Adult Age Differences in Event Memory Updating: The Roles of Prior-Event Retrieval and Prediction

David Stawarczyk[1,2], Christopher N. Wahlheim[3], Jeffrey M. Zacks[1]

[1]Department of Psychological & Brain Sciences, Washington University in St. Louis

[2]Department of Psychology and Neuroscience of Cognition Research Unit, University of Liège, Liège, Belgium

[3] Department of Psychology, University of North Carolina at Greensboro

Author Note

This project was funded by NIH grant R21AG05231401. David Stawarczyk is currently supported by a Marie Curie Global Fellowship grant from the European Commission (Grant Agreement number: 798109). We thank Hannah Aaron, Nora Cahill, Ryan Kahle, Katie Muson, Madeleine Schroedel, Devin Werner, and Jon Wingens for assistance with data collection, transcribing, and coding.

Correspondence concerning this article should be addressed to David Stawarczyk, Department of Psychology, University of Liège, Quartier Agora, place des Orateurs 1, 4000 Liège 1, Belgium. Email: d.stawarczyk@uliege.be. Tel.: 0032 4 3663991.

The stimulus materials, anonymized data, and analysis scripts are available on the Open Science Framework at https://osf.io/5pn67. This study was registered on the Open Science Framework at https://doi.org/10.17605/OSF.IO/NX47R.

The ideas and data presented in this manuscript were previously presented in 2018 during a poster session at the 59th Annual Meeting of the Psychonomic Society in New-Orleans (LA).

Abstract

Remembering past events can lead to predictions of what is to come, and to experiencing prediction errors when things change. Previous research has shown enhanced memory updating for ongoing events that are inconsistent with predictions based on past experiences. According to the Event Memory Retrieval and Comparison (EMRC) Theory, such memory updating depends on the encoding of configural representations that bind retrieved features of the previous event, changed features, and the relationship between the two. We investigated potential age-related differences in these mechanisms by showing older and younger adults two movies of everyday activities. Activities in the second movie either repeated from the first movie or included changed endings. During the second movie, before activities ended, participants were instructed to predict the upcoming action based on the first movie. One week later, participants were instructed to recall activity endings from the second movie. For younger adults, having predicted endings consistent with the first movie before seeing changed endings was subsequently associated with better recall of these changed endings and recollection that activities had changed. Conversely, for older adults, making such predictions prior to changes was associated with intruding details from the first movie endings, and was less strongly associated with change recollection. Consistent with EMRC, these findings suggest that retrieval of relevant experiences when events change can trigger prediction errors that prompt associative encoding of existing memories and current perceptions. These mechanisms were less efficient in older adults, which may account for their poorer event memory updating than younger adults.

Keywords: action observation, change detection, event cognition, integrative encoding, memory updating

Public Significance Statement

When events change, such as when a physician’s office updates their check-in system, memories need to be updated to avoid confusion with the past. Our research suggests that, when updating memories, older adults may not benefit as much as younger adults from predictions based on past events because of difficulties in linking together the features of past and new changed events. This finding can inform interventions for age-related memory decline.

Adult Age Differences in Event Memory Updating: The Roles of Prior-Event Retrieval and Prediction

Memories of past experiences can guide future behaviors. Suppose when you visit a friend, they hang your coat in a closet near the kitchen. When you visit them later, walking by the closet reminds you of the previous visit, and you anticipate that the closet will contain clothes. If you find the closet full of food because your friend has converted it to a pantry, you may experience a discrepancy between your prediction and current perception—a prediction error. Prediction errors that result from memory-based expectations have been termed *mnemonic prediction errors* (Bein et al., 2020, 2021; Sinclair & Barense, 2019). When a previous memory leads to a prediction error, it would be adaptive to update your memory to prevent you from looking for your coat in the pantry before leaving. Older adults are sometimes impaired in their ability to update existing memories when events change (e.g., Stawarczyk et al., 2020; Wahlheim & Zacks, 2019), as in the example above. Although mechanisms underlying age-related event memory differences have been well studied (for reviews, see Rubin & Umanath, 2015; Zacks, 2020), the causes of such memory differences remain unclear. Some possibilities are that aging leads to changes in (a) the formation of predictions during comprehension, (b) the use of prior event retrieval to form those predictions, or (c) the registering of prediction errors when changes occur, or (d) the use of prediction errors to prompt memory updating.

Prominent theories of perception and cognition view prediction errors as a core mechanism underlying efficient learning and promoting adaptive behaviors across multiple domains ranging from basic sensory processing to more complex cognitive functions (e.g., Clark, 2013; Den Ouden et al., 2012; Friston, 2009; Henson & Gagnepain, 2010). Several theories supported by empirical evidence assign prediction errors a key role in memory updating (e.g., Exton-McGuinness et al., 2015; Greve et al., 2017; Kim et al., 2014; Sinclair & Barense, 2019;

van Kesteren et al., 2012). However, these accounts propose different mechanisms and sometimes focus on different aspects of memory functioning.

The Schema-Linked Interactions between Medial Prefrontal and Medial Temporal Lobe model (SLIMM; van Kesteren et al., 2012) focuses on predictions stemming from semantic memory. SLIMM proposes that new information that conflicts with established semantic knowledge–and therefore induces prediction errors–can lead to memory updating. This view has been tested by having participants learn arbitrary rules about the pairing of stimuli that are later violated (e.g., Greve et al., 2019; Kafkas & Montaldi, 2018), or using stimuli that contradict established schemas–for example, seeing a picture of a toaster on a bed rather than a kitchen counter (e.g., Quent et al., 2022). In line with SLIMM, these studies often show better memory for new stimuli that strongly conflict with the established rules or schemas. However, they generally do not assess how memory for the rules or schemas used to generate the predictions may be updated when prediction errors occur.

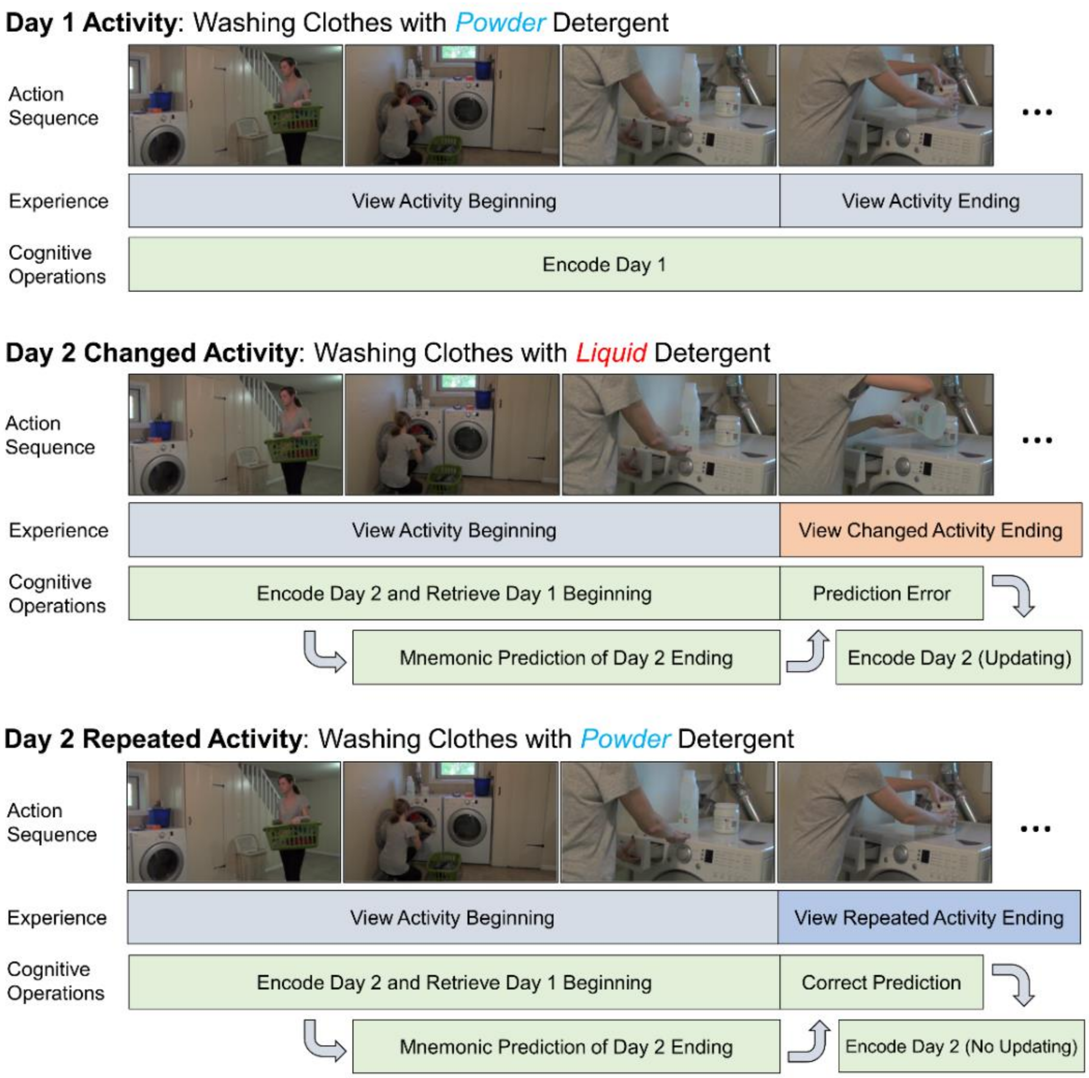
In the domain of episodic memory, two main mechanisms have been proposed to explain memory updating. *Pruning* accounts propose that prediction errors weaken memories for items that fail to appear in the expected context, thus weakening irrelevant memories (Kim et al., 2014). Evidence for this was shown when repeating picture sequences with a changed ending diminished recognition of the original ending relative to a control condition in which expectations were not violated (Kim et al., 2014, see also 2017). *Reconsolidation* accounts propose that prediction errors stimulate a period of lability in which reactivated memories are disrupted. This disruption may be followed by either (1) the preservation of the original memory when no new information is presented following the prediction error or (2) new learning that weakens the original memory in a way similar to the pruning account when new information is presented. This view has been supported by studies showing that inducing prediction errors by interrupting video clips before their expected outcomes led to more subsequent intrusions from clips showing related but different actions, but only when these new clips appeared after the interruptions. These increased intrusions did not occur when new related clips appeared before the interruptions, suggesting that memories for the interrupted outcomes were preserved rather than altered in this condition (Sinclair & Barense, 2018; see also, Sinclair et al., 2021; Sinclair & Barense, 2019). Thus, according to these views, efficient episodic memory updating following mnemonic prediction errors happens through a weakening of the memory of the specific previous experience that was used to generate the prediction, or by altering those memories based on new information.

In contrast to pruning and reconsolidation mechanisms, Event Memory Retrieval and Comparison theory (EMRC; Wahlheim & Zacks, 2019), proposes that prediction errors lead to the formation of new memory representations that incorporate the original prediction, the error, and the new information into an integrated whole. EMRC starts from the proposal that, when viewing ongoing activities, observers construct stable representations of what is happening now called *event models* (Zacks et al., 2007). Event models are one basis for observers to predict how events will unfold, notably by serving as cues to trigger the episodic retrieval of previous similar events (Hintzman, 2011). EMRC diverges from pruning or reconsolidation mechanisms in proposing that when a changed event leads to a prediction error, a new representation is formed that includes the prediction based on the previous experience (e.g., the closet contained clothes), the conflicting new information (e.g., the closet is now a pantry), and a trace of the conflict between what was predicted and what occurred. This conflict trace encodes temporal order information because the predicted features correspond to the earlier event, whereas the actual features correspond to the later event. This kind of *configural representation* can facilitate encoding the new information, reducing interference between the two experiences in memory, and judging the temporal order of the two experiences (Jacoby et al., 2015; Wahlheim & Jacoby, 2013). Forming a configural representation enriches the encoding of the second event, rendering it more distinct from the similar previous event—a process akin to pattern separation mechanisms proposed to be a function of the hippocampus (e.g., Yassa & Stark, 2011) and its broader interconnections with cortical structures (for a review, see Amer & Davachi, 2023). Note that pruning, reconsolidation, and the formation of a configural representation are not mutually exclusive operations; all three could potentially impact memory updating in different situations.

EMRC has been tested using an *everyday changes* paradigm (Wahlheim & Zacks, 2019). This paradigm adapts paired-associate learning designs (Jacoby et al., 2015; Wahlheim & Jacoby, 2013) to naturalistic materials, allowing semantic knowledge to play a role–as in real-life memory updating, and allowing investigation of the temporal dynamics of memory updating. Participants watch two movies of an actor performing everyday activities with each movie depicting a day in her life. Some activities repeat exactly across movies, others are new in the second movie, and a final set includes activity beginnings that repeat and endings that change across movies. Subsequent memory for activity features from the second movie is then tested. Participants also indicate if the ending features changed between movies, and if so, attempt to recall the original ending features. An example of changed endings across the two days/movies can be seen in an activity depicting the actor washing clothes (Figure 1). The actor approaches the washer identically on both days but washes the clothes with *powder detergent* on the first day and *liquid detergent* on the second day when the activity is changed.

Figure 1

Theoretical Processing Chain Leading to Event Memory Updating



Note. Schematic depicting the processing chain according to Event Memory Retrieval and Comparison (EMRC) Theory. Action sequences depict examples of everyday activities that include the same beginning and changed or repeated endings. *Experience* indicate viewers' subjective experiences. *Cognitive operations* indicate the relevant cognitive processes underlying the viewing experiences that lead to mnemonic predictions (i.e., predictions that are consistent with the Day 1 activity ending) and event memory updating. The individual whose face appears here gave signed consent for her likeness to be published in this article.

Consistent with EMRC, remembering both that the activities had changed and the ending features from the first movie (i.e., a pattern of retrieval referred to as *change recollection*) was associated with correct recall of changed features. This facilitation in memory updating was observed for both younger and older adults, but older adults recollected fewer changes on the final memory test. These results suggest that older adults were less able to encode and retrieve configural representations. However, the original *everyday changes* task did not examine whether participants generated predictions that were consistent with the first movie ending before viewing the changed endings in the second movie. This precluded inferences about age-related differences in how prediction errors relate to change recollection indicating the presence of configural representations, and more generally how such differences relate to established age-related differences, such as associative memory (e.g., Burton et al., 2019; Ebert & Anderson, 2009; Naveh-Benjamin, 2000) and feature binding (Mitchell & Johnson, 2009; for a meta-analyses, see Old & Naveh-Benjamin, 2008) that may affect configural representations.

We modified the *everyday changes* paradigm to examine the role of mnemonic prediction errors in event memory updating in younger adults. In one study, we assessed the relationship between predictive looking errors and memory updating using eye tracking (Wahlheim et al., 2022). Wahlheim and colleagues hypothesized that when an object was contacted during an action in the first movie, the beginning of that action in the second movie would induce viewers to retrieve the previous ending and to predict that the new event would end similarly—a mnemonic prediction. This would lead to looking ahead to the previously-contacted object. If the ending action changed, such looks would be evidence for mnemonic prediction errors. As predicted, looks ahead to previously-contacted objects were associated with higher rates of memory updating (i.e., accurate memory for changed features from the second movie). They were also associated with better change recollection (i.e., correctly remembering that the event had changed combined with remembering the original feature that had changed). In a related study, participants made overt predictions after each activity beginning in the second movie by choosing between two still shots depicting possible endings (Hermann et al., 2021). Choosing a picture of the original ending was evidence for making a mnemonic prediction. Such choices were associated with better memory updating and change recollection after event changes, consistent with looks ahead.

Finally, we examined age-related differences in the role of mnemonic predictions in event memory updating using a task suited for functional magnetic resonance imaging (Stawarczyk et al., 2020). The second movie was paused shortly after each activity began, and participants were instructed to mentally replay the relevant activity ending from the first movie during that time. Neural pattern reinstatement in the medial temporal lobes and posterior medial cortex during mental replay of first-movie endings was taken as evidence for memory retrieval that could enable mnemonic predictions. Such reinstatement predicted subsequent memory updating and change recollection in younger adults, but did so less for older adults. However, the extent that neural reinstatement indexes prediction errors remains to be precisely determined (cf. Bein et al., 2020); comprehenders could retrieve but fail to use the retrieved information to predict.

Thus, previous results in young adults suggest that retrieval of a relevant previous event can drive predictions about how related events will unfold, which are related to memory updating (Hermann et al., 2021; Wahlheim et al., 2022). They also suggest age differences in how neural reinstatement of previous events during new event encoding is associated with memory updating (Stawarczyk et al., 2020). However, an important open question is whether predictions that are consistent with previous events (i.e., mnemonic predictions) similarly benefit memory updating in younger and older adults when confronted with event changes that lead to prediction errors. The present study was designed to address this question using an overt prediction method.

The Present Study

Participants watched two movies in a first experimental session. The movies included the actor performing activity sequences on two separate days that we henceforth refer to as *Day 1* and *Day 2*. In the Day 2 movie, each activity paused after the beginning segment. Participants were instructed to predict aloud how the activity would end if the actor repeated her Day 1 actions (i.e., to make a mnemonic prediction). This procedure improves on previous procedures (e.g., Hermann et al., 2021; Stawarczyk et al., 2020; Zacks et al., 2011) by using an open-ended response procedure. For example, Hermann et al. (2021) asked participants to make a prediction by selecting from two pictures depicting the two possible endings. A limitation of that procedure is that It exposes participants to the alternate ending, which could contaminate memory. As in prior variants of the everyday changes paradigm, the activity endings in the Day 2 movie sometimes repeated from the Day 1 movie and other times included actions changes. These action changes in the Day 2 movie always directly followed the pauses where participants made predictions in order to induce the experience of prediction errors. After a week, participants were instructed to recall Day 2 endings, indicate if they had changed, and if so, recall Day 1 endings.

This paradigm allowed us to test four pairs of hypotheses about the processes underlying age differences in event memory updating based on prior findings and EMRC:

- Older adults show generally poorer episodic recall than younger adults (for a review, see Craik, 2022). Older adults should therefore make fewer mnemonic predictions. Thus, when asked to predict the end of an activity during Day 2 viewing, they should be less likely to produce a response consistent with the Day 1 ending of that activity (Hypothesis 1a) and should show overall poorer final recall of Day 2 activity endings (Hypothesis 1b) than younger adults.
- EMRC assumes that retrieval of a relevant previous event ending enables encoding of a configural representation when the activity ending changes (Wahlheim & Zacks, 2019).

This leads to the hypothesis that viewers should have a better final recall of changed Day 2 endings for activities where they could make predictions that were consistent with their previous Day 1 movie experience (Hypothesis 2a). This effect should be reduced in older adults due to their poorer formation of configural representations (Hypothesis 2b).

- The assumption that retrieval of a relevant previous event enables encoding of a configural representation also entails that predicting an activity ending consistent with the Day 1 movie should be associated with better change recollection (i.e., remembering both that the activities had changed and the ending features from the first movie; Hypothesis 3a). Parallel to Hypothesis 2b, this should be reduced in older adults (Hypothesis 3b).
- EMRC proposes that mnemonic predictions lead to configural representations that support memory updating. If so, then memory for changed Day 2 movie features following predictions consistent with Day 1 activity endings should be improved when participants recollect change and impaired when they do not recollect change (Hypothesis 4a). This relationship should be weaker in older adults reflecting their weaker configural representations (Hypothesis 4b).

Importantly, hypotheses 2a and 2b contrast with the prediction of pruning (Kim et al., 2014) and reconsolidation (Sinclair & Barense, 2018) accounts. Those mechanisms should lead memory for Day 1 movie endings to be poorer following mnemonic prediction errors, and thus predict opposing effects. These hypotheses also contrast with interference accounts of episodic memory, which predict that more accessible Day 1 movie endings should lead to heightened competition at retrieval (e.g., Anderson & Neely, 1996).

Method

Transparency and Openness

We report how we determined our sample size and describe all data exclusions, manipulations, and all measures used in the study. The stimulus materials, anonymized data, and analysis scripts are available on the Open Science Framework (Stawarczyk et al. 2023). The research reported here was approved by the Institutional Review Board of Washington University in St. Louis with the study name "Change Detection in Healthy Aging and Early AD: Behavioral pilot testing" and protocol number 201705076. We pre-registered the design, hypotheses, and analyses. The links to the preregistration, materials, data, and analysis scripts appear in the Author Note.

Participants

The sample consisted of 44 younger adults (33 females; M_{Age} = 20.02 years, *SD* = 1.72, range: 18-26) from Washington University in St. Louis and 47 older adults (36 females; M_{Age} = 70.77 years, *SD* = 3.09, range: 65-76) from the St. Louis community[1] who participated in 2017 and 2018. There were 36 White (77%) and 10 Black (21%) older adult participants (one participant did not report their ethnicity), as well as 14 White (32%), 13 Black (30%), 15 Asian (34%), one Asian and Black (2%) and one Asian and White (2%) younger adult participants. Three younger adults and one older adult reported Hispanic ethnicity. All older adults scored 25 or above (*M* = 28.89, *SD* = 1.31, range: 25-30) on the Mini-Mental State Exam (MMSE; Folstein et al., 1975). Vocabulary scores from the synonym and antonym task of Salthouse (1993) were significantly higher for older (*M* = 14.95, *SD* = 4.83) than younger (*M* = 12.37, *SD* = 3.90) adults, $t(81) = 2.68$, $p = .009$, $d = 0.59$. Seven older adults and one younger adult were not given the

[1] In addition to these 91 participants, 13 others participated in the study but were not included in the final sample. Three were younger adults: one was excluded because of experimenter error leading to unusable data, one because English was not their native language (and they claimed to not understanding the memory questions), and one because they had already seen the movies before. Ten were older adults: two were excluded because of experimenter errors or computer issues leading to unusable data, two did not return for the second session, one reported having been recently diagnosed with vascular dementia, and five did not follow task instructions.

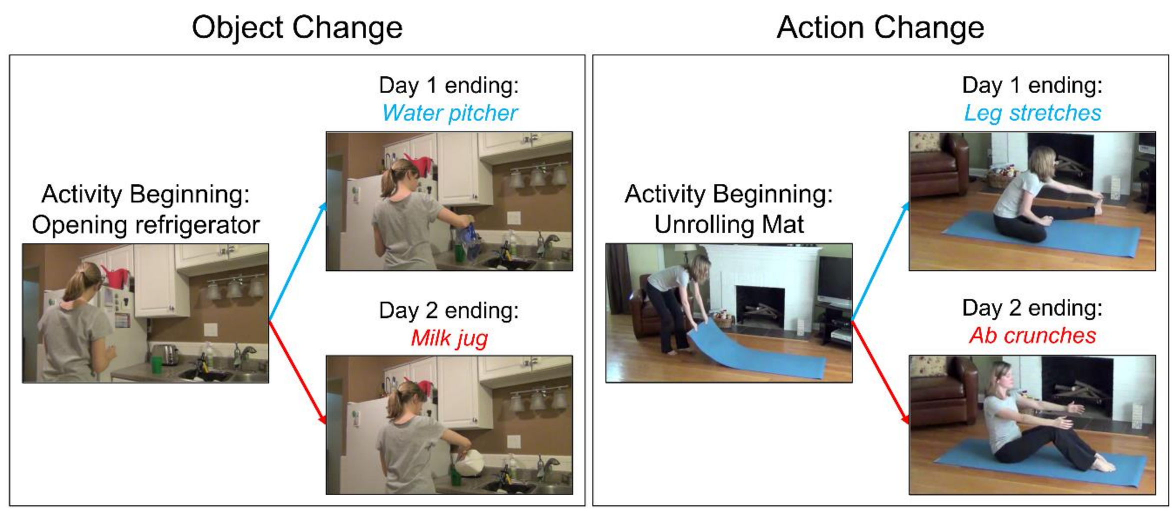
vocabulary test due to an oversight. Although there were age differences in vocabulary scores, this variable was not related to our outcome measures and so was not included as a covariate in our analyses. Years of education were also significantly higher for older (M = 16.57, SD = 2.56) than younger (M = 14.30, SD = 1.56) adults, $t(89) = 5.08$, $p < .001$, $d = 1.07$. This variable was not included as covariate in our analyses as it reflects younger adults being undergraduates that had not finished their studies rather than an absolute difference in education level between the two groups. For participating, younger adults received $10 per hour or course credit, and older adults received $10 per hour, plus $5 for travel reimbursement. We chose our sample size a priori based on a power analysis using G*Power 3 (Faul et al., 2007). Our goal was to achieve 80% power (alpha = .05) to detect a medium-to-large between-group effect size (d = 0.65) based on previous findings of age-related differences in the detection of changes across movies (Wahlheim & Zacks, 2019). Note that this sample size estimation should be considered an approximation, because we analyzed the data using mixed effect models for which G*Power 3 could not estimate power.

Materials

We showed movies of a woman performing everyday activities on two fictive days of her life. As in prior studies (e.g., Garlitch & Wahlheim, 2021; Stawarczyk et al., 2020; Wahlheim et al., 2022; Wahlheim & Zacks, 2019), we described the movies to the participants as *Day 1* and *Day 2,* to make clear that they were intended to depict two different days in the actor’s life. There were two versions (A and B) of each activity that differed on a thematically central feature. Each activity comprised a beginning segment that was identical in both versions followed by an ending segment including the changed feature. For some activities, the feature that changed between movies was an object that the actor contacted (e.g., pouring water or milk). For others, the action itself changed (e.g., that actor performed stretches or sit-ups on the same exercise mat). Figure 2 displays still shots from two example activities, one featuring an object change (left) and the other featuring an action change (right). The activities included an audio track but no dialogue (except when the actor interacted with her dog) because there were no other people in the movies.

Figure 2

Changed Activity Examples Showing Object and Action Changes



Note. The images above are still shots from movies in the everyday changes paradigm. These shots depict activities with endings that changed between days—one that included an object change (left panel) and another that included an action change (right panel). The individual whose face appears here gave signed consent for her likeness to be published in this article

The complete material set comprised 45 activity pairs. All the activities were common everyday tasks rather than more age-specific activities (e.g., setting up a video game system) that may be difficult for older adults to encode because they lack relevant knowledge (Smith et al., 2020). The complete list of activities is provided in SM (Table S1). The critical manipulation was whether the activity endings changed or repeated between movies (see Figure 1). Given our interest in how people process activity changes, we included twice as many changed activities (30) as repeated activities (15). For counterbalancing, the 45 activities were divided into 3 groups of 15 activities and rotated through conditions, such that each participant viewed two groups of

15 changed activities and one group of 15 repeated activities. We also alternated, across participants, whether the A or B version of each activity appeared in the first movie. This arrangement produced six experimental formats. The mean duration of complete activities was 22.9 s (SD = 7.8, range: 10.8 - 47.6) and did not differ between the A (M = 22.9 s, SD = 8.0) and B (M = 22.8 s, SD = 7.6) versions of activities, $t(44) = 0.28$, $p = .78$, $d < 0.01$. The total duration of the movies was between 18 and 20 minutes depending on the experimental format. All movies appeared in 960 × 540 resolution in the center of the screen on a 21.5-in monitor with a 1920 × 1080 resolution. The participants sat approximately 60 cm away from the monitor. Stimuli were presented using E-Prime 2 software (Psychology Software Tools, Pittsburgh, PA, USA; Schneider et al., 2012).

Procedure

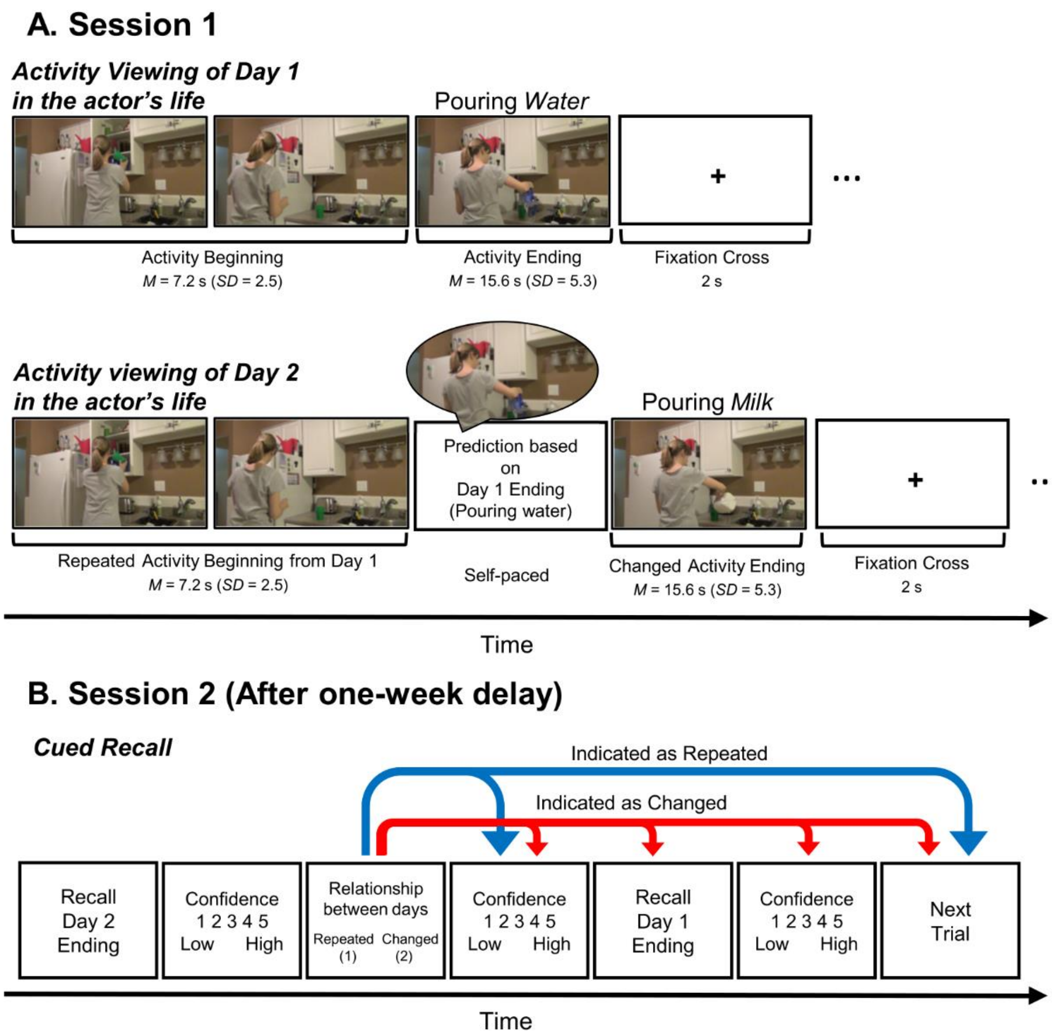
Figure 3 displays a schematic of the procedure. Participants completed the experiment in two sessions separated by one week.

Session 1: Movie viewing. During Session 1, participants watched the two movies of the actor performing everyday activities (panel A). Both the Day 1 and Day 2 movies were presented during Session 1. As mentioned above, the terms “Day 1” and “Day 2” refer to the different days in the movies on which actions were performed, not to the days on which participants watched the movies. Before viewing the Day 1 movie, we informed participants that they would see an actor performing a series of daily activities throughout the course of a day and that they should pay attention to each activity. During viewing, each activity appeared without interruption; from the participant’s viewpoint each activity was a continuous single narrative event. A fixation cross appeared in the middle of the screen for 2 s after each activity; this demarcation was necessary for the prediction task in the subsequent Day 2 movie. These fixation crosses were similar to cuts between scenes in commercial movies and no specific instructions about them were given to the participants. The activities appeared in a fixed random sequence such that each third of the task (15 activities) contained five repeated items and no more than six consecutive changed activities. Two example activities separated by a fixation cross appeared before the Day 1 movie.

The Day 2 movie appeared immediately after the Day 1 movie in the first session. Before the Day 2 movie began, we told participants that activities would appear in the same order as in the Day 1 movie, but that the movie would stop intermittently so that they could predict the upcoming actions. We instructed participants to predict the exact actions from the relevant activity in the Day 1 movie. After each activity beginning, the movie stopped, and participants made a verbal prediction aloud about the ending action. An experimenter recorded predictions with an audio recorder. Participants were encouraged to provide detailed descriptions. There was no time limit for each prediction. Participants then pressed the space bar to resume viewing and were told to think about whether the next actions repeated or changed from Day 1. These instructions were meant to induce mnemonic prediction errors when the predicted and changed endings did not match. A fixation cross appeared for 2 s between each complete activity. Prior to the test trials for the Day 2 movie, participants made predictions for the two example activities. To create naturalistic viewing conditions and to reduce the risk of ceiling performance in younger adults, we did not tell participants during Session 1 that we would administer memory tests during Session 2. We did this to allow participants to encode the movies as they would naturally and to prevent the use of external aids (e.g., taking notes) during the week between sessions.

Figure 3

Schematic Depiciting a Changed Activity and a Prediction Consistent with the Day 1 Ending (Mnemonic Prediction)



Note. (A) During the first session, participants watched two consecutive movies depiciting everyday activities on two days in the actor's life. The *changed* activity example above shows the actor pouring water on Day 1 and milk on Day 2. On Day 2, the movie stopped just before the activity ending, and participants predicted the ending based on memory for the Day 1 ending. Predicting that the actor would pour water would have led to a prediction error when the actor instead poured milk next. (B) After one week, during the second session, participants completed a cued recall test. Participants were instructed to recall the Day 2 activity endings, indicate whether activities had changed across movies, and if so, recall the Day 1 activity endings. Participants made confidence ratings for all memory responses. There were no time limits to respond.

Session 2: Memory testing. During Session 2, one week later, participants first completed a cued recall test (Figure 3B) for activities from Session 1. The test cues always queried features that varied between the two activity versions (the cues and correct responses appear in Supplemental Material [SM], Table S1). The test items appeared in the same order as the activities in the Day 2 movie. Participants were instructed to first recall the Day 2 ending action feature by typing their response. For example, for the question, "What did the actor eat for breakfast?", participants could have typed, "A banana." After each response, participants indicated if the activity had repeated (1) or changed (2) by pressing the corresponding keys. When participants indicated that the activity had changed, they were prompted to type the Day 1 ending action feature. Most participants typed their own responses using the keyboard, but the experimenter typed for some older adults. Participants rated the confidence in each response on a scale from 1 (Low) to 5 (High). Analyses of confidence ratings are provided in SM6.

After the cued recall test, participants completed a two-alternative forced choice (2AFC) recognition memory test (SM4). Age-related memory differences are typically larger on recall than recognition tests (Rhodes et al., 2019). We therefore wanted to examine whether predicting action endings based on prior relevant experiences differentially benefits memory for changed action endings for younger and older adults in a recognition task. The results from cued recall and 2AFC recognition tests were very similar with respect to our four set of hypotheses. For brevity, we provide detailed analyses for cued recall and summarize the 2AFC recognition results below. A complete description of the 2AFC recognition results is available as SM4. After 2AFC recognition, participants completed a vocabulary test (Salthouse, 1993), and then only older adults completed the Mini Mental State Exam (Folstein et al., 1975).

Response Coding

The first author and a rater blind to conditions coded the open-ended responses. Predictions during movie pauses before Day 2 activity endings were *Day 1 consistent* when they included relevant features from Day 1 endings. The prediction codes for the other responses appear in SM2. In addition, responses for cued recall of Day 2 features were coded in four ways. *Correct Day 2 recalls* included Day 2 ending features. *Day 1 intrusions* included Day 1 ending features. *No responses* were self-evident but also included phrases such as “I don’t remember.” *Incorrect* responses included inter-activity and extra-experimental features. Finally, responses for cued recall of Day 1 features were coded in three ways. *Correct Day 1 recall* included Day 1 ending features. *No Responses* and *Incorrect* responses were analogous to Day 2 cued recall. For cued recall of Day 2 and Day 1 features, we only report analyses of responses relevant to hypothesis testing (i.e., Correct Day 2 recall; Day 1 intrusions; Correct Day 1 recall). We report analyses of other responses in SM3.1. The interrater agreements (*Cohen’s* κ = .92 [predictions], .89 [Day 2 cued recall], and .90 [Day 1 cued recall]) were near perfect (Landis & Koch, 1977).

Statistical Methods

We performed all analyses using R software (R Core Team, 2021). Predictions and cued recall responses were analyzed with logistic mixed effect models using the glmer function from the lme4 package (Bates et al., 2015). Mixed effects models are preferred to standard regression models for several reasons, most notably because mixed effect models allow random effects of both participants and items within a single model (for a review on this topic, see Hoffman & Walters, 2022). In the present study, we defined the random effect structure of the model following the procedure proposed by Matuschek and colleagues (2017) that balances the need to prevent false positives with the reduction of statistical power resulting from more complex models. We first attempted to fit the maximal model (with random intercepts of participants and activities and random slopes for each relevant fixed effect and their interactions) and gradually decreased the complexity of the random effect structure in case of convergence or singularity issues until a fit occurred. We then continued to use this backward-selection method to reduce the models until further reduction of the random effect structure led to a significant decrease in the goodness-of-fit (tested with a likelihood ratio test and a liberal threshold of alpha = .20; Matuschek et al., 2017). We report the final model for each analysis in SM (Table S4). The fixed effects for each model are specified below. Hypothesis tests were performed using the default setting of the Anova function from the car package (Type-II ANOVA with Wald test; Fox & Weisberg, 2019). The emmeans (Lenth, 2019) package was used to conduct post hoc comparisons using the Tukey method that controlled for multiple comparisons and to obtain model-estimated probabilities.

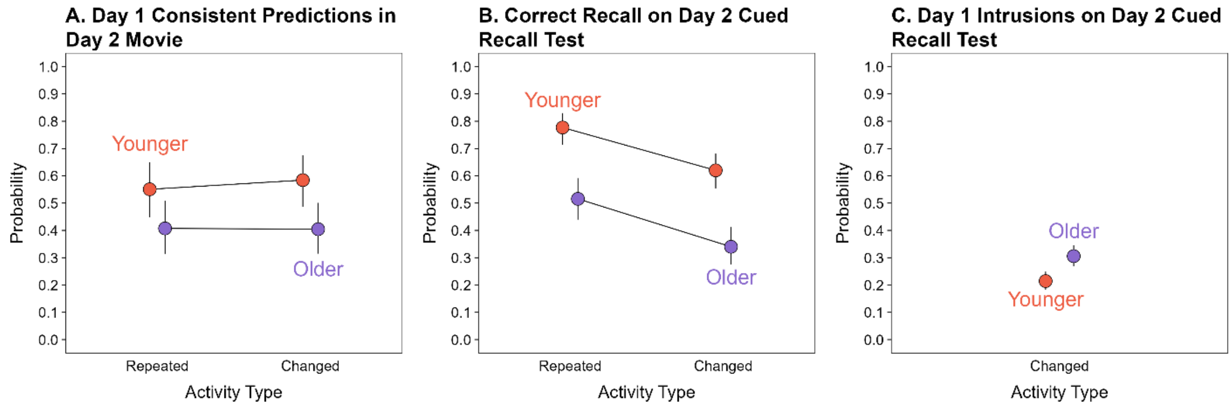
Results

Older Adults Made Fewer Day 1 Consistent Predictions

We first tested the hypothesis that, although all participants were instructed to use their memories for Day 1 to make predictions about Day 2 activity endings, older adults should make fewer Day 1 consistent predictions than younger adults (Hypothesis 1a). A 2 (Age: younger vs. older) × 2 (Activity: repeated vs. changed) model (Figure 4A; Model 1 in Table S4) indicated evidence for this hypothesized Age effect $\chi^2(1) = 18.11$, $p < .001$, no significant effect of Activity, $\chi^2(1) = 0.61$, $p = .44$, and no significant interaction, $\chi^2(1) = 0.94$, $p = .33$.

Figure 4

Day 1 Consistent Predictions and Cued Recall Responses



Note. The points are estimated probabilities from mixed effects models. The error bars are 95% confidence intervals.Panel A shows that older adults made significantly fewer predictions that were consistent with the Day 1 movie than younger adults. Panels B and C show that older adults showed significantly poorer correct recall and experienced more intrusions from the Day 1 movie than younger adults when recalling activity features from the Day 2 movie.

Older Adults Showed Poorer Subsequent Recall of Activity Features

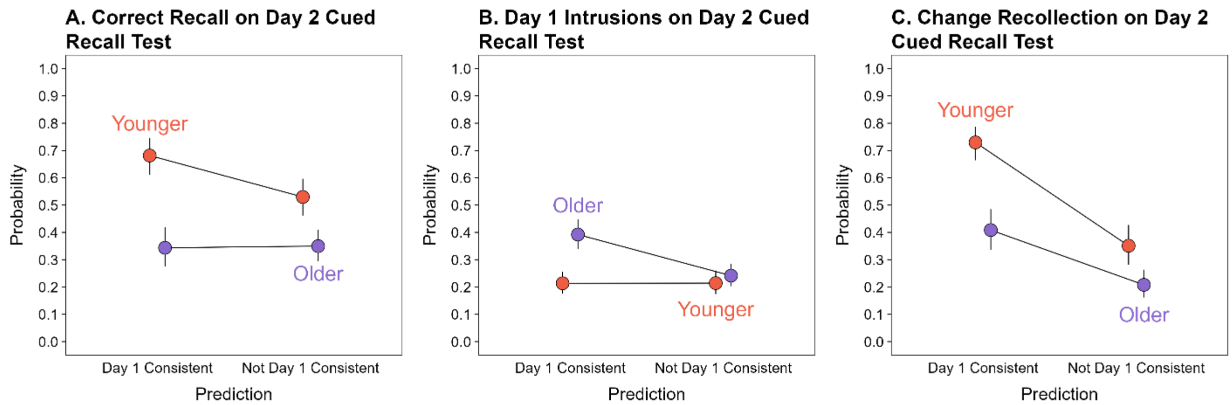
We next tested the hypothesis that older adults should recall fewer Day 2 features than younger adults (Hypothesis 1b) using a 2 (Age: younger vs. older) × 2 (Activity: repeated vs. changed) model for correct Day 2 recall (Figure 4B, Model 2 in Table S4). Older adults showed significantly lower recall than younger adults, $\chi^2(1) = 50.42$, $p < .001$, and recall was lower for changed than repeated features, $\chi^2(1) = 42.89$, $p < .001$. There was no interaction, $\chi^2(1) = 0.03$, $p = .86$. We examined age differences in Day 1 intrusions for only changed activities using a 2 (Age: younger vs. older) model (Figure 4C, Model 3 in Table S4). Older adults produced significantly more intrusions than younger adults, $\chi^2(1) = 18.67$, $p < .001$.

Only Younger Adults' Updating Benefitted from Day 1 Consistent Predictions

The main goal of the present study was to determine whether prediction error plays a role in age-related differences in event memory updating. We tested the hypotheses that correct Day 2 recall should be higher following Day 1 consistent predictions, and that this benefit should be greater for younger than older adults (Hypotheses 2a & 2b). For changed activities, we fitted separate 2 (Age: younger vs. older) × 2 (Prediction: Day 1 consistent vs. not Day 1 consistent) models to correct Day 2 recall and Day 1 intrusions (Models 4 and 5 in Table S4). The Day 2 recall model (Figure 5A) indicated a significant Prediction effect, $\chi^2(1) = 8.49$, $p = .004$, qualified by a significant interaction with Age, $\chi^2(1) = 12.29$, $p < .001$. Day 1 consistent predictions led to significantly higher Day 2 recall for only younger adults, z ratio = 4.53, $p < .001$ (older adults: z ratio = -0.20, $p = 0.84$). The opposite pattern emerged for Day 1 intrusions (Figure 5B): The model indicated a significant Prediction effect, $\chi^2(1) = 16.55$, $p < .001$, qualified by a significant interaction with Age, $\chi^2(1) = 14.84$, $p < .001$. Day 1 consistent predictions led to significantly higher intrusions for only older adults, z ratio = 5.59, $p < .001$ (younger adults: z ratio = -0.03, $p = 0.98$). These findings support the hypothesis that prediction errors should benefit memory updating less for older than younger adults. EMRC accounts for this age-related impairment by assuming that older adults formed configural representations less effectively when viewing changed endings. We evaluated this possibility further by examining change recollection and its association with Day 1 consistent predictions and Day 2 recall.

Figure 5

Cued Recall Responses for Changed Activities Conditioned on Prediction Type



Note. The points are estimated probabilities from mixed effects models. The error bars are 95% confidence intervals. Panels A and B show that, for younger adults, Day 1 consistent predictions while viewing the Day 2 movie were significantly associated at test with better recall of the changed activity features from the Day 2 movie and were not related to the experience of intrusions from the Day 1 movie. The opposite pattern was observed in older adults: Day 1 consistent predictions in this age group were associated at test with the experience of more intrusions from the Day 1 movie and were not related to recall of the changed activity features from the Day 2 movie. Panel C shows that Day 1 consistent predictions were significantly associated at test with better change recollection (i.e., recall that the activity changed across movie along with the Day 1 activity feature) for both age groups altough the effect was larger in younger than older adults.

Older Adults Showed Weaker Associations Between Day 1 Consistent Predictions and Change Recollection

If prediction errors prompt the formation of configural representations, leading to subsequent change recollection, then Day 1 consistent predictions should be associated with better change recollection (Hypothesis 3a). Moreover, if older adults have a weaker path from prediction errors to configural representations, then older adults should recollect fewer changes following Day 1 consistent predictions (Hypothesis 3b). Consistent with these hypotheses, a 2 (Age: younger vs. older) × 2 (Prediction: Day 1 consistent vs. not Day 1 consistent) model (see Figure 5C, Model 6 in Table S4) indicated significantly better change recollection for younger than older adults $\chi^2(1) = 41.90$, $p < .001$, and following Day 1 consistent predictions, $\chi^2(1) = 157.98$, $p < .001$. A significant interaction, $\chi^2(1) = 11.76$, $p < .001$, showed greater benefits of Day 1 consistent predictions on change recollection for younger adults, z ratio = 11.44, $p < .001$, than older adults, z ratio = 7.17, $p < .001$.

Older Adults Showed Weaker Associations Among Day 1 Consistent Predictions, Change Recollection, and Correct Day 2 Recall for Changed Activities

EMRC also predicts that age differences in the pathway from prediction errors to memory updating should depend on whether changes can be recollected. This leads to the hypothesis that correct Day 2 recall should be higher when predictions are consistent with Day 1 endings and changes are recollected (Hypothesis 4a). EMRC also proposes that age-related deficits in encoding configural representations should lead to weaker associations between change recollection and Day 2 recall for older than younger adults (Hypothesis 4b). We tested these hypotheses by examining age differences in correct Day 2 recall conditioned on whether predictions included Day 1 endings and if change was subsequently recollected (Figure 6, Models 7 & 8 in Table S4).

We tested for these differences using a 2 (Age: younger vs. older) × 2 (Prediction: Day 1 consistent vs. not Day 1 consistent) × 2 (Change Recollection: recollected vs. not recollected) model (Table 1, top rows). Recall was significantly higher when changes were recollected than when they were not (green vs. red points). Supporting Hypothesis 4a, a Prediction × Change Recollection interaction showed that when participants recollected change, recall was better when participants had earlier predicted the Day 1 ending, z ratio = 2.54, $p = .01$ (higher green points in the left than right panel); otherwise recall for changed activity features was significantly worse, z ratio = -6.12, $p < .001$ (lower red points in the left than right panel). Supporting Hypothesis 4b, an Age × Change Recollection interaction showed significantly higher recall for younger than older adults when changes were recollected, z ratio = 7.32, $p < .001$, and no significant age difference when changes were not recollected, z ratio = 1.87, $p = .06$.

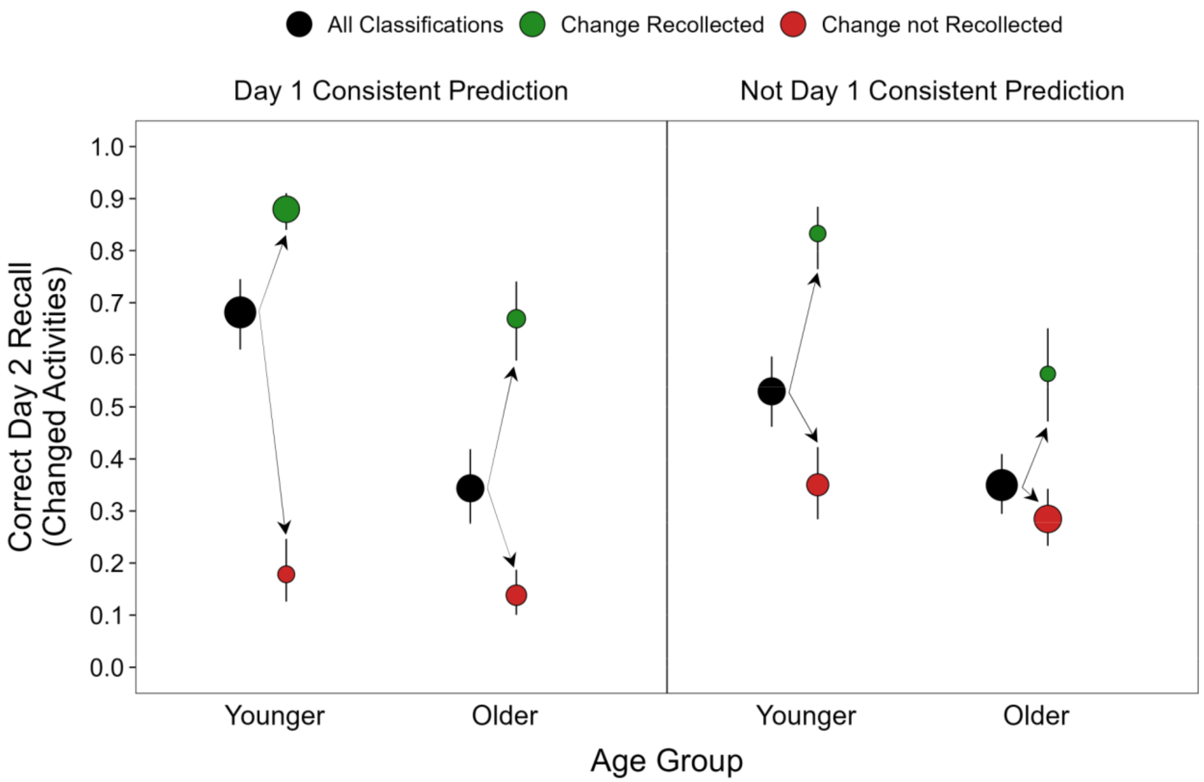
Table 1

Model Results for Cued Recall Measures Conditioned on Mnemonic Prediction and Memory for Change: Changed Activities Only

Measure	Effect	χ^2	*df*	*p*
Correct Day 2 recall	Age	31.33	1	< .001
	Prediction	8.27	1	< .01
	Change Recollection	420.92	1	< .001
	Age × Prediction	0.02	1	= .89
	Age × Change Recollection	22.24	1	< .001
	Prediction × Change Recollection	39.36	1	< .001
	Age × Prediction × Change Recollection	0.02	1	= .87
Day 1 intrusions	Age	0.68	1	= .41
	Prediction	84.62	1	< .001
	Change Recollection	225.53	1	< .001
	Age × Prediction	0.14	1	= .71
	Age × Change Recollection	16.26	1	< .001
	Prediction × Change Recollection	48.58	1	< .001
	Age × Prediction × Change Recollection	0.02	1	= .88

Figure 6

Day 2 Recall Conditioned on Predictions and Change Recollection for Changed Activities



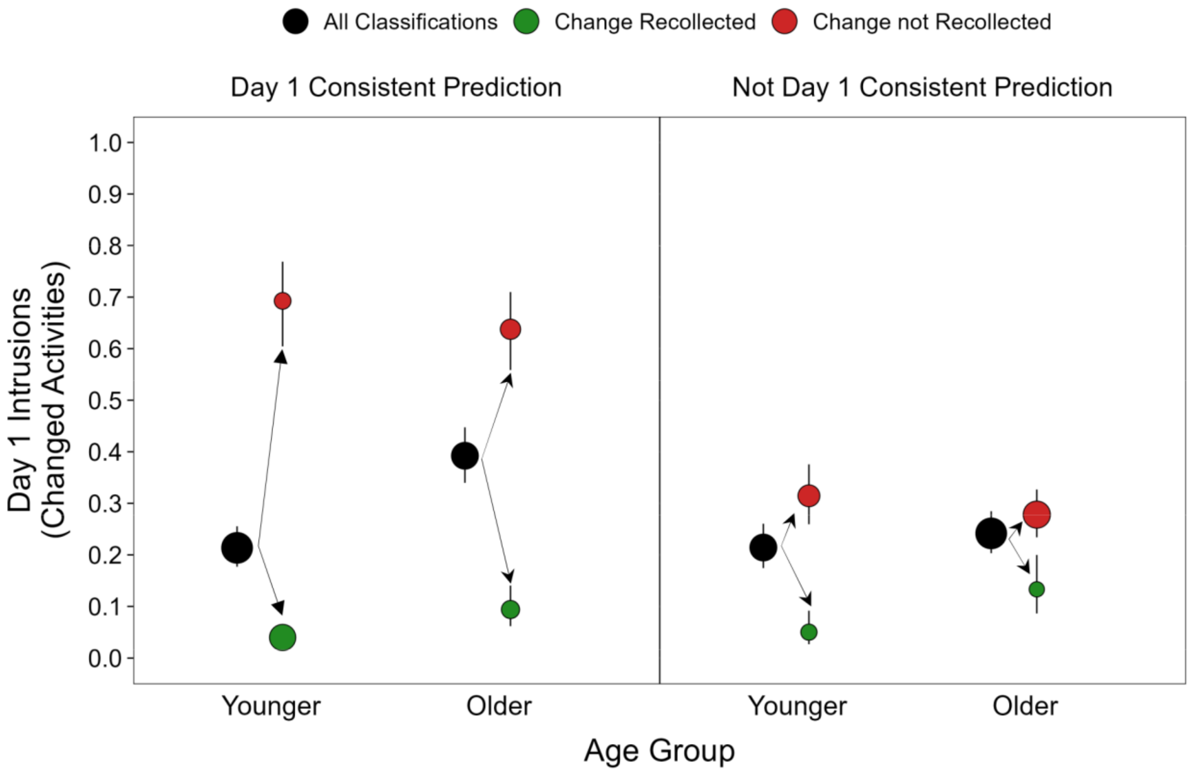
Note. The black points are model-estimated probabilities conditioned on prediction during Day 2 viewing and age group. These estimates are the same as those displayed in Figure 5A. The colored points are probabilities conditioned on both prediction and change recollection. The green points are changed activities that were identified as such followed by correct recall of Day 1 features (change recollected). The red points are changed activities that were identified as such followed by incorrect recall of Day 1 features or not identified as such (change not recollected). The error bars are 95% confidence intervals.

Finally, we also examined conditional Day 1 intrusions (Figure 7) using the same model as in the previous analysis (Table 1, bottom rows). We focus only on the effects relevant to hypothesis testing. The results supported Hypothesis 4a that Day 1 consistent predictions should lead to more intrusions when changes are not recollected. A Prediction × Change Recollection interaction showed significantly more intrusions for changes that were not recollected when prior predictions were consistent with Day 1 than when they were not (higher red points in the left than right panel), smallest *z* ratio = 11.04, $p < .001$. Conversely, intrusions did not depend on predictions when change was recollected (green points in left and right panels), *z* ratio = -1.24, $p = .21$. Change recollection was thus associated with a greater memory accuracy benefit when predictions were Day 1 consistent. An Age × Change Recollection interaction showed significantly fewer intrusions for younger than older adults for recollected changes, *z* ratio = 4, $p < .001$ (supporting Hypothesis 4b) and no significant age difference when change was not recollected, *z* ratio = -1.43, $p = .15$. Day 1 intrusions associated with change recollection occurred when Day 1 features were output twice for the same activity, which could reflect guessing or some other cause. These results nonetheless support the hypothesis that change recollection should be associated with reduced benefits to memory accuracy for older than younger adults.

Collectively, the analyses reported in this section indicate that if predictions lead to change recollection, this benefits memory for the changed features. But if predictions are not followed by the recollection of change, it leads to more intrusions. Older adults recollected change less often than younger adults following Day 1 consistent predictions (see Figure 5C) which explains why they overall only experienced more Day 1 intrusions following Day 1 consistent predictions (see Figure 5B and black points in Figure 7) whereas younger adults overall only made more correct recalls in the same situation (see Figure 5A and Black points in Figure 6).

Figure 7

Day 1 Intrusions Conditioned on Predictions and Change Recollection for Changed Activities



Note. The black points are model-estimated probabilities conditioned on prediction during Day 2 viewing and age group. These estimates are the same as those displayed in Figure 5B. The colored points are probabilities conditioned on both prediction and change recollection. The green points are changed activities that were identified as such followed by correct recall of Day 1 features (change recollected). The red points are changed activities that were identified as such followed by incorrect recall of Day 1 feature or not identified as such (change not recollected). The error bars are 95% confidence intervals and are obscured when the intervals are smaller than the point diameters.

Summary of Additional Preregistered Analyses

The 2AFC recognition memory results converged with the cued recall results on key measures that tested our hypotheses. The findings supported all except Hypothesis 2b, which stated that if older adults are less able to form configural representations following a prediction error, then they should show weaker associations between Day 1 consistent predictions and correct Day 2 recall. However, younger and older adults did not differ in how Day 1 consistent predictions benefited 2AFC recognition of changed Day 2 features (SM4). This could be because the environmental support provided by the recognition test reduced age-related differences (cf. Rhodes et al., 2019). Response time analyses for change classifications during cued recall and for all recognition responses revealed faster responses for correct answers following Day 1 consistent predictions (SM5). Finally, confidence rating analyses revealed no association between prediction and monitoring resolution, with the latter measuring the extent to which ratings distinguished correct from incorrect responses (SM6).

Discussion

This study examined how younger and older adults update their memories when confronted with changes in events. Viewers were induced to experience prediction errors by exposing them to activities with repeated beginnings but changed endings. If one sees a repeated beginning and anticipates a previously seen ending, but then a changed ending is presented, this creates a prediction error. Compared to younger adults, older adults were less likely to make predictions that were based on their memory for the previous ending, and they recalled fewer changed activities. Predicting endings consistent with the first movie before seeing changed endings was associated with better memory for changed features on a final recall test for younger adults. In contrast, making such predictions prior to changed endings was associated with more intrusions of the previous ending, and lower recollection rates that the activity had changed across movies for older adults. These results suggest that older adults used episodic memory to guide encoding of changed event details less effectively tan younger adults, which may have reflected fewer memory-based prediction errors prompting memory updating.

In both age groups, successfully retrieving and predicting the previous ending before seeing a changed ending was associated with contrasting final recall for the changed ending, depending on whether the participant later recollected the change. When participants recollected that endings had changed, predictions consistent with original endings were associated with better recall of changed features. However, when participants failed to recollect that the activity had changed, they intruded more of the original ending features. This suggests that retrieval of relevant previous events during the encoding of a change can protect against interference if it triggers the formation of a representations that includes the conflict between the two endings—a configural representation comprising the previous ending, the new ending, and the conflict. However, retrieval of the previous ending to generate predictions can *increase* interference if it does not trigger such memory updating, presumably because it boosts accessibility of the competing memory trace. Older adults recollected less changes following predictions that were consistent with the previous endings, thus supporting the proposal that older adults were less likely to use prediction errors to prompt memory updating.

These results generally support the predictions of EMRC. The theory proposes a processing chain including event model construction, event memory retrieval during ongoing perception, prediction errors following detected changes, and recollection of event changes (Wahlheim & Zacks, 2019). The model's most controversial proposal pertains to the fate of event memories reinstated prior to changed endings. EMRC proposes that event memories become more accessible when perceptual features trigger their retrieval (for a review, see Roediger & Butler, 2011). When a reactivated event memory used to generate a prediction is coactivated with perceived action changes, integrative encoding may occur, thus producing a configural representation (Hintzman, 2011; Wahlheim & Jacoby, 2013). This kind of memory updating serves to differentiate memories and reduce interference, in contrast to mechanisms such as pruning (Kim et al., 2014) and reconsolidation (Sinclair & Barense, 2019), which would lead to weakened or altered memory for the prior event that failed to appear when expected.

The present results do not constitute evidence that pruning or reconsolidation do not contribute to memory updating; these have been found to be important mechanisms of memory in other situations. For example, after learning sequences of faces and scenes, violating expectations with a different ending exemplar enhanced recognition for the changed exemplar and impaired recognition for the original exemplar when neural prediction strength was high (Kim et al., 2014). These results suggested that prediction errors stimulated pruning of existing memories. Relatedly, creating prediction errors by interrupting repeated movie clips before their outcomes has led to more intrusions from related movies (e.g., Sinclair et al., 2021; Sinclair & Barense, 2018). These findings suggest that reconsolidation modified reactivated memories (for a review, see Elsey et al., 2018). However, the present findings are incompatible with accounts proposing these mechanisms, because prediction errors were positively associated with recall of original and changed endings in younger adults. Both age groups also showed better memory for Day 1 features associated with prediction errors, which would not be produced by pruning or reconsolidation. Future studies should attempt to clarify when memory updating following prediction errors leads to weakened memories for the event used to generate the prediction vs integrative encoding of this prior event memory with the new changed event.

The present findings of age differences in retrieval of Day 1 activity endings are important, and are consistent with broad findings of age-related impairments in associative binding (for a review and meta-analysis, see Old & Naveh-Benjamin, 2008; see also Mitchell & Johnson, 2009). However, other age-related deficits that are well-established in the cognitive aging literature such as controlled attention (e.g., McCabe et al., 2010) or processing speed (e.g., Salthouse, 1993, 2000) might also contribute to age differences found here.

Such age differences may also reflect older adults' tendency to encode event gist, resulting in less precise memory for specific actions (for reviews, see Brainerd & Reyna, 2015;

Devitt & Schacter, 2016; Greene & Naveh-Benjamin, 2023). Here, most of the tested features were conceptually central to the activities, as opposed to being perceptual background features. For example, in an activity in which the actor stocked a refrigerator with drinks, the drinks chosen were different in the changed endings. Despite this, participants may have encoded the objects in terms of high-level gist and failed to register the difference between the drinks across movies. Older adults in particular focus more on gist in some situations (e.g., Greene & Naveh-Benjamin, 2022; Grilli & Sheldon, 2022). However, the fact that older adults were more likely to intrude the specific previous event feature (e.g., the specific drink) after viewing a changed ending argues against this possibility. Instead, it suggests that older adults encoded specific information, albeit weakly associated with its source.

When interpreting the age-relate updating differences seen here, one must also consider older adults' heightened dependence on self-relevance for supporting everyday memories (Hess, 2014; Hess & Emery, 2012). One could argue that the younger adult actor performed activities that were less self-relevant to older adults. Indeed, there were some activities of this kind, such as when the actor prepared a book bag for school. The actor was also closer in age to younger than older adult viewers. This may have induced a sense of personal distance from the movie content that diminished attention during encoding and subsequent memory updating. The roles of these self-relevance variables in memory updating could be examined by varying actor and participant ages or by providing self-referential encoding strategies (e.g., asking participants to judge if they would perform activities the same way as the actor; Gutchess et al., 2007; Hamami et al., 2011).

Although encoding precision and self-relevance may contribute to age-related event memory updating differences, we argue that such differences primarily reflect how the strength of prediction errors promotes integrative encoding and subsequent recollection. Several younger adult studies have shown greater memory updating following stronger prediction errors, such as when violating expectations about object sequences improves the ability to indicate that lures are similar but not identical to studied objects (Bein et al., 2021). However, this was only observed when the original ending object was correctly recognized as such (a kind of change recollection) and when original sequences were learned well enough to evoke strong prediction errors (also see Chen et al., 2015). Consistent with the idea that prediction strength modulates memory updating, stronger prediction errors have been shown to enhance associative memory for updated scene-face pairs (Greve et al., 2017) and memory for new information that updates knowledge (Brod et al., 2018). This enhancement may be characterized by improved recollection, such as when prediction errors selectively increase reports of context retrieval in associative recognition (Kafkas & Montaldi, 2018). These studies suggest that the present age differences partly reflected older adults' weaker reactivation of existing memories. This may have led to prediction errors that prompted less integrative encoding and subsequent recollection of changed actions.

We assessed mnemonic prediction errors here by explicitly asking participants to retrieve details of a previous event and use the details to predict the outcome of a related new event. Some studies have induced errors by other means, including interrupting movies before expected endings (Sinclair & Barense, 2019; Sinclair et al., 2021) or omitting and adding movie segments (Yazin et al., 2021) without explicitly asking participants to make predictions. Other studies varied object sequences and repetition frequency to vary expectations that were later violated with unexpected ending objects (e.g., Bein et al., 2020, 2021; Kim et al., 2014). Despite consistent reports of memory updating associated with prediction errors, the generalizability across approaches and regarding age differences remains unclear. Cross-task comparisons using identical measures could enable systematic investigations of the role of prediction strength.

More generally, the fundamental role of prediction in core neural operations suggests that prediction error processes may be shared across multiple domains (e.g., Clark, 2013; Den Ouden et al., 2012; Friston, 2009; Henson & Gagnepain, 2010). The prediction errors here are likely to differ from prediction errors in domains such as reward and motivation (Chowdhury et al., 2013; for a review, see Pearson et al., 2011) or language processing (Federmeier et al., 2002, 2010), where older adults also sometimes show impaired predictive abilities. Nevertheless, theoretical development may benefit from characterizing similarities and differences among the effects of various kinds of predictions errors on event memory updating. For instance, Event Segmentation Theory (Zacks et al., 2007) suggests a key role for prediction errors in chunking ongoing perceptions into discrete units that can be encoded as events. Older adults sometimes show heterogeneity in event segmentation that is negatively associated with memory for everyday activities (Bailey et al., 2013; Kurby & Zacks, 2011; Zacks et al., 2006; but see Kurby & Zacks, 2018; Sargent et al., 2013). One possibility is that the updating mechanism that leads to changes in long-term memory observed here also leads to the subjective experience of segmentation.

In conclusion, the present study supports a role for prediction errors in prompting the formation of a configural representation of the previous related event, the new event, and the discrepancy between the two. This memory updating mechanism can benefit subsequent memory differentiation. These findings join related studies in implicating links between prior-event retrieval while viewing changed actions and subsequent recollection of updated event memories and in suggesting that these links are weaker in older than younger adults (Garlitch & Wahlheim, 2021; Stawarczyk et al., 2020; Wahlheim & Zacks, 2019). When mnemonic predictions do not stimulate integrative encoding that successfully leads to change recollection, the retrieved features upon which predictions were based are more likely to later intrude. This accounts for the lack of prediction error benefits to recall for older adults who experienced impairment in this cascade of processes. These findings could inspire techniques for repairing everyday memory updating deficits: A key may be to educate viewers to look back to relevant previous experiences during encoding, promoting strong, specific predictions that can trigger robust memory updating when the new experience conflicts with predictions based on the past.

References

Amer, T., & Davachi, L. (2023). Extra-hippocampal contributions to pattern separation. *ELife*, *12*, e82250. https://doi.org/10.7554/eLife.82250

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (2nd ed., pp. 237–313). Academic Press.

Bailey, H. R., Kurby, C. A., Giovannetti, T., & Zacks, J. M. (2013). Action perception predicts action performance. *Neuropsychologia*, *51*(11), 2294–2304. https://doi.org/10.1016/j.neuropsychologia.2013.06.022

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bein, O., Duncan, K., & Davachi, L. (2020). Mnemonic prediction errors bias hippocampal states. *Nature Communications*, *11*(1), 3451. https://doi.org/10.1038/s41467-020-17287-1

Bein, O., Plotkin, N. A., & Davachi, L. (2021). Mnemonic prediction errors promote detailed memories. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *28*(11), 422–434. https://doi.org/10.1101/lm.053410.121

Brainerd, C. J., & Reyna, V. F. (2015). Fuzzy-Trace Theory and Lifespan Cognitive Development. *Developmental Review: DR*, *38*, 89–121. https://doi.org/10.1016/j.dr.2015.07.006

Brod, G., Hasselhorn, M., & Bunge, S. A. (2018). When generating a prediction boosts learning: The element of surprise. *Learning and Instruction*, *55*, 22–31. https://doi.org/10.1016/j.learninstruc.2018.01.013

Burton, R. L., Lek, I., Dixon, R. A., & Caplan, J. B. (2019). Associative Interference in Older and Younger Adults. *Psychology and Aging*, *34*(4), 558–571. https://doi.org/10.1037/pag0000361

Chen, J., Cook, P. A., & Wagner, A. D. (2015). Prediction strength modulates responses in human area CA1 to sequence violations. *Journal of Neurophysiology*, *114*(2), 1227–1238. https://doi.org/10.1152/jn.00149.2015

Chowdhury, R., Guitart-Masip, M., Lambert, C., Dayan, P., Huys, Q., Düzel, E., & Dolan, R. J. (2013). Dopamine restores reward prediction errors in old age. *Nature Neuroscience*, *16*(5), 648–653. https://doi.org/10.1038/nn.3364

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Craik, F. I. M. (2022). Reducing age-related Memory Deficits: The Roles of Environmental Support and self-initiated Processing Activities. *Experimental Aging Research*, *48*(5), 401–427. https://doi.org/10.1080/0361073X.2022.2084660

Den Ouden, H., Kok, P., & De Lange, F. (2012). How Prediction Errors Shape Perception, Attention, and Motivation. *Frontiers in Psychology*, *3*. https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00548

Devitt, A. L., & Schacter, D. L. (2016). False memories with age: Neural and cognitive underpinnings. *Neuropsychologia*, *91*, 346–359. https://doi.org/10.1016/j.neuropsychologia.2016.08.030

Ebert, P. L., & Anderson, N. D. (2009). Proactive and retroactive interference in young adults, healthy older adults, and older adults with amnestic mild cognitive impairment. *Journal of the International Neuropsychological Society: JINS*, *15*(1), 83–93. https://doi.org/10.1017/S1355617708090115

Elsey, J. W. B., Van Ast, V. A., & Kindt, M. (2018). Human memory reconsolidation: A guiding framework and critical review of the evidence. *Psychological Bulletin*, *144*(8), 797–848. https://doi.org/10.1037/bul0000152

Exton-McGuinness, M. T. J., Lee, J. L. C., & Reichelt, A. C. (2015). Updating memories—The role of prediction errors in memory reconsolidation. *Behavioural Brain Research*, *278*, 375–384. https://doi.org/10.1016/j.bbr.2014.10.011

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. MEDLINE. https://doi.org/10.3758/BF03193146

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language*, *115*(3), 149–161. https://doi.org/10.1016/j.bandl.2010.07.006

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, *39*(2), 133–146. https://doi.org/10.1017/S0048577202001373

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third). Sage.

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005

Garlitch, S. M., & Wahlheim, C. N. (2021). Directing attention to event changes improves memory updating for older adults. *Psychology and Aging*, *36*(4), 475–490. https://doi.org/10.1037/pag0000503

Greene, N. R., & Naveh-Benjamin, M. (2022). Adult age differences in specific and gist associative episodic memory across short- and long-term retention intervals. *Psychology and Aging*, *37*(6), 681–697. https://doi.org/10.1037/pag0000701

Greene, N. R., & Naveh-Benjamin, M. (2023). Adult age-related changes in the specificity of episodic memory representations: A review and theoretical framework. *Psychology and Aging*, *38*, 67–86. https://doi.org/10.1037/pag0000724

Greve, A., Cooper, E., Kaula, A., Anderson, M. C., & Henson, R. (2017). Does prediction error drive one-shot declarative learning? *Journal of Memory and Language*, *94*, 149–165. https://doi.org/10.1016/j.jml.2016.11.001

Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge Is Power: Prior Knowledge Aids Memory for Both Congruent and Incongruent Events, but in Different Ways. *Journal of Experimental Psychology. General*, *148*(2), 325–341. https://doi.org/10.1037/xge0000498

Grilli, M. D., & Sheldon, S. (2022). Autobiographical event memory and aging: Older adults get the gist. *Trends in Cognitive Sciences*, *26*(12), 1079–1089. https://doi.org/10.1016/j.tics.2022.09.007

Gutchess, A. H., Kensinger, E. A., Yoon, C., & Schacter, D. L. (2007). Ageing and the self-reference effect in memory. *Memory (Hove, England)*, *15*(8), 822–837. https://doi.org/10.1080/09658210701701394

Hamami, A., Serbun, S. J., & Gutchess, A. H. (2011). Self-referencing enhances memory specificity with age. *Psychology and Aging*, *26*, 636–646. https://doi.org/10.1037/a0022626

Henson, R. N., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus*, *20*(11), 1315–1326. https://doi.org/10.1002/hipo.20857

Hermann, M. M., Wahlheim, C. N., Alexander, T. R., & Zacks, J. M. (2021). The role of prior-event retrieval in encoding changed event features. *Memory & Cognition*, *49*(7), 1387–1404. https://doi.org/10.3758/s13421-021-01173-2

Hess, T. M. (2014). Selective Engagement of Cognitive Resources: Motivational Influences on Older Adults' Cognitive Functioning. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *9*(4), 388–407. https://doi.org/10.1177/1745691614527465

Hess, T. M., & Emery, L. (2012). Memory in context: The impact of age-related goals on performance. In M. Naveh-Benjamin & N. Ohta (Eds.), *Memory and Aging* (pp. 183–215). Psychology Press.

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

Hoffman, L., & Walters, R. W. (2022). Catching Up on Multilevel Modeling. *Annual Review of Psychology*, *73*(1), 659–689. https://doi.org/10.1146/annurev-psych-020821-103525

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

Kafkas, A., & Montaldi, D. (2018). Expectation affects learning and modulates memory experience at retrieval. *Cognition*, *180*, 123–134. https://doi.org/10.1016/j.cognition.2018.07.010

Kim, G., Lewis-Peacock, J. A., Norman, K. A., & Turk-Browne, N. B. (2014). Pruning of memories by context-based prediction error. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(24), 8997–9002. MEDLINE. https://doi.org/10.1073/pnas.1319438111

Kim, G., Norman, K. A., & Turk-Browne, N. B. (2017). Neural Differentiation of Incorrectly Predicted Memories. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(8), 2022–2031. https://doi.org/10.1523/JNEUROSCI.3272-16.2017

Kurby, C. A., & Zacks, J. M. (2011). Age differences in the perception of hierarchical structure in events. *Memory & Cognition*, *39*(1), 75–91. MEDLINE. https://doi.org/10.3758/s13421-010-0027-2

Kurby, C. A., & Zacks, J. M. (2018). Preserved neural event segmentation in healthy older adults. *Psychology and Aging*, *33*(2), 232–245. https://doi.org/10.1037/pag0000226

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lenth, R. (2019). *Emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.3.5.1*. https://CRAN.R-project.org/package=emmeans

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222–243. https://doi.org/10.1037/a0017619

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin*, *135*(4), 638–677. https://doi.org/10.1037/a0015849

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*(5), 1170–1187. https://doi.org/10.1037//0278-7393.26.5.1170

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*(1), 104–118. https://doi.org/10.1037/0882-7974.23.1.104

Pearson, J. M., Heilbronner, S. R., Barack, D. L., Hayden, B. Y., & Platt, M. L. (2011). Posterior cingulate cortex: Adapting behavior to a changing world. *Trends in Cognitive Sciences, 15*(4), 143–151.

Quent, J. A., Greve, A., & Henson, R. N. (2022). Shape of U: The Nonmonotonic Relationship Between Object-Location Memory and Expectedness. *Psychological Science*, *33*(12), 2084–2097. https://doi.org/10.1177/09567976221109134

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rhodes, S., Greene, N. R., & Naveh-Benjamin, M. (2019). Age-related differences in recall and recognition: A meta-analysis. *Psychonomic Bulletin & Review*, *26*(5), 1529–1547. https://doi.org/10.3758/s13423-019-01649-y

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

Rubin, D. C., & Umanath, S. (2015). Event memory: A theory of memory for laboratory, autobiographical, and fictional events. *Psychological Review*, *122*(1), 1–23. MEDLINE. https://doi.org/10.1037/a0037907

Salthouse, T. A. (1993). Speed and knowledge as determinants of adult age differences in verbal tasks. *Journal of Gerontology*, *48*(1), P29-36.

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological Psychology*, *54*(1), 35–54. https://doi.org/10.1016/S0301-0511(00)00052-1

Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., Eisenberg, M. L., & Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition*, *129*(2), 241–255. https://doi.org/10.1016/j.cognition.2013.07.002

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime user's guide*. Psychology Software Tools Inc.

Sinclair, A. H., & Barense, M. D. (2018). Surprise and destabilize: Prediction error influences episodic memory reconsolidation. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *25*(8), 369–381. https://doi.org/10.1101/lm.046912.117

Sinclair, A. H., & Barense, M. D. (2019). Prediction Error and Memory Reactivation: How Incomplete Reminders Drive Reconsolidation. *Trends in Neurosciences*, *42*(10), 727–739. https://doi.org/10.1016/j.tins.2019.08.007

Sinclair, A. H., Manalili, G. M., Brunec, I. K., Adcock, R. A., & Barense, M. D. (2021). Prediction errors disrupt hippocampal representations and update episodic memories.

Proceedings of the National Academy of Sciences of the United States of America, *118*(51), e2117625118. https://doi.org/10.1073/pnas.2117625118

Smith, M. E., Newberry, K. M., & Bailey, H. R. (2020). Differential effects of knowledge and aging on the encoding and retrieval of everyday activities. *Cognition*, *196*, 104159. https://doi.org/10.1016/j.cognition.2019.104159

Stawarczyk, D., Wahlheim, C. N., Etzel, J. A., Snyder, A. Z., & Zacks, J. M. (2020). Aging and the encoding of changes in events: The role of neural activity pattern reinstatement. *Proceedings of the National Academy of Sciences*, *117*(47), 29346–29353. https://doi.org/10.1073/pnas.1918063117

Stawarczyk, D., Wahlheim, C. N., & Zacks, J. M. (2023). How memory-based predictions affect change comprehension in everyday life activities for older and younger adults [Data Set, Analysis Script, and Stimulus Materials]. Washington University in St. Louis. https://osf.io/5pn67/

Stawarczyk, D., Wahlheim, C. N., & Zacks, J. M. (2023). *How memory-based predictions affect change comprehension in everyday life activities for older and younger adults [Data set, analysis script, and stimulus materials]*. Washington University in St. Louis.

van Kesteren, M. T. R., Ruiter, D. J., Fernandez, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211–219. MEDLINE. https://doi.org/10.1016/j.tins.2012.02.001

Wahlheim, C. N., Eisenberg, M. L., Stawarczyk, D., & Zacks, J. M. (2022). Understanding Everyday Events: Predictive-Looking Errors Drive Memory Updating. *Psychological Science*, *33*(5), 765–781. https://doi.org/10.1177/09567976211053596

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. MEDLINE. https://doi.org/10.3758/s13421-012-0246-9

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

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, *34*(10), 515–525. https://doi.org/10.1016/j.tins.2011.06.006

Yazin, F., Das, M., Banerjee, A., & Roy, D. (2021). Contextual prediction errors reorganize naturalistic episodic memories in time. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-90990-1

Zacks, J. M. (2020). Event Perception and Memory. *Annual Review of Psychology*, *71*(1), 165–191. https://doi.org/10.1146/annurev-psych-010419-051101

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction Error Associated with the Perceptual Segmentation of Naturalistic Events. *Journal of Cognitive Neuroscience*, *23*(12), 4057–4066. https://doi.org/10.1162/jocn_a_00078

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind/brain perspective. *Psychological Bulletin*, *133*(2), 273–293. MEDLINE.

Zacks, J. M., Speer, N. K., Vettel, J. M., & Jacoby, L. L. (2006). Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychology and Aging*, *21*(3), 466–482. https://doi.org/10.1037/0882-7974.21.3.466