Retrieval of Past Events During Change Experience is Associated with Memory for Change

Mary Hermann

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Part of the Cognitive Psychology Commons

Recommended Citation
https://openscholarship.wustl.edu/art_sci_etds/1986

This Thesis is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
Retrieval of Past Events During Change Experience is Associated with Memory for Change
by
Mary M. Hermann

A thesis presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Master of Arts

December 2019
St. Louis, Missouri
# Table of Contents

List of Figures .......................................................................................................................... iv
List of Tables ............................................................................................................................. v
Acknowledgments .................................................................................................................. vi
Abstract of the Thesis ............................................................................................................ vii

Chapter 1: Introduction ........................................................................................................... 1

Chapter 2: Method (Experiment 1) ....................................................................................... 10
  2.1 Participants ...................................................................................................................... 10
  2.2 Task and Materials ......................................................................................................... 11
  2.3 Design .............................................................................................................................. 13
  2.4 Procedure ......................................................................................................................... 14

Chapter 3: Results (Experiment 1) ....................................................................................... 16
  3.1 Recall ............................................................................................................................... 17
  3.2 Prediction Error ................................................................................................................ 17
  3.3 Memory-Based Predictions ........................................................................................... 18
  3.4 Change Recollection ...................................................................................................... 19
  3.5 Changed Activities Conditionalized on Change Recollection ...................................... 20

Chapter 4: Discussion (Experiment 1) .................................................................................. 23

Chapter 5: Method (Experiment 2) ....................................................................................... 25
  5.1 Participants ...................................................................................................................... 26
  5.2 Task and Materials ......................................................................................................... 26
  5.3 Design .............................................................................................................................. 28
  5.4 Procedure ......................................................................................................................... 29

Chapter 6: Results (Experiment 2) ...................................................................................... 31
  6.1 Recall ............................................................................................................................... 31
  6.2 Prediction Error ................................................................................................................ 32
  6.3 Memory-Based Predictions ........................................................................................... 32
  6.4 Change Recollection ...................................................................................................... 33
  6.5 Change Activities Conditionalized on Change Recollection ...................................... 33
List of Figures

Figure 1: Narrative Video Endings ..................................................................................52

Figure 2: Probabilities of Day 2 Recall ..............................................................................53

Figure 3: Probabilities of Day 1 Intrusions ........................................................................54

Figure 4: Difference on Day 2 Recall for Memory-Based Predictions and Non-Memory-Based Predictions as a Function of Average Memory-Based Prediction Report ..................................55

Figure 5: Breakdown of Experiment 2 Conditions ..............................................................56

Figure 6: Probabilities of Day 2 Recall ..............................................................................57

Figure 7: Probabilities of Day 1 Intrusions ........................................................................58
List of Tables

Table 1: Recall of Novel Activities as a Function of Prediction Error Type ........................................48
Table 2: Recall and Intrusions as a Function of Activity Type and Memory-Based Prediction...49
Table 3: Probabilities of Change Recollection as a Function of Activity Type .................................50
Table 4: Recall as a Function of Activity Type and Memory-Based Predictions ..............................51
Acknowledgments

First, I would like to thank my advisor, Dr. Jeffrey Zacks, for all of his support during this process. I would also like to thank Drs. Ian Dobbins and Todd Braver for serving on my thesis committee. Finally, I would like to thank my collaborator, Dr. Christopher Wahlheim.

Mary Hermann

Washington University in St. Louis

December 2019
ABSTRACT OF THE THESIS

Retrieval of Past Events During Change Experience is Associated with Memory for Change

by

Mary Hermann

Master of Arts in Psychology

Washington University in St. Louis, 2019

Professor Jeffrey Zacks, Chair
Professor Ian Dobbins
Professor Todd Braver

A novel theory has emerged that examines how people process and comprehend change. Two experiments examined these theoretical mechanisms used to detect and recollect changes in everyday activities. Participants viewed movies of an actor performing narrative activities across two fictitious days. During the second movie, participants also completed a prediction task in which they were asked to predict what they thought was going to happen next. In Experiment 1, some activities repeated identically across the two days, some were repeated but changed on a critical feature (e.g. waking up to an alarm from a clock or a phone), and some were only shown on the second day. In Experiment 2, some activities during the second day matched the participant prediction of the critical feature from the prediction task, some activities did not match the critical feature (e.g. predicting the alarm clock and being shown the phone), and some activities were only shown on the second day. After a delay, participants completed a cued recall test for the activities on the second day. Making a prediction that matched the critical feature shown during the first day was associated with encoding of the subsequent ending, both when it repeated and when it changed so that it mismatched the prediction. Contrary to hypothesis,
experiencing prediction error during the viewing of the second day was not associated with improved memory. These results support the proposal successful retrieval of relevant previous instances benefits the subsequent encoding of situational changes.
Chapter 1: Introduction

The environment is constantly changing. In order to survive in this dynamic world, it is necessary to monitor, recognize, and adapt to changes. However, rather than being drastic differences, changes are more often just minor variations of the same basic event. For example, say you drive the same route to work every day, but then construction begins on a portion of that route. The next time you drive to work, you would predict that you could make it through to your destination without any problems, and then experience an error when that prediction is wrong and you encounter the closed road. Recognizing and adapting to this change will help you better predict the next day when you’re driving to work to save yourself some unnecessary time in traffic. There must be a mechanism in place for people to use their past memories to recognize, process, and adapt to these changes in their present environment.

In research on episodic memory, changes are usually conceptualized in terms of interference (Underwood, 1957). Theories of memory interference generally predict that episodic changes will lead to reduced memory when there are different features associated with a common context. The fan effect addresses this type of interference by demonstrating that similar features can lead to competition during retrieval and can impair memory (Anderson & Neeley, 1996). In our construction example, the two different routes are both associated with the same context of driving to work. Because of this overlapping context, you may struggle to remember which route you took after the construction appeared because of proactive interference—memory of the first route competing with the retrieval of the second route. Retroactive interference refers to the reverse relationship, when memory of the second route competes with the retrieval of the first route (McGeoch & McDonald, 1931). While the negative consequences of interference are well-reported in the literature (Anderson & Neeley, 1996), there are situations in which multiple
features associated with a common context can facilitate, rather than inhibit, memory.

*Retroactive facilitation* has been found in paired associated learning when identical cues were presented with similar responses in separate lists, A-B, A-B’, (Barnes and Underwood, 1959) and when identical cues were presented with different responses, A-B, A-D, but the participants were informed about the relationship between the pairs (Robbins & Bray, 1974; Bruce & Weaver, 1973). Similar patterns were found outside paired associate learning in reading passages where interpolated learning of similar, non-identical materials resulted in retroactive facilitation (Ausubel, Robbins, & Blake Jr., 1957). These findings suggest that changes involving common contexts, but different features, do not always lead to interference, but can result in facilitation. In the fan effect paradigm, it has been shown that interference can be reduced if the things to be remembered can be integrated into an event model that is coherent in space and time (Radvansky, 2005). One mechanism that could produce such resistance to interference, or event facilitation, is integration of the multiple encoded items into a compound representation.

This hypothesis inspired the memory-for-change framework, which addresses when change will improve memory or impair memory (Jacoby, Wahlheim, & Kelley, 2015; Wahlheim & Jacoby, 2013; Jacoby, Wahlheim, & Yonelinas, 2013). The memory-for-change framework relies on the idea that experiencing an event can involuntarily trigger a person to remember past events that are similar to the current event, and this triggering process can then be encoded into the representation alongside memory of both the current event and the similar past event. This is known as *recursive reminding* (Hintzman, 2011). The experience of the current event causes a reminding of past events that is then incorporated into a memory trace that includes the current event, that process of reminding, and the past events. Recursive reminding allows for not only the joint encoding of the two sets of event features, but also of the temporal order of events; this
is because the earlier events must precede the later events for the later events to have triggered a reminding of the earlier events.

In the memory-for-change framework, changes in events can lead to either facilitation or interference. If a current event triggers the automatic retrieval of similar past events and the changed features between the different events are successfully detected, both of these events will be integrated into a single memory trace along with the memory of the retrieval process, or the reminding. This integrated trace will facilitate memory recollection for the temporal order of the events because of the inclusion of the later event triggering a reminding of the earlier event. In the construction example, remembering the first route and the construction along it while driving the second route will allow for the preservation of the temporal order between the two routes. However, if there is a failure to detect the changes that occurred between the two events, interference will occur because the same retrieval process that brought the similar earlier events to mind will make the competing responses more accessible and, therefore, make them more able to interfere during retrieval. The ability to retrieve past events during processing is necessary for facilitation to occur.

Research on learning word pairs by Jacoby and colleagues showed support for these facilitative effects of successful retrieval and change detection on memory. In one study, participants were tasked with studying two lists of word pairs, List 1 (e.g. knee – bend) and List 2 (e.g. knee – bone) (Jacoby, Wahlheim, & Yonelinas, 2013). For some of the word pairs, the right-hand word was changed between List 1 and List 2. Participants were instructed to indicate when they detected that such a change had occurred between the two lists. At test, participants were asked to report whether or not a pair was presented in List 2 and whether or not the right-hand word had changed between the lists. Results showed a strong relationship between change detection
during studying and change recollection during test such that better change detection was associated with better subsequent recollection. Participants showed facilitation when the change was recollected, and interference when the change was not recollected. The retrieval of List 1 while studying List 2 allowed for increased change detection and the integration of both lists with the experience of reminding into a single trace that facilitated memory recollection for list membership.

Alongside retrieval, another important theorized mechanism for event change comprehension is the construct of prediction error. Prediction error is thought to underlie methods of human learning and behavior, particularly in reinforcement learning, through phasic firing of dopamine cells (Garrison, Erdeniz, & Done, 2013; Glimcher, 2011; Cavanagh et al., 2010). These studies found links between the magnitude of prediction error, indicated by dopaminergic activity, and behavioral learning. The theoretical framework posits that the amount of dopamine released carries information about prediction errors that occurred from discrepancies between the current state and the predicted state (Montague, Dayan, & Sejnowski, 1996). Although this system was originally identified in the context of unpredicted rewards and absences of reward, it has been proposed that dopaminergic activity carries unsigned error signals as well (Nasser et al., 2017). Prediction errors can influence behavioral decisions as behavior is adjusted to minimize these errors and form more accurate representations (Glimcher, 2011). This process is reflected in reinforcement learning as participants take in reward or punishment information, represented in the brain with dopaminergic activity, and use it to minimize their errors over time by reducing the magnitude of activity. These behavioral adaptations are evident in reaction time slowing after errors and reaction time speeding following the correct choice (Cavanagh et al., 2010).
Prediction error is a key player in the encoding of event memories. Event Segmentation Theory (EST; Zacks et al., 2007) addresses how the continuous activity present in the environment is segmented into meaningful chunks, called events, and how those events are encoded into memory. EST proposes that people form representations, called event models, of what is currently happening. These models are formed from perceptual information taken in from the environment and knowledge from long-term memory to form a working model of the current state of affairs. EST posits that event models serve an advantageous purpose as people are able to use their current model to make predictions about what will happen in the immediate future. This prediction mechanism allows people to be proactive to potentially avoid harmful situations before they occur and to take advantage of impending opportunities. However, this adaptive function is only beneficial if the predictions are accurate. Event models need to veridically represent the current situation in order to drive accurate predictions. Therefore, there must be a process in place to recognize and update inaccurate event models. EST proposes that a spike in prediction error provides a natural signal that can indicate when a working event model is no longer an accurate representation of the current state. Therefore, the spike triggers the event model to be updated to reflect the latest state. This updating is represented as a segmentation between two, now distinct, events in memory.

Event Memory Retrieval and Comparison Theory (EMRC; Wahlheim & Zacks, 2019) combines elements from the memory-for-change framework and EST to create a novel explanation for how memories can guide ongoing comprehension and the processing of changes in events. EST claims that, as people interact with the environment, they are constantly making predictions about what will happen in the immediate future based on their event model representations. According to the memory-for-change framework, as people observe everyday events, they may
experience features that cue recollection of similar earlier events. These recollections of earlier events are now available to influence the predictions made about the immediate future. If the predictions are accurate because the earlier event is the same as the current event, it will result in a maintained event model. If the new event differs from the previous one such that predictions based on the old event are inaccurate, a spike in prediction error will result. This experience of prediction error should result in detection of the changed features and an updating of the current event model.

Note that the memory trace that results from updating based on prediction error encodes the current event, the experience of reminding, and the earlier event; this recursive representation hypothesis comes from the memory-for-change framework. As in that framework, the formation of a recursive representation leads to resistance to interference and benefits memory for temporal order; however, if no recursive representation is formed then interference should occur and memory for temporal order should be poor.

Wahlheim and Zacks (2019) tested EMRC using an everyday changes paradigm that presented participants with a series of narrative movies of an actor performing everyday activities over the course of two days. Each activity had two versions that differed from one another in a central feature (i.e. an alarm going off from a phone or from a clock). The critical manipulation was that, across the two days, some of the activities repeated (e.g., the phone alarmed on both days) whereas some changed (e.g., the phone alarmed on the first day, but the clock alarmed on the second day). In addition, some novel events were presented only on the second day. At test, participants were asked to recall the activity feature from Day 2, whether the activity changed from Day 1 to Day 2, and the activity feature from Day 1 for all activities for which they reported that a change occurred.
Wahlheim and Zacks found that participants’ ability to detect and recollect changes, which hinged on the ability to successfully retrieve similar past events during processing, was associated with subsequent memory facilitation. Detecting and recollecting a change, where a change was correctly reported to have occurred and the features from Day 1 were remembered, was associated with better memory for Day 2 features and fewer intrusions of Day 1 features into Day 2 recall. Successful change detection and recollection allowed for a memory trace that included the features from both days and the experience of prediction error while maintaining the temporal relationship, leading to facilitated memory recall. Detecting a change without recollection, where a change was correctly reported to have occurred but the features from Day 1 were not remembered, was associated with worse Day 2 memory and more intrusions of Day 1 features into Day 2 recall. Detecting but not recollecting led to proactive interference as the Day 1 features competed with Day 2 features upon retrieval.

EMRC proposes that change detection arises as a result of prediction error monitoring and the successful retrieval of similar past events. As participants are watching Day 2 and they see the beginning of the clip, it triggers the retrieval of the activity from Day 1 and they will use that event model to make predictions as to what will happen on Day 2. When that prediction is false because the activity changed from Day 1 to Day 2, the spike in prediction error leads to the creation of a new event for Day 2 that is integrated into a temporally organized trace that contains both Day 1 and Day 2 features and the experience of prediction error, which facilitates memory.

Wahlheim & Zacks (2019) has already established the facilitative effects of successful retrieval during event comprehension. Alongside retrieval, the experience of prediction error is another integral part of the mechanism proposed by EMRC, which has not been investigated directly to
date. These experiments sought to isolate, measure, and manipulate prediction error in order to examine its direct effect on subsequent memory recall. This was accomplished by introducing a new prediction task to the paradigm. Participants watched naturalistic videos of an actor performing activities across two different days, just as in Wahlheim & Zacks (2019). However, during Day 2 viewing, playback was stopped, and participants were asked to make an explicit prediction about what they thought was going to happen during Day 2. Experiment 1 manipulated which activities changed or repeated across the two days, which resulted in the experience of prediction error for some activities and no prediction error for some activities. Experiment 2 directly manipulated whether prediction error was or was not experienced, which resulted in some of the activities changing and some repeating.

Replicating the results of Wahlheim & Zacks (2019), we hypothesized that successful change detection and recollection, due to retrieval during processing, would facilitate memory and offset interference effects. This would appear as better recall of Day 2 features when participants accurately reported that a change occurred and were able to remember Day 1 features. We also hypothesized that memory would be impaired when participants detected but did not recollect change, because of interference from Day 1 features. We posited that this facilitation relied first on the ability to retrieve memories of past events during processing and then on the experience of prediction error. Successful change detection relies on experiencing prediction error during Day 2 viewing and integrating that experience into the memory trace. Therefore, the experience of prediction error would be associated with better memory for Day 2 features and fewer intrusions of Day 1 features. The theoretical construct of prediction error from EMRC would be functionally beneficial for memory through this facilitation, but if the explicit prediction task does not capture this concept of prediction error, it may not result in improved memory. Even
when participants are experiencing a conscious prediction error through the prediction task paradigm, they may not be experiencing the successful retrieval and theoretical prediction error that EMRC posits as necessary for successful change detection and recollection. This could result in no relationship between the experimental experience of prediction error and memory recall.
## Chapter 2: Method (Experiment 1)

Experiment 1 examined the effects of change on memory for naturalistic activities. Participants viewed two movies of an actor performing everyday activities across two different days in her life. These movies will be referred to as Day 1 and Day 2. The activities that made up the Day 2 movie included *repeated* activities that were identical to clips shown in the Day 1 movie, and also *changed* activities that began the same as the corresponding Day 1 activity but ended differently. In addition, we included *novel* activities that were only shown in the Day 2 movie. During Day 2 viewing, participants also completed a prediction task where they were asked to predict the Day 2 ending after being presented with the beginning of the activity to serve as a cue. Participants then returned after a two-day delay to complete a cued recall test that asked about the activities they had seen, as well as a change recollection test that asked whether the activities changed or repeated between Day 1 and Day 2.

### 2.1 Participants

We recruited 43 students recruited from psychology courses at the University of North Carolina at Greensboro. Five participants did not complete the second session of the experiment and could not be included in analyses, resulting in a total sample of 38 participants (21 females, $M_{age} = 18.76$, $SD = 0.93$, range = 18-21). The sample size was determined based on earlier studies examining the role of change recollection in moderating proactive effects of memory (Wahlheim, Maddox, & Jacoby, 2014; Wahlheim & Zacks, 2019). Participants were compensated for time spent with partial course credit. All participants provided informed consent in accordance with the university’s Institutional Review Board.
2.2 Task and Materials

The materials were a series of videos of a female actor performing daily activities around the actor’s home on two fictional days in her life. For each activity, there were two versions of the video that differed in a critical feature. To create two versions of each activity, these scenes were divided into two segments. The first segment, which appeared in both versions of the activity, displayed identical footage that ended just before the changed feature appeared on-screen. These clips were on average 10.07 s (min: 1 s, max: 40 s) in duration. The second segment of each scene displayed the changed feature of the activity. Second segments averaged 26.85 s (min: 2 s, max: 88 s) in duration. Figure 1 depicts an example of an activity. The first segment shows the actor sleeping; second segment A the alarm clock goes off, whereas in second segment B the alarm on her phone goes off.

The material set was made up of 59 total activities (45 tested, 14 untested filler items). To counterbalance activities across the experimental manipulation, the 45 critical items were divided into 3 groups of 15. Each activity was predetermined to appear once in each activity type condition—Changed, Repeated, and Novel across the three counterbalancing sets. Assignment of items to conditions was implemented by changing which activities and which versions of those activities were presented in the Day 1 movie; thus, each group saw a different Day 1 movie but all participants saw the same Day 2 movie. Repeated items were repeated exactly in the two movies. For changed items, the version shown during Day 1 ended differently than the version during Day 2. Novel items were not shown in the Day 1 movie. For example, if in the first format, the alarm activity from Figure 1 was designated as Changed, participants would see the alarm going off from the phone during Day 1 and the clock during Day 2. If in the next format, it was Repeated, participants would see the alarm going off from the clock during both Day 1 and
Day 2. In the final format, the activity was Novel, and participants would only see the alarm going off from the *clock* during Day 2 and would not see this activity during Day 1.

Each Day 1 movie featured the actor performing 44 activities (30 tested, 14 untested fillers). The durations of the Day 1 movies were as follows: 26 min and 33 s (Format 1), 26 min and 14 s (Format 2), and 22 min and 58 s (Format 3). The Day 2 movie portrayed the actor performing 59 activities (45 tested, 14 untested fillers). The Day 2 movie had a total duration of 35 min 31 s. The activity type conditions appeared in a fixed-random order restricted so that no more than three consecutive critical activities appeared as the same item type in any experimental format. Untested filler items were added to the movies to increase the coherence of the movie as a whole. All filler items were repeated across movies.

Day 1 was viewed straight through as one continuous film; the segmentation into activities was not emphasized by the editing, and the overall impression was intended to be of a continuous narrative. During Day 2 movie viewing, the activities were interrupted by the additional prediction task. For this task, participants were shown the identical beginning for each activity, ending just before the divergence point at which the two versions became discriminably different. Playback was then stopped and participants were asked to make a prediction about the ending of the activity for Day 2. They did so by selecting one of two still images, which were extracted from the two second segments of the current activity. Each picture clearly displayed the changed feature of the activity. Participants made a prediction about which shown event would occur during Day 2 by selecting either the left-side picture by pressing “Q” or the right-side picture by pressing “P” on the computer keyboard. Following the participant’s prediction, a black slide appeared asking the participants to imagine the selected actions as they would play
out. After five seconds, the Day 2 movie resumed playback with the predetermined ending that reflected the correct activity type condition relationship to Day 1.

Memory was measured with cued recall and change recollection tests. Participants were presented with 59 activity test cues, one for each activity shown during Day 2. The cue question asked about the critical feature of the activity in the Day 2 movie (e.g. “What did the actor use to clean the counter?”). Participants typed their answer into a provided box beneath the question. After entering each response, a prompt appeared asking whether that activity had been different in the Day 2 movie from its representation in the Day 1 movie. Participants pressed the “1” key to indicate that the activity had changed and the “2” key to indicate that the activity had not changed. When the participant indicated that a change did occur, a cue appeared asking the participant to recall the Day 1 feature by typing it into a provided answer box. After entering the Day 1 response, or indicating that the activity had no changed, the cue corresponding to the next activity appeared. All computerized aspects of this experiment were presented using E-Prime 3 software (Psychology Software Tools, Pittsburgh, PA) on a desktop computer with a screen size of 24 inches (60.96 cm). Participants sat approximately 2 feet (61 cm) away from the computer.

2.3 Design
This study utilized a within-subjects design in which we manipulated activity type as our independent variable (Repeated, Changed, Novel). As there was a single, identical Day 2 movie for all participants, the Day 1 movie manipulated the activity endings to reflect one of these activity type conditions. Changed trials depicted different activity endings between the two movies, Repeated trials depicted the same ending between the movies, and Novel trials were only shown during the Day 2 movie. In addition to the direct experimental manipulation, for Repeated and Changed items participants could make either a Memory-Based Prediction or a
Non-Memory-Based Prediction. This was determined by whether or not the participant’s prediction was consistent with the ending they had seen during Day 1 (Memory-Based) or not (Non-Memory-Based). This did not apply to the novel activities as they were not viewed on Day 1.

Depending upon the combination of the above factors, activities could either result in the experience of prediction error or in no experience of prediction error. For example, a Memory-Based Prediction on a Changed trial would lead to prediction error because the prediction would not match the Day 2 ending. A Not-Memory-Based Prediction on a Changed Trial would not lead to prediction error because the participant prediction would match the Day 2 ending.

2.4 Procedure
Participants completed the tasks in two experimental sessions, separated by two days. Both sessions were conducted individually.

During the first session, participants watched movies of an actor performing everyday activities around the actor’s home on two fictional days in her life. Prior to viewing the first movie (Day 1), participants were told to pay careful attention to the actor’s actions and that they would be tested later. After the instructions were presented, participants watched a practice activity to become familiar with the type of activity they would see, and then watched the Day 1 movie corresponding to their counterbalancing condition.

Prior to viewing the second film (Day 2), participants were informed that while some of the activities the actor performed would be identical to Day 1, some actions would be similar but performed in a different way, and some actions would be new altogether. Although participants were informed of the presence of changed and repeated items, no instructions were given as to
what strategy the participant should use during the prediction task. After instructions, participants proceeded to an example of the prediction task before the full Day 2 movie and prediction task.

During the second session, participants completed a computerized version of the Shipley vocabulary test and the cued recall test for Day 2 activities that included change recollection measures. Before beginning the recall test, participants viewed two film clips of one activity to represent the type of change which they were being asked to report. In the example video, the actor wipes down the kitchen counter using a dish rag on the first day but using a paper towel on the second day. Participants then completed the cued recall and change recollection tests for all activities.
Chapter 3: Results (Experiment 1)

The level for significance was set at $\alpha = .05$. Participants’ recall responses were classified into four different types for scoring: Correct responses were accurate descriptions of the critical feature in the Day 2 movie; Day 1 Intrusion were descriptions that included the critical feature from Day 1 instead of the critical feature from Day 2, for repeated and novel activities these were characterized as descriptions in which participants reported the opposite critical feature, regardless of whether or not is was depicted on Day 1; Ambiguous responses were answers that were correct descriptions of the activity’s action but did not include any information necessary to distinguish between the two critical features from the days; finally, Other Error responses included descriptions that did not include any critical features for the specific activity, whether because they referenced a non-target activity or because they referenced an activity that was never viewed, and descriptions that were omitted. As an example, consider potential responses to an event in which participants saw the actor wake up to an alarm clock on Day 1 and to a phone on Day 2. A Correct response might be “she turned off the alarm on her phone;” a Day 1 Intrusion might be “she turned off the alarm on her clock;” an Ambiguous response might be “she turned off the alarm;” and an Other Error might be “she slept.” We did not have any a priori hypotheses regarding Ambiguous and Other Error responses, so we report only analyses on Correct and Day 1 Intrusion responses. Recall was scored such that only Correct answers received credit and all other responses were marked as incorrect.

For analysis, we computed response probabilities for Day 2 recall and Day 1 intrusions for each activity type and prediction error type using logistic mixed effects models. Recall performance was analyzed for critical activities only and did not include any of the filler items. Activity Type and Prediction Error Type were modeled as fixed effects, with subject and activity as random
effects. We included activity as a random effect because of the potential for item selection effects. Models were fitted using the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2015), hypothesis tests were run using the Anova function of the car package (Fox & Weisberg, 2011), and post hoc pairwise comparisons were performed using the emmeans package (Lenth, 2019). Figures depict the model estimated probabilities with 95% confidence intervals.

### 3.1 Recall

The analysis of Day 2 recall (Figure 2) showed that repeated activities were remembered best and changed items were remembered worst, leading to a significant main effect of Activity Type, $\chi^2(2) = 18.732, p < .001$. There was no significant main effect of Prediction Error Type, $\chi^2(1) = 2.130, p = .144$. Post hoc tests showed that, within activity type groups, none of the prediction error type conditions differed significantly from one another. There was no significant Activity Type x Prediction Error Type interaction, $\chi^2(2) = 1.306, p = .521$.

The analysis of Day 1 intrusions (Figure 3) showed that changed activities produced more Day 1 intrusions than novel and repeated activities, leading to a significant main effect of Activity Type, $\chi^2(2) = 25.509, p < .001$. There was no significant main effect of Prediction Error Type, $\chi^2(1) = 0.387, p = .534$. Post hoc tests showed that, within activity type groups, none of the prediction error type conditions differed significantly from one another. There was no significant Activity Type x Prediction Error Type interaction, $\chi^2(2) = 1.130, p = .569$.

### 3.2 Prediction Error

The models depicted in Figures 2 & 3 provided no evidence for memory benefit or impairment for activities during which participants experienced prediction error. However, the experience of prediction error relied on the ability to accurately remember Day 1. Prediction error could only
be experienced if a participant was able to remember the critical feature from Day 1 and consequently experience an error when that prediction was wrong. In response, we examined the construct of prediction error in the novel activities, which eliminated the need for successful retrieval of Day 1. In order to experience prediction error. This “pure” construct was tested with a model that had Prediction Error Type as a fixed effect, subjects and activities as random effects, for Day 2 recall of the novel activities only. Contrary to hypothesis, activities associated with prediction error were not remembered better than activities with no prediction error, $\chi^2(1) = 1.007, p = .316$ (see Table 1).

### 3.3 Memory-Based Predictions
EMRC entails that successful retrieval of Day 1 features during Day 2 viewing is necessary in order to detect changes, experience prediction error, and form the integrated trace that contains features from both events along with the prediction error experience. Therefore, we hypothesized that being able to accurately retrieve Day 1 features during Day 2 viewing would be associated with better subsequent memory. We believed this retrieval process would influence the predictions made during Day 2 viewing, such that participants would make predictions that were consistent with their memory for Day 1. We refer to these predictions that are consistent with what happened on Day 1 as memory-based predictions, and predictions that were inconsistent with what happened on Day 1 as non-memory-based predictions. A memory-based prediction indicates likely successful retrieval of Day 1, whereas non-memory-based prediction indicates a likely failure to retrieve Day 1 features. In Experiment 1, participants were not instructed to make predictions based on their memory for Day 1 and so no specific a priori hypotheses were made about the effect of memory-based predictions on Day 2 recall for Experiment 1. Rather, we examined the characterization of memory-based predictions with post hoc analyses to illuminate
any potential relationship. Participants made memory-based predictions 65% of the time for both changed and repeated activities. Figure 4 depicts the relationship between a participant’s average tendency to make memory-based predictions, or how often they made memory-based predictions during the prediction task, and the magnitude of the memory-based prediction effect on Day 2 recall, or the difference between Day 2 recall performance on memory-based prediction trials and non-memory-based prediction trials. It is clear, from this figure, that participants who were able to successful retrieve Day 1 features and adopted the strategy of using those memories to guide their predictions and make memory-based predictions showed a benefit for Day 2 recall on memory-based prediction trials. Participants who made memory-based predictions at well above the chance level reported better Day 2 recall on trials in which they made a memory-based prediction than trials in which they made a non-memory-based prediction, indicating a relationship between memory-based predictions and subsequent memory. In models that examined Day 2 recall and Day 1 intrusions with Activity Type and Memory-Based Prediction Type as fixed effects, subjects and activities as random effects, no significant main effects were found for Memory-Based Prediction Type for Day 2 recall, $\chi^2(1) = 1.152, p = .283$, or Day 1 intrusions, $\chi^2(1) = 0.004, p = .950$. No significant interactions were found for Activity Type x Memory-Based Prediction Type for either Day 2 recall, $\chi^2(1) = 1.324, p = .250$, or Day 1 intrusions, $\chi^2(1) = 1.490, p = .222$. Table 2 shows the model estimated probabilities for Day 2 recall and Day 1 intrusions broken down by activity type and memory-based prediction type.

3.4 Change Recollection

We crafted separate logistic mixed effect models for change recollection that estimated the probability of participants indicating during test that the activities had changed between Day 1 and Day 2. These models had Activity Type as the fixed effect and subjects and activities as
random effects. For changed activities, these change recollection probabilities represented correct classifications while the probabilities of novel and repeated activities represented incorrect classifications. The probability of reporting a change was significantly greater on changed activities than on novel activities, $z$ ratio = 11.109, $p < .001$, significantly greater on novel activities than repeated activities, $z$ ratio = 8.594, $p < .001$, and significantly greater on changed activities than repeated activities, $z$ ratio = 17.008, $p < .001$, leading to a significant main effect of Activity Type, $\chi^2(2) = 305.11$, $p < .001$. Table 3 displays the probability of change recollection for each activity type.

3.5 Changed Activities Conditionalized on Change

Recollection

We hypothesized that correct change detection and recollection would be associated with better memory for Day 2. EMRC proposes that successful retrieval of Day 1 features during Day 2 viewing would allow for change detection and the integration of a single trace that contains Day 1 features, the experience of prediction error, and Day 2 features. This would facilitate memory for the temporal order of these events and reduce interference of features from the Day 1 event into Day 2 recall. To test this, we examined recall of changed activities only, conditionalized on change recollection. Changed activities could be classified as Change Recollected, where the change was reported to have occurred and Day 1 features were recalled, Change Remembered but Not Recollected, where the change was reported to have occurred but Day 1 features were not recalled, and Change Not Remembered, where the change was not reported to have occurred and Day 1 features were not recalled. We built logistic models with Change Classification and Prediction Error Type as fixed effects, subjects and activities as random effects.
The right pane of Figure 2 plots correct Day 2 recall for changed items conditionalized on change recollection, and the right pane of Figure 3 plots Day 1 intrusions conditionalized in the same way. For Day 2 recall, memory was better when change was recollected than when change was remembered but not recollected and when change was not remembered, leading to a significant main effect of Change Classification, $\chi^2(2) = 114.233, p < .001$. There was no significant main effect of Prediction Error Type, $\chi^2(1) = 1.367, p = .242$, and no significant Change Classification x Prediction Error Type interaction, $\chi^2(2) = 1.863, p = .394$. Post hoc tests revealed that within both prediction error types, memory was significantly better when change was recollected than when change was not remembered, smallest $z$ ratio = 6.782, $p < .001$, and when change was recollected than when change was remembered but not recollected, smallest $z$ ratio = 5.337, $p < .001$, but there was no significant difference in memory when change was not remembered and when change was remembered but not recollected, largest $z$ ratio = 1.681, $p = .5441$.

There were fewer Day 1 Intrusions when change was recollected than when change was remembered but not recollected and when change was not remembered, leading to a significant main effect of Change Classification, $\chi^2(2) = 54.479, p < .001$. There was no significant main effect of Prediction Error Type, $\chi^2(1) = 2.390, p = .122$, but there was a significant Change Classification x Prediction Error Type interaction, $\chi^2(2) = 12.544, p < .01$. Post hoc tests revealed that for activities associated with prediction errors there were significantly more intrusions when change was not remembered than when change was recollected, $z$ ratio = 7.004, $p < .001$, significantly more intrusions when change not remembered but not recollected than when change was recollected, $z$ ratio = 5.834, $p < .001$, but no significant difference between when change was remembered but not recollected and when change was not remembered, $z$ ratio
= 0.332, \( p = .9995 \). For no-prediction-error trials, there were significantly more intrusions when change was not remembered than when change was recollected, \( z \) ratio = 3.946, \( p < .01 \), significantly more intrusions when change was remembered but not recollected than when change was recollected, \( z \) ratio = 4.597, \( p < .001 \), and significantly more intrusions when change was remembered but not recollected than when change was not remembered, \( z \) ratio = 3.206, \( p < .05 \). These results indicate that enhanced Day 2 recall and fewer Day 1 intrusions relies on the ability to correctly recall Day 1 features.
Chapter 4: Discussion (Experiment 1)

The results of Day 2 recall performance indicated that there was a benefit of repetition in viewing events in memory and that changing a critical feature resulted in proactive facilitation when the change was remembered and recollected, but proactive interference when the change was not recollected. The pattern of intrusions of features from Day 1 mirrored the results for correct Day 2 recall: When change was recollected such intrusions were rare, but when it was not, they were increased. Successful retrieval of Day 1 during Day 2 viewing allowed for the creation of an integrated trace that included Day 1 and Day 2 features preserved in their temporal order. Therefore, successful retrieval of Day 1 will allow for the creation of this trace that will reduce interference, and facilitate memory, when asked to recall Day 2.

We characterized successful retrieval of Day 1 with memory-based predictions. While we observed no significant benefit of memory-based predictions on recall performance, there is clearly a relationship between memory-based predictions and subsequent memory, indicated by the post hoc analysis of individual differences, such that making memory-based predictions is associated with better memory. This relationship was found in post hoc analysis because no specific hypotheses were made about memory-based predictions for Experiment 1. During Day 2 viewing, participants were instructed to make their predictions based on what they believed was “more likely” to happen and with the knowledge that some activities would be different from Day 1. Participants were not instructed to make their predictions for Day 2 based on their memory for Day 1, which is assumed in the characterization of memory-based and non-memory-based predictions. Without this instruction, there is no indication whether participants were able to retrieve Day 1 features, but chose instead to predict the opposite ending, which would mask successful retrieval. In Experiment 2, we modified the instructions to tell participants to make
their predictions based on their memory for Day 1 and made the specific hypothesis that memory-based predictions would be associated with better memory performance.

Finally, we found no benefit of experiencing prediction error on subsequent memory performance, except for a significant Change Classification x Prediction Error Type interaction for Day 1 intrusions. In this experiment, the experimental manipulation was activity type, with the experience of prediction error as a consequence. The combination of activity type with prediction accuracy allowed prediction error to be measured as a resulting variable and compare its effect on memory between trials during which prediction error was experienced and when it was not experienced. Experiment 2 reverses this relationship and directly manipulates the experience of prediction error in order to eliminate a potential confound caused by the interaction between items and the experimental manipulation.
Chapter 5: Method (Experiment 2)

Experiment 1 showed that the experience of prediction error on trials when participants’ predictions did not match the Day 2 ending was not associated with better subsequent memory than on trials where participants’ predictions did match the Day 2 ending and no prediction error was experienced, and that memory-based predictions were not associated with better memory performance. However, in Experiment 1, the experience of prediction error was not directly manipulated; instead, it was measured correlationally in response to the experimental manipulation of changing the endings of events. This entails a subtle confound: Prediction error is confounded with the interactions between items and the experimental manipulations. For example, suppose that participants were biased based on their prior knowledge to remember the actor waking up to a cell phone rather than to a clock (Figure 1). If so, they would be more likely to experience prediction error when the alarm clock was presented on Day 1 and repeated, or when the cell phone was presented on Day 1 and then changed, relative to when the cell phone was presented on Day 1 and repeated or the clock on Day 1 and then changed. Such interactions could distort the actual effects of prediction error on performance (or lack thereof).

To address this, we modified the paradigm to directly manipulate the experience of prediction error using a prediction-contingent display paradigm. Each trial with a recurring activity was randomly assigned to either a Prediction Error or No Prediction Error condition. During Day 2 viewing, participants made a prediction response and then were shown either the ending that they selected (for No Prediction Error trials) or the alternate (for Prediction Error trials). For example, if an activity were assigned to the Prediction Error condition and the participant predicted the alarm going off from the clock, they would then be shown the alarm going off from the phone. No Prediction Error trials would always show the ending that was consistent with the
participant’s prediction. If the participant predicted the *clock*, the ending would show the *clock*. Thus, whereas in Experiment 1 whether each item was Changed or Repeated was randomized, in this case prediction error was randomized and whether each item was Changed or Repeated depended on the participant’s response. We also included Novel activities that were only shown on Day 2, as in Experiment 1.

### 5.1 Participants
We recruited 42 undergraduate students from the Washington University in St. Louis subject pool. Five participants did not complete all the experimental sessions and the data of two participants were lost due to technical errors, leaving a total of 35 participants for analysis (23 females; $M_{age} = 19.80$, $SD = 1.41$, range = 17-24). Participants were either paid $10 per hour or given course credit. All participants provided informed consent in accordance with the university’s Institutional Review Board.

### 5.2 Task and Materials
The tasks and materials were almost identical to Experiment 1. The materials used came from the same pool of videos used in Experiment 1, in which a woman performed a series of activities across two fictitious days in the actor’s life. There was a total of 59 activities (45 tested, 14 untested fillers) and participants watched two separate movies consisting of these activities, Day 1 and Day 2. The clips had identical beginnings with two different possible endings for each activity.

As in Experiment 1, participants first watched a movie consisting of a series of narrative activities, called Day 1, and then a second movie of narrative clips, called Day 2. Of the 45 tested activities, 36 were shown on both Day 1 and Day 2, while 9 were presented as novel activities only on Day 2. To counterbalance the assignment of items to prediction error conditions and
which version of each activity was shown on Day 1, the 36 recurring activities were divided into
groups of 18 so each activity appeared once in each combination of Prediction Error or No
Prediction Error and ending version across four different experimental formats. This ensured
each activity was shown as both a Prediction Error and No Prediction Error activity. The 9
activities shown as novel clips on Day 2 also differed between these two formats and were
divided similarly in accordance with the manipulation. Each of the four resulting Day 1 movies
showed the actor completing 50 activities (36 tested, 14 untested fillers). The durations of each
of the Day 1 movie formats were: 21 min and 37 s (Format 1), 21 min and 19 s (Format 2), 22
min and 28 s (Format 3), 22 min and 21 s (Format 4).

Participants again watched the Day 1 movie straight through as one continuous film as in
Experiment 1. Also, as in Experiment 1, during the viewing of the Day 2 movie participants were
asked to make predictions about the endings before they were shown. However, in Experiment 2
these Day 2 endings were contingent upon participant prediction. For each of the 59 activities of
Day 2 (45 tested [36 recurring, 9 novel], 14 untested fillers), the participant was shown the
beginning of each activity, up to the divergence point at which the two versions became
discriminably different, at which point playback was stopped. Participants were then presented
with two still images extracted from the endings of the two versions of the current activity. The
stills clearly depicted the feature that differed between the two versions. Participants were asked
to predict what they thought was going to happen on Day 2, based on their memory for Day 1.
This is an important difference from Experiment 1, where participants were given no additional
instructions on strategies for the prediction task. Participants made their predictions by pressing
either the “1” or “2” key, corresponding to whichever number was listed underneath the image of
their prediction. After the prediction was made, the Day 2 movie resumed playback. This was
repeated for each of the 59 activities. The total duration of the Day 2 movie depended on participants’ predictions, but could range from 23 min and 57 s, if the participant prediction resulted in the shorter of the two possible endings for every activity, to 26 min and 50 s, if the participant prediction resulted in the longer of the two possible endings for every activity. These duration ranges only include the running time for the Day 2 movies and exclude pauses for the prediction task. In short, what participants saw and what they did was essentially the same as in Experiment 1; however, in this case the ending of each activity was manipulated based on their prediction to control whether they experienced a prediction error.

Similar to Experiment 1, memory was measured with cued recall and change recollection tests. Participants were presented with the same activity test cues that were presented on Day 2. After the cue played, participants were asked to describe the ending of the activity for Day 2 by typing their response on the keyboard. After reporting the Day 2 ending, participants were asked to report whether this activity changed or repeated from Day 1 to Day 2. Participants pressed the “1” if they believed an activity changed, and the “2” if they believed it repeated. If a participant reported that the activity changed, they were then asked to report what the ending was for Day 1 by typing their response on the keyboard. This was repeated for all activities.

All computerized aspects of this experiment were presented using PsychoPy software (Peirce et al., 2019) on a desktop computer with a screen size of 21.5 inches (54.61 cm). Participants sat approximately 2 feet (61 cm) away from the computer.

5.3 Design
This study utilized a within-subjects design in which we manipulated the participants’ experience of prediction error as an independent variable (Prediction Error, No Prediction Error). The Day 2 ending for each activity was contingent upon the combination of prediction error condition and
the participant prediction. On Prediction Error trials, we presented participants with the activity version opposite of their prediction during Day 2 viewing (e.g. an alarm clock ending following a phone alarm prediction). On No Prediction Error trials, participants were presented with the activity version that matched their prediction during Day 2 viewing (e.g. an alarm clock ending following an alarm clock prediction). Participants’ predictions could be described, as in Experiment 1, as either Memory-Based (consistent with what happened on Day 1) or a Non-Memory-Based (inconsistent with what happened on Day 1). This did not apply to the novel activities as they were not viewed on Day 1.

Each activity ended either having Repeated what happened on Day 1 or having Changed. Whereas in Experiment 1 the choice of Repeated or Changed was randomly assigned, in Experiment 2, Prediction Error was randomly assigned, whether an activity was Repeated or Changed depended on whether it was assigned to the Prediction Error or No Prediction Error condition, combined with the participant’s prediction response. For example, a Memory-Based Prediction on a Prediction Error trial would lead to a changed activity from Day 1 to Day 2, whereas a Non-Memory-Based Prediction on a Prediction Error trial would lead to a repeated activity. For the novel clips, the variables describing memory-based prediction and change do not apply. For the untested fillers, we showed the same ending across Day 1 and Day 2, regardless of participant prediction; this had the effect of increasing the number of times in which an activity repeated across the two days, and was intended to increase the motivation of participants to make predictions consistent with the Day 1 ending.

5.4 Procedure
Participants completed the tasks in three sessions, each separated by a week. Videos, instructions, and memory test were all presented with PsychoPy software on a desktop computer.
During the first session, participants were told that they would be watching a movie depicting activities occurring over a single day and were instructed to pay close attention and to try to commit the activities to memory. After the instructions were presented, participants watched two practice activities to become familiar with the type of activity they would see, and then watched the Day 1 movie corresponding to their counterbalancing condition.

During the second session, participants were instructed that they would be watching another day’s activities, and that some of the activities for Day 2 would be the same as Day 1 while some would be different. The addition of the novel activities was not specifically mentioned beyond the emphasis that the activities were performed on a second, different day. After Day 2 instructions, participants performed the prediction task on two practice examples to ensure they understood the instructions. One practice example was a Prediction Error trial and one was a No Prediction Error trial. After completing the practice examples, the experimenter explicitly pointed out the change that occurred from Day 1 to Day 2 in the Prediction Error trial and the repetition of activity from Day 1 to Day 2 in the No Prediction Error trial, drawing attention to the type of change participants were supposed to notice. Participants then performed the prediction task with each of the Day 2 activities.

During the third session, participants completed the cued recall test for Day 2 and Day 1 activities as well as change recollection judgments.
Chapter 6: Results (Experiment 2)

Experiment 2 used the same recall response classification and scoring criteria as Experiment 1, analyzing only Correct and Intrusion responses. We compared Day 2 correct recall and Day 1 intrusions for each activity type and prediction error type using the same logistic mixed effects models as in Experiment 1.

6.1 Recall

The analysis of Day 2 recall (Figure 6) showed that repeated activities were remembered best and changed items were remembered worst, leading to a significant main effect of Activity Type, $\chi^2(2) = 12.276, p < .01$. There was no significant main effect of Prediction Error Type, $\chi^2(1) = .433, p = .511$. Post hoc tests showed that, within activity type groups, none of the prediction error type conditions differed significantly from one another. There was a significant Activity Type x Prediction Error Type interaction, $\chi^2(2) = 8.065, p < .05$.

The analysis of Day 1 intrusions (Figure 7) showed that changed activities elicited more Day 1 intrusions than novel and repeated activities, leading to a significant main effect of Activity Type, $\chi^2(2) = 13.514, p < .01$. There was no significant main effect of Prediction Error Type, $\chi^2(1) = 2.607, p = .106$. Post hoc tests showed that, within activity type groups, the only significant difference was between the prediction error types for changed activities. There were significantly more intrusions with the prediction error type than the no-prediction-error type, $z$ ratio = 3.277, $p < .05$. There was a significant Activity Type x Prediction Error Type interaction, $\chi^2(2) = 9.371, p < .01$. 
6.2 Prediction Error
The models depicted in Figures 5 & 6 resulted in no significant main effects for Prediction Error Type on Day 2 recall and Day 1 intrusions. Again, it was necessary to create the same model as Experiment 1 that tested “pure” prediction error with the novel activities that do not rely on accurate memory of Day 1 to experience prediction error during Day 2 viewing. It was revealed that there was no significant main effect of Prediction Error Type, $\chi^2(1) = 2.158, p = .142$. This result also trended in the opposite direction of our hypothesis, reporting that not experiencing prediction error was associated with better memory.

6.3 Memory-Based Predictions
Experiment 2 made a specific a priori hypothesis that memory-based predictions would be associated with better subsequent memory. We used the same models as Experiment 1 to examine the effect of memory-based predictions on Day 2 recall. Participants made memory-based predictions 75% of the time for both changed and repeated activities. Memory-based predictions were associated with better Day 2 recall than non-memory-based predictions, leading to a significant main effect for Memory-Based Prediction Type, $\chi^2(1) = 7.956, p < .01$, and significantly fewer Day 1 intrusions, $\chi^2(1) = 5.114, p < .05$. Post hoc tests found no significant differences between memory-based predictions and non-memory-based predictions for either activity type group in Day 2 recall. For Day 2 recall, there was no significant interaction between Activity Type and Memory-Based Prediction Type, $\chi^2(1) = 0.094, p = .759$. For Day 1 intrusions, it was not possible to estimate the interaction because there were no Day 1 intrusions observed in the repeated condition. The main effect of Memory-Based Prediction Type on Day 1 intrusions came from a reduced model fit only with Memory-Based Prediction Type. Table 4
shows the model estimated probabilities for Day 2 recall broken down by activity type and memory-based prediction type.

### 6.4 Change Recollection
We used the same change recollection models as Experiment 1. The probability of reporting a change was significantly greater on changed activities than on novel activities, $z$ ratio = 4.137, $p < .001$, significantly greater on novel activities than repeated activities, $z$ ratio = 2.551, $p < .05$, and significantly greater on changed activities than repeated activities, $z$ ratio = 9.318, $p < .001$, indicating a significant main effect of Activity Type, $\chi^2(2) = 88.209, p < .001$. Table 3 displays the probability of change recollection for each activity type.

### 6.5 Change Activities Conditionalized on Change Recollection
For Day 2 recall, memory was better when change was recollected than when change was remembered but not recollected and when change was not remembered, leading to a significant main effect of Change Classification, $\chi^2(2) = 150.683, p < .001$. There was no significant main effect of Prediction Error Type, $\chi^2(1) = 0.345, p = .557$, and no significant Change Classification x Prediction Error Type interaction, $\chi^2(2) = 2.650, p = .266$. Post hoc tests revealed that within both prediction error types, memory was significantly better when change was recollected than when change was not remembered, smallest $z$ ratio = 4.015, $p < .001$, and when change was recollected than when change was remembered but not recollected, smallest $z$ ratio = 2.972, $p < .05$, but there was no significant difference in memory when change was not remembered and when change was remembered but not recollected, largest $z$ ratio = 1.530, $p = .6446$. 
There were fewer Day 1 intrusions when change was recollected than when change was remembered but not recollected and when change was not remembered, leading to a significant main effect of Change Classification, $\chi^2(2) = 32.022, p < .001$. There was a significant main effect of Prediction Error Type, $\chi^2(1) = 14.421, p < .001$, but there was not a significant Change Classification x Prediction Error Type interaction, $\chi^2(2) = 3.379, p = .185$. Post hoc tests revealed that for prediction error trials there were significantly more intrusions when change was not remembered than when change was recollected, $z_{ratio} = 5.389, p < .001$, significantly more intrusions when change was remembered but not recollected than when change was recollected, $z_{ratio} = 5.311, p < .001$, but no significant difference between when change was remembered but not recollected and when change was not remembered, $z_{ratio} = 0.601, p = .9910$. For no-prediction-error trials, there were no significant differences between any of the change classification groups.
Chapter 7: Discussion (Experiment 2)

As in Experiment 1, the results of Day 2 recall performance demonstrated that there was a benefit of repetition in viewing events in memory and that changing a critical feature resulted in proactive facilitation when the change was remembered and recollected, but proactive interference when the change was not recollected. This is also reflected in fewer Day 1 intrusions when change was recollected. In Experiment 2, where participants were instructed to make their predictions based on their memory for Day 1, we found a benefit of memory-based predictions on recall performance. Finally, we again found no benefit of experiencing prediction error on subsequent memory performance, despite the experimental design directly manipulating the experience of prediction error.
Chapter 8: General Discussion

In the present experiments, we examined the theoretical mechanisms underlying change processing, specifically the effects of experiencing prediction error and making memory-based predictions. We found that memory was better when retrieval was successful and changes were detected and recollected, but not because of the experience of prediction error. There was no evidence supporting a benefit or impairment of memory when prediction error was experienced during the prediction task in Day 2 viewing. There was a benefit of making memory-based predictions, indicating that successful retrieval during processing was associated with better subsequent memory.

Our primary goal was to investigate the direct effects of successful retrieval during processing and experiencing prediction error on subsequent memory for events. We did this by adapting the everyday changes paradigm from Wahlheim & Zacks (2019) to include a prediction task during Day 2 viewing. This allowed us to characterize successful retrieval and effectively control which activities were associated with prediction error and which were not, which created the opportunity to compare the two on memory performance. Previous research demonstrated that when change was detected and recollected, memory performance was enhanced through proactive facilitation, and when change was not recollected, memory was impaired through proactive interference (Wahlheim & Zacks, 2019). EMRC posits that successful change detection and recollection relies on retrieval of past events and experiencing prediction error during event comprehension. Our basic assumption was that when retrieval occurred and the theoretical construct of prediction error was experienced during processing, this would allow for change detection and recollection and would facilitate memory, while a failure to retrieve and not experiencing prediction error would not allow for change detection and recollection and
would impair memory. Both experiments demonstrated that successful retrieval during processing was associated with better subsequent memory. Experiment 1 showed that, when prediction error was a consequence of the activity type manipulation, it had no significant effect on memory performance. Experiment 2 directly manipulated the experience of prediction error, and still demonstrated no significant effect of prediction error on memory performance.

8.1 Prediction Error
EMRC combines EST (Zacks et al., 2007) and the memory-for-change framework (Jacoby et al., 2015) and theorizes that this retrieval of Day 1 features during ongoing processing allows for change detection by way of prediction error. EMRC takes from the memory-for-change that retrieval of similar past events when experiencing current events will facilitate memory, which was very strongly supported in these data, but it adds that the driving mechanism behind this retrieval process is prediction error, which was not supported in these experiments. Retrieval of previous events guides predictions for upcoming events, and any unexpected features due to changes lead to the experience of prediction error. This prediction error triggers updating of the current event model with the changed features, which are then integrated into a single trace with the Day 1 features and the experience of prediction error. Therefore, EMRC posits prediction error as the driving mechanism behind successful change detection and recollection. This led to the hypothesis that experiencing prediction error during Day 2 viewing would be associated with better memory performance.

Although the findings we report fail to support EMRC because prediction error was not associated with better memory performance, the experiments did not entirely reject prediction error as an important force behind change processing in everyday event memory. Previous research that has examined prediction in comprehension has utilized eye tracking as a means of
measuring predictions. Eisenberg, Zacks, & Flores (2018) demonstrated predictive looking during narrative activities with eye tracking by showing that viewers looked at objects that were about to be touched shortly before the objects were touched. Eye tracking is an effective and unobtrusive way of measuring predictions during ongoing processing, without halting a participant’s engagement.

In the present experiments, the prediction task did not involve unobtrusive means, but rather halted participants’ event processing and forced them to make an explicit prediction by choosing one of two options. The video playback had been stopped, removing the participant from the context of the narrative, and a conscious choice had to be made. Rather than allowing the participant to continue a natural interaction with the scene, the experiment stopped ongoing processing with the artificial prediction task. This experimental task does not accurately parallel the theoretical constructs of predictive processing and prediction error from EMRC. In normal event comprehension, people are not removed from the scene context and asked to make an explicit prediction as to what is going to happen in the immediate future, but rather are unaware of the continuous prediction process that underlies event processing. Therefore, it is possible that the lack of a benefit of prediction error on memory performance, which was contrary to hypothesis, is due to an inaccurate representation of the predictive process. The artificial prediction error that we induced may not parallel the prediction error posited in EMRC. With that in mind, future directions should focus on eye tracking, or other less intrusive means of measuring and manipulating prediction error without interrupting ongoing processing and comprehension.
8.2 Memory-Based Predictions

The present results supported previous findings that successful retrieval would benefit memory as Day 2 recall was improved and Day 1 intrusions were reduced when Day 1 features were recalled at test, with correct Day 1 recall, and during the prediction task, with memory-based predictions. The creation of integrated memory traces relies on the successful retrieval of earlier memories during processing. EMRC assumes that these traces are formed when Day 1 features are recollected during Day 2 viewing and directly compared with Day 2 features, allowing for the detection of change and the experience of prediction error to create the integrated trace that includes Day 1 and Day 2 as separate, but related, events with a temporal relationship. The present study examined these conscious recollections with memory-based predictions.

We quantified the conscious accessibility of Day 1 features during Day 2 viewing by characterizing memory-based and non-memory-based predictions. Memory-based predictions assumed the successful retrieval of Day 1 features and that these retrieved features then guided the Day 2 prediction. In Experiment 1, participants were instructed to predict what they thought was likely to happen. Since Experiment 1 made no a priori hypotheses regarding memory-based predictions, our characterization and analyses were post hoc examinations. We interpreted the prediction behavior such that Day 1 consistent predictions, or memory-based predictions, represented a successful conscious recollection of Day 1 features and the use of that recollection to guide predictions for Day 2. Therefore, memory-based predictions indicated that the Day 1 features were successfully retrieved and available for change detection and the integration into a single memory trace to facilitate memory. However, using memory-based predictions as a proxy for recollection of Day 1 features may be inappropriate as several behaviors could cause recollection to be misrepresented as a non-memory-based prediction or unsuccessful recollection.
as a memory-based prediction. Participants in Experiment 1 were not instructed to make their predictions based off of Day 1 or to use Day 1 features in any way to guide their Day 2 predictions, so participants could have recollected Day 1 features but chose to predict the opposite ending because they knew some activities were going to change. There is evidence to support this possible interpretation as, in Experiment 1, when participants were not instructed to make Day 2 predictions based off their memory for Day 1, there was a 65% rate of memory-based predictions. In Experiment 2, when this instruction was included, this rate rose to 75% memory-based predictions. While we found a relationship between memory-based predictions and Day 2 recall with our post hoc analysis of individual differences in Experiment 1, this lack of instruction to utilize Day 1 memories could potentially account for the lack of a significant benefit of memory-based predictions, while Experiment 2 did report a significant associated benefit for memory. Another potential artifact that both experiments were likely subject to was participant guessing. On activities where participants could not successfully recollect Day 1 and simply made their prediction by guessing the Day 2 ending, they had a 50% chance of guessing “correctly” and choosing the ending that was consistent with Day 1. This guess would then be misinterpreted as a memory-based prediction when, in fact, there was no Day 1 recollection. These misinterpretations have the potential to mask the true effect of successful recollection of Day 1 features during Day 2 viewing on subsequent memory performance. A more representative proxy of Day 1 retrieval and recollection is necessary to further elucidate the nature of this effect but finding support for memory-based predictions facilitating memory despite these potential shortcomings illustrates the robustness of the beneficial effects of retrieval during processing.
8.3 Conclusion
The present studies examined underlying theoretical mechanisms behind change processing and memory with naturalistic activities. These findings, by demonstrating support for one of the mechanisms and failing to find support for another, open the door for continued research in event cognition. These results also bring to light potential shortcomings of current experimental approaches and call for new paradigms to better parallel the theoretical process. Future directions should focus on paradigms that more accurately represent the theoretical constructs to help further develop a representative framework to explain how people recognize and process change.
References


Tables and Figures

Table 1. Recall of Novel Activities as a Function of Prediction Error Type (Experiments 1 & 2)

<table>
<thead>
<tr>
<th>Prediction Error Type</th>
<th>Prediction Error</th>
<th>No Prediction Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td>0.583 [.428, .723]</td>
<td>0.638 [.488, .765]</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>0.209 [.108, .367]</td>
<td>0.372 [.220, .555]</td>
</tr>
</tbody>
</table>

Notes. 95% confidence intervals are presented in the brackets.
Table 2. Recall and Intrusions as a Function of Activity Type and Memory-Based Predictions

<table>
<thead>
<tr>
<th></th>
<th>Activity Type</th>
<th>Changed</th>
<th>Repeated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 2 Recall</td>
<td>Memory-Based</td>
<td>0.586 [.465, .697]</td>
<td>0.747 [.642, .830]</td>
</tr>
<tr>
<td></td>
<td>Non-Memory-Based</td>
<td>0.587 [.452, .710]</td>
<td>0.675 [.546, .782]</td>
</tr>
<tr>
<td>Day 1 Intrusions</td>
<td>Memory-Based</td>
<td>0.165 [.115, .231]</td>
<td>0.068 [.043, .107]</td>
</tr>
<tr>
<td></td>
<td>Non-Memory Based</td>
<td>0.675 [.546, .782]</td>
<td>0.089 [.053, .146]</td>
</tr>
</tbody>
</table>

Notes. 95% confidence intervals are presented in the brackets.
Table 3. Probabilities of Change Recollection as a Function of Activity Type (Experiments 1 & 2)

<table>
<thead>
<tr>
<th>Activity Type</th>
<th>Changed</th>
<th>Novel</th>
<th>Repeated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td>0.659 [.581, .730]</td>
<td>0.292 [.227, .366]</td>
<td>0.088 [.061, .125]</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>0.429 [.356, .505]</td>
<td>0.264 [.195, .346]</td>
<td>0.183 [.143, .232]</td>
</tr>
</tbody>
</table>

Notes. Probabilities for changed activities are correct change classifications while probabilities for novel and repeated activities are false change classifications. 95% confidence intervals are presented in the brackets.
Table 4. Recall as a Function of Activity Type and Memory-Based Predictions (Experiment 2)

<table>
<thead>
<tr>
<th>Activity Type</th>
<th>Changed</th>
<th>Repeated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 2 Recall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory-Based</td>
<td>0.370 [.279,.471]</td>
<td>0.480 [.379,.582]</td>
</tr>
<tr>
<td>Non-Memory Based</td>
<td>0.283 [.188,.402]</td>
<td>0.359 [.251,.484]</td>
</tr>
</tbody>
</table>

Notes. 95% confidence intervals are presented in the brackets.
Figure 1. Narrative Video Endings

Figure 1: Example still frames from activities. The top image depicts the beginning cue that was identical across all activities and the bottom images depict the two possible endings, differing by the critical features.
Figure 2. Probabilities of Day 2 Recall

*Figure 2. Probabilities of Day 2 recall. Conditional points (green and red) refer to the change recollection probabilities. Point areas represent the proportion of observations contributing to each point. Error bars are 95% confidence intervals. Bolded text represents the experimental manipulation on activity type.*
Figure 3. Probabilities of Day 1 Intrusions

Figure 3. Probabilities of Day 1 intrusions. Conditional points (green and red) refer to the change recollection probabilities. Point areas represent the proportion of observations contributing to each point. Error bars are 95% confidence intervals. Bolded text represents the experimental manipulation on activity type.
Figure 4: The difference in Day 2 Recall memory performance on Memory-Based Prediction (MBP) trials and Non-Memory-Based Prediction (NMBP) trials as a function of the individual’s average tendency to make Memory-Based Predictions, for each participant.
Figure 5: Breakdown of Experiment 2 Conditions

<table>
<thead>
<tr>
<th>Recurring</th>
<th>Novel</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prediction Error</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Memory-Based</strong></td>
<td><strong>Non-Memory-Based</strong></td>
</tr>
<tr>
<td>Day 1</td>
<td>Day 1</td>
</tr>
<tr>
<td>Clock</td>
<td>Clock</td>
</tr>
<tr>
<td><strong>Changed</strong></td>
<td><strong>Reverted</strong></td>
</tr>
<tr>
<td>Clock</td>
<td>Phone</td>
</tr>
<tr>
<td>Clock</td>
<td>Clock</td>
</tr>
<tr>
<td><strong>Repeated</strong></td>
<td><strong>Changed</strong></td>
</tr>
</tbody>
</table>

Figure 5: Breakdown of Experiment 2 conditions by Prediction Error Condition (Prediction Error, No Prediction Error) and participant Memory Condition (Memory-Based, Non-Memory-Based) with examples of how the condition would be manifested.
Figure 6. Probabilities of Day 2 Recall

Figure 6. Probabilities of Day 2 recall. Conditional points (green and red) refer to the change recollection probabilities. Point areas represent the proportion of observations contributing to each point. Error bars are 95% confidence intervals. Bolded text represents the experimental manipulation on prediction error type.
Figure 7. Probabilities of Day 1 Intrusions

Figure 7. Probabilities of Day 1 intrusions. Conditional points (green and red) refer to the change recollection probabilities. Point areas represent the proportion of observations contributing to each point. Error bars are 95% confidence intervals. Bolded text represents the experimental manipulation on prediction error type.