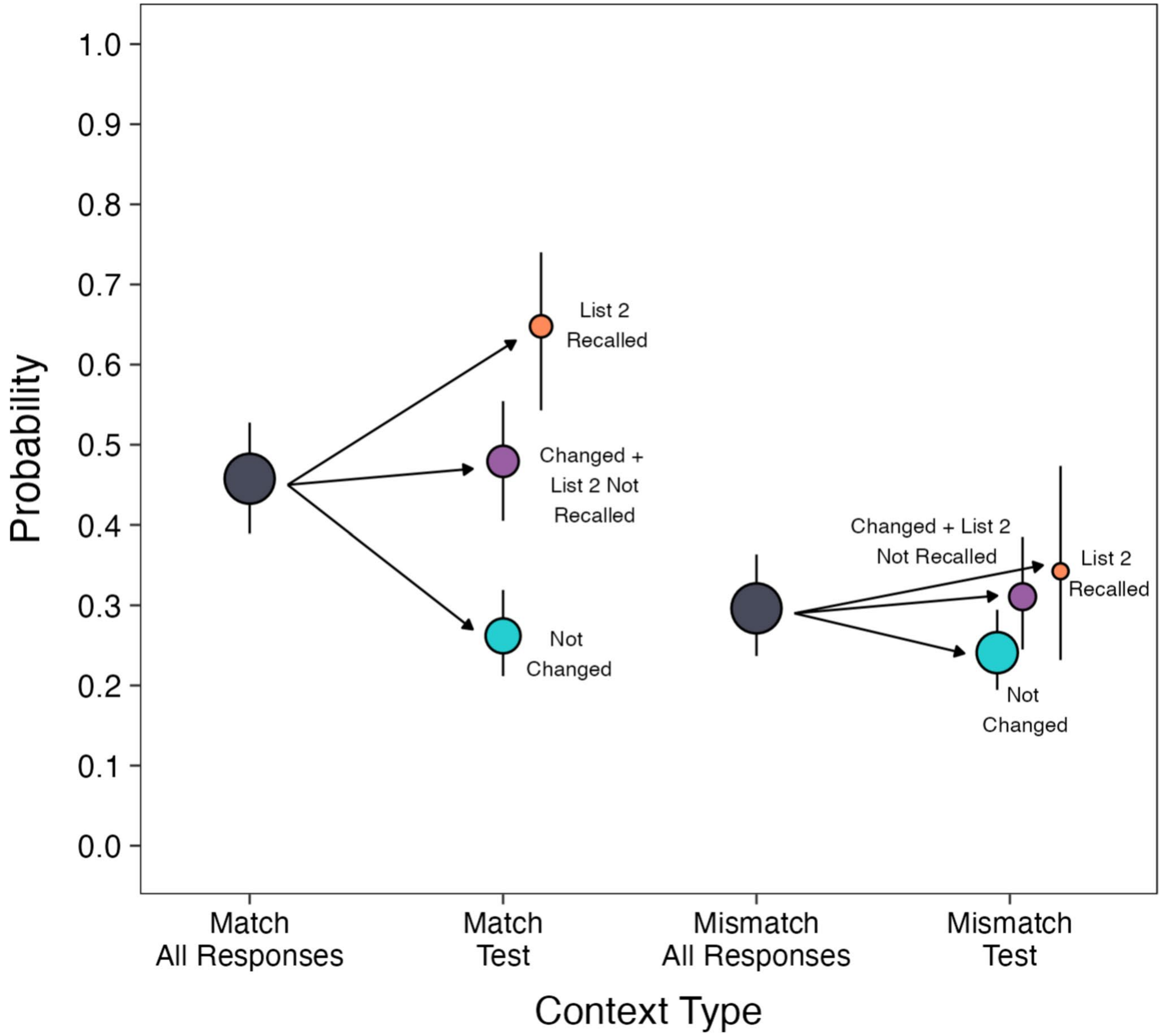

Fig. 7 Conditional recall of B responses from List 1 for A-B, A-D items: Experiment 3. Points are estimated probabilities from a mixed-effects model. Point areas reflect the relative observation counts. The aggregate values from Fig. 4 appear here in dark gray (Match and Mismatch All Responses). Recall conditioned on "changed" classification at test appear in orange, purple, and teal (Match and Mismatch Test). Error bars are 95% confidence intervals. (Color figure online)

without these measures (Experiment 3). Based on prior findings (Cox et al., 2021; Negley et al., 2018), we hypothesized that repeating background scenes would evoke more study-phase retrievals during new learning, thus promoting the encoding of and subsequent memory for changes. This outcome would be consistent with integration theory, generally, and the MFC framework (Wahlheim & Jacoby, 2013), specifically. We found that repeating background scenes increased change detection and study-phase retrievals, suggesting that that context reinstatement supported associative encoding of responses and that they changed. Moreover, context reinstatement increased B-response recall, "changed" classifications, and D-response recall at test, showing that context reinstatement enhanced episodic memory updating. Finally, conditional analyses of B-response recall revealed dependencies suggesting that detecting changes by comparing A-D pairs with retrieved B responses promoted integrative encoding: B-response recall was consistently highest when changes were remembered at test, especially when participants earlier detected changes or reported study-phase retrievals. Collectively, these findings verified that context reinstatement promotes study-phase retrievals and change awareness that leads to integrative encoding, which supports memory updating.

The present findings add to a growing body of work showing that the encoding conditions that promote study-phase retrievals can produce retroactive facilitation in situations that could otherwise produce retroactive interference. For example, enhancement of B-response recall in retroactive memory paradigms has been shown when participants were instructed to look back in memory for A-B items during A-D study (Jacoby et al., 2015, Experiment 1), longer study durations were provided for A-D pairs (Garlitch & Wahlheim, 2020; Negley et al., 2018), A-B pairs received retrieval practice with feedback prior to A-D study (Wahlheim et al., 2023), and semantic associations between A-D and A-B pairs were stronger (Antony et al., 2022). These findings are also consistent with classic studies showing retroactive facilitation under conditions where study-phase retrievals and change awareness were likely (cf. Bruce & Weaver, 1973; Robbins & Bray, 1974a, b). Taken together with other studies that used converging assays of the contact made between related episodes based on retrieval dependencies (Yu et al., 2025), our findings suggest that integrative encoding contributed to retroactive facilitation over and above the benefits of retrieval practice, as proposed by the MFC framework.

Although the retrieval dependences in conditional B-response recall here implicate roles for integration and retrieval practice, the dependencies varied based on the conditionalization method. Experiment 1 showed that for detected changes, B-response recall was higher when changes were remembered, especially when D responses were also recalled. Experiment 2 showed that for verified study-phase retrievals, B-response recall was higher when changes were remembered, but this advantage did not vary with D response recall success. Experiment 3 showed that when not measuring study-phase retrievals, B-response recall was higher when changes were remembered, especially when D responses were also recalled, as in Experiment 1. The discrepancies across experiments replicate differences in patterns across earlier studies (e.g., Garlitch & Wahlheim, 2020; Negley et al., 2018) and raise a theoretical issue about the characteristics of the memory representations that can be inferred. It has been argued that the most successful integrative encoding should result in recollection of changes, defined as being able to remember that a change occurred, and recall the alternative response, defined as what the target event changed to or from (Negley et al., 2018; Wahlheim et al., 2019). However, as mentioned above, remembering changes without recalling the alternative response can still indicate that participants integratively encoded the changed experience, even if some of the content was lost. The present discrepancies can be interpreted as showing that the lost content will be more likely to influence final recall when the conditionalization method does not first constrain the observations to successful study-phase retrievals. This constraint resulted in ceiling effects in B-response recall in Experiment 2 that limited the sensitivity to detect recall differences across classifications. Such differences, if they exist, may be detectable using a task procedure with a longer retention interval that brings B-response recall down from the ceiling.

The MFC framework also posits that memory for an item's order of occurrence, which is a type of source memory (i.e., items in List 1 appeared before items List 2), will be best preserved when the original encoding operations supported recollection-based retrieval of changes. Our measure of intrusion rates of D responses could be considered one assay of source (or temporal) memory failure. All three experiments clearly showed that fewer intrusions occurred with repeated as compared with different scenes, indicating better source/temporal memory following context reinstatement. Taken with the increased remembering of changes promoted by context reinstatement, these intrusion differences support the MFC framework proposal that study-phase retrievals counteract interference by engendering integrative encoding that supports memory for responses and their relative order.

By overtly measuring change detection and study-phase retrieval here, we also changed the way that participants approached the task. The consequences of these measures are apparent when comparing overall D-response recall patterns across the current experiments. The absence of retroactive interference in overall D-response recall when background scenes mismatched in Experiments 1 and 2—a finding that deviates from prior work (Cox et al., 2021)—suggests that requiring participants to reflect on how List 2 pairs related to List 1 pairs promoted controlled study-phase retrievals beyond the effects of context reinstatement. Conversely, the finding of retroactive interference in overall D-response recall when scenes mismatched in Experiment 3—one that replicates previous work (Cox et al., 2021)—suggests that participants were less likely to engage controlled study-phase retrievals when doing so was unsupported by perceived contexts. This discrepancy across experiments may seem surprising at first given that other paradigms with similar dual study-list structures did not show such reactivity when the presence of an overt change detection measure was manipulated between subjects (Jacoby et al., 2013). However, the overt judgments reflecting study-phase retrievals and the subjective experiences created by relatively long exposure durations in Experiments 1 and 2 here may have instead been more similar to looking back procedures that manipulated the use of controlled study-phase retrievals (Jacoby, 1974; Jacoby et al., 2015; Jacoby & Wahlheim, 2013).

The degree to which the present procedures required looking back may have modulated neural mechanisms that determine whether participants were in encoding or retrieval states. Prior work suggests that encoding and retrieval cannot be engaged simultaneously to the same extent (Duncan et al., 2012; Patil & Duncan, 2018) because they are supported by overlapping neural mechanisms (Long & Kuhl, 2019). Moreover, retrieval states triggered either by task instructions or stimulus-driven signals may linger for seconds (Duncan et al., 2012). Here, reinstating context and prompting scene recognition in List 2 may have induced retrieval states that lingered into the presentation of word pairs, thereby heightening sensitivity to stimulus-driven signals (e.g., repeated cues). Moreover, as suggested above, the List 2 measures in Experiments 1 and 2 may have generalized retrieval states across all items, including those with mismatched background scenes. However, mismatched background scenes should have still undermined the rapid initiation of a retrieval state, especially in Experiment 3, which included the least encouragement to look back. Although this neurocognitive account is plausible, the clearest evidence for mnemonic brain states comes from neuroimaging data and neurocomputational models (e.g., Long & Kuhl, 2019; O'Reilly & McClelland, 1994). Therefore, the present results can at best provide targets for work using neurocomputational methods to characterize these mechanisms.

Limitations

Two key limitations merit consideration. First, a major theoretical claim here is that retrieval practice and integrative encoding both contributed to enhance memory updating. These separate contributions were inferred from conditional analyses showing that successful study-phase retrievals led to enhanced B-response recall, with such enhancement being greater when accompanied by change awareness. However, the experimental paradigms did not allow us to disentangle the extent to which the acts of study-phase retrievals and detecting changes during study separately supported these effects. These processes could potentially be isolated using multivariate neuroimaging approaches that decode activation into cognitive processes using machine learning (e.g., Chanales et al., 2019) and could be augmented using computational modeling that separates item content from contextual features, which may include changes (e.g., Lohnas et al., 2015). Second, we chose the present word stimuli because the use of these materials in similar paradigms suggested their suitability for assessing context-cued integration and interference effects here. We chose the background scenes based on prior research which suggested that such scenes could cue successful recognition when repeated. We pseudo-randomly paired words with scenes without attempting to control for semantic relatedness of words and scenes or the degree to which particular combinations were imageable. We assumed that idiosyncratic effects would filter into the error variance given our counterbalancing scheme, and all models included items as random intercept effects to control for variance created by inherent differences in item memorability.

Future directions

We showed that reinstating background scene contexts enhances episodic memory updating, but it is unclear how these effects generalize across different context types. Context can take many forms. For example, temporal context can be manipulated by changing the time between study phases (e.g., Underwood & Ekstrand, 1966; Underwood & Freund, 1968). Internal states, such as thoughts, emotions, or mood can be manipulated using induction techniques (e.g., Kiley & Parks, 2022; Macht et al., 1977). Finally, physical and mental contexts can be manipulated by changing testing environments and the extent to which thoughts are congruent with tasks occurring in those environments (e.g., Sahakyan & Kelley, 2002; Smith, 1979). Future research could catalogue how variations in the degree to which these and potentially other types of context congruence induce change awareness that promotes memory updating.

We used single-shot learning phases and short delays among phases to prevent multiple instances of change detection and study-phase retrievals that would complicate conditional analyses. However, the prior study showing that reinstatement of background scenes promotes retroactive facilitation used muti-trial learning phases that allowed participants to reach a performance criterion, and separated each phase by 24 hours (Cox et al., 2021). Despite these differences, both studies showed near identical major results and provided support for integration over interference theory. However, reconsolidation theory is also relevant here (for a review, see Schroyens et al., 2023). Reconsolidation theory uniquely proposes a role for post-study-phase-retrieval memory malleability that determines updating success. To evaluate this mechanism, studies must include a long enough delay between learning phases for memories to become consolidated then reconsolidated after reactivation. Cox et al. (2021) discussed this theory because their design included 24-hour retention intervals and concluded that reconsolidation theory could not explain their findings of retroactive facilitation. Indeed, other perspectives from that literature have challenged the reconsolidation account and argue that integration is a more viable alternative (Gisquet-Verrier & Riccio, 2018). We could not adjudicate between integration and reconsolidation accounts here because of the short delays among phases. Future studies using creative designs with longer delays and precise overt List 2 measures may accomplish this.

Conclusion

The present experiments examined the roles for study-phase retrievals and change awareness in the facilitative effects of perceptual context reinstatement on episodic memory updating. Our findings are compatible with the variant of integration theory proposed by Larry Jacoby and colleagues (Jacoby & Wahlheim, 2013; Wahlheim & Jacoby, 2013) in supporting the proposal that reinstating context promotes study-phase retrievals that engender the encoding of changes leading to memory representations with qualities that counteract retroactive interference and promote retroactive facilitation. According to this MFC framework, reinstating context aided in the establishment of cross-episode associations and subsequent recollection-based retrieval of items and change attributes that allows participants to infer the sources (i.e., lists) of items. These findings further suggest that the assumption from classic interference theory that interference is best mitigated by differentiating contexts needs to be updated to consider how uniting contexts can accomplish the same goal by another means.

Author note This article is based on data presented as a poster at the North Carolina Cognition Conference (2024) and the Context and Episodic Memory Symposium (2024). We thank Hua Bai, Macy Faw, Eshan Singh, Qingyang Sun, Xiangling Chen, Abhay Chukkaluri, and Aliyaa Pathan for their assistance with data collection. We thank Drs. Aslihan Imamoglu and Min Sung Seo for helpful comments and insightful discussions.

Author contributions **Jennifer L. Fiedler:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing: Original Draft, Writing: Review & Editing, Visualization, Project administration.

Kelly S. Giovanello: Conceptualization, Methodology, Investigation, Validation, Resources, Writing: Review & Editing, Supervision, Project administration.

Robert W. Wiley: Validation, Investigation, Writing: Review & Editing, Visualization, Supervision.

Christopher N. Wahlheim: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing: Review & Editing, Visualization, Supervision, Project administration.

Funding The authors did not receive support from any organization for the submitted work.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Ethics approval This study was approved by the Institutional Review Board at the University of North Carolina at Chapel Hill (IRB number 22-2244).

Consent to participate Written informed consent was obtained from all individual participants included in the study.

Consent to publish Participants signed informed consent regarding publishing their deidentified data.

References

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 237–313). Academic Press.

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

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

Bilodeau, I. M., & Schlosberg, H. (1951). Similarity in stimulating conditions as a variable in retroactive inhibition. *Journal of Experimental Psychology, 41*(3), 199–204. https://doi.org/10.1037/h0056809

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

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.

Chanales, A. J. H., Dudukovic, N. M., Richter, F. R., & Kuhl, B. A. (2019). Interference between overlapping memories is predicted by neural states during learning. *Nature Communications, 10*(1), 5363. https://doi.org/10.1038/s41467-019-13377-x

Cox, W. R., Dobbelaar, S., Meeter, M., Kindt, M., & van Ast, V. A. (2021). Episodic memory enhancement versus impairment is determined by contextual similarity across events. *Proceedings of the National Academy of Sciences of the United States of America, 118*(48), e2101509118. https://doi.org/10.1073/pnas.2101509118

Duncan, K., Sadanand, A., & Davachi, L. (2012). Memory's penumbra: Episodic memory decisions induce lingering mnemonic biases. *Science, 337*(6093), 485–487. https://doi.org/10.1126/science.1221936

Fox, J., & Weisberg, S. (2019). *An R companion to applied regressio* (3rd ed.). 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

Gisquet-Verrier, P., & Riccio, D. C. (2018). Memory integration: An alternative to the consolidation/reconsolidation hypothesis. *Progress in Neurobiology, 171*, 15–31. https://doi.org/10.1016/j.pneurobio.2018.10.002

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth." *Perspectives on Psychological Science, 6*(3), 253–271. https://doi.org/10.1177/1745691611406924

Hintzman, D. L., Summers, J. J., & Block, R. A. (1975). Spacing judgments as an index of study-phase retrieval. *Journal of Experimental Psychology: Human Learning and Memory, 104*(1), 31–40.

Jacoby, L. L. (1974). The role of mental contiguity in memory: Registration and retrieval effects. *Journal of Verbal Learning* *and Verbal Behavior, 13*(5), 483–496. https://doi.org/10.1016/S0022-5371(74)80001-0

Jacoby, L. L., & Wahlheim, C. N. (2013). On the importance of looking back: The role of recursive remindings in recency judgments and cued recall. *Memory & Cognition, 41*(5), 625–637. https://doi.org/10.3758/s13421-013-0298-5

Jacoby, L. L., Wahlheim, C. N., & Yonelinas, A. P. (2013). The role of detection and recollection of change in list discrimination. *Memory & Cognition, 41*(5), 638–649. https://doi.org/10.3758/s13421-013-0313-x

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

Kiley, C., & Parks, C. M. (2022). Mechanisms of memory updating: State dependency vs. reconsolidation. *Journal of Cognition, 5*(1), Article 7. https://doi.org/10.5334/joc.198

Koen, J. D., & Rugg, M. D. (2016). Memory reactivation predicts resistance to retroactive interference: Evidence from multivariate classification and pattern similarity analyses. *The Journal of Neuroscience, 36*(15), 4389–4399. https://doi.org/10.1523/JNEUROSCI.4099-15.2016

Kuhl, B. A., Shah, A. T., DuBrow, S., & Wagner, A. D. (2010). Resistance to forgetting associated with hippocampus-mediated reactivation during new learning. *Nature Neuroscience, 13*(4), 501–506. https://doi.org/10.1038/nn.2498

Lenth, R. V. (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.10.2. https://rvlenth.github.io/emmeans/

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

Long, N. M., & Kuhl, B. A. (2019). Decoding the tradeoff between encoding and retrieval to predict memory for overlapping events. *NeuroImage, 201*, Article 116001. https://doi.org/10.1016/j.neuroimage.2019.07.014

Macht, M. L., Spear, N. E., & Levis, D. J. (1977). State-dependent retention in humans induced by alterations in affective state. *Bulletin of the Psychonomic Society, 10*(5), 415–418. https://doi.org/10.3758/BF03329377

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 pack* (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.

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. 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

O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus, 4*(6), 661–682. https://doi.org/10.1002/hipo.450040605

Patil, A., & Duncan, K. (2018). Lingering cognitive states shape fundamental mnemonic abilities. *Psychological Science, 29*(1), 45–55. https://doi.org/10.1177/0956797617728592

Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory & Cognition, 1*(1), 19–40. https://doi.org/10.3758/BF03198064

R Core Team. (2025). *R Software*. R Foundation for Statistical Computing. https://www.R-project.org/

Robbins, D., & Bray, J. F. (1974). Repetition effects and retroactive facilitation: Immediate and delayed recall performance. *Bulletin of the Psychonomic Society, 3*(5), 347–349. https://doi.org/10.3758/BF03333492

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

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(6), 1064–1072.

Sareen, P., Ehinger, K. A., & Wolfe, J. M. (2016). CB database: A change blindness database for objects in natural indoor scenes. *Behavior Research Methods, 48*(4), 1343–1348. https://doi.org/10.3758/s13428-015-0640-x

Schroyens, N., Beckers, T., & Luyten, L. (2023). Appraising reconsolidation theory and its empirical validation. *Psychonomic Bulletin & Review, 30*(2), 450–463. https://doi.org/10.3758/s13423-022-02173-2

Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory, 5*(5), 460–471.

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8*(2), 203–220. https://doi.org/10.3758/BF03196157

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

Underwood, B. J., & Ekstrand, B. R. (1966). An analysis of some shortcomings in the interference theory of forgetting. *Psychological Review, 73*(6), 540–549. https://doi.org/10.1037/h0023883

Underwood, B. J., & Freund, J. S. (1968). Effect of temporal separation of two tasks on proactive inhibition. *Journal of Experimental Psychology, 78*(1), 50–54. https://doi.org/10.1037/h0026157

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., 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., 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

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.

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

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open practices statement The data and analysis scripts are available on the Open Science Framework (https://osf.io/2q7ry/) and Experiments 1 (https://osf.io/9sbwq) and 2 (https://osf.io/gtva3) were preregistered.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.