Research Article

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
1–17

DOI: 10.1177/09567976211053596
www.psychologicalscience.org/PS
SAGE

Understanding Everyday Events: Predictive-Looking Errors Drive Memory Updating

Christopher N. Wahlheim[1], Michelle L. Eisenberg[2], David Stawarczyk[2,3], and Jeffrey M. Zacks[2]

[1]Department of Psychology, University of North Carolina at Greensboro; [2]Department of Psychological and Brain Sciences, Washington University in St. Louis; and [3]Department of Psychology, Psychology and Neuroscience of Cognition Research Unit, University of Liège

Abstract

Memory-guided predictions can improve event comprehension by guiding attention and the eyes to the location where an actor is about to perform an action. But when events change, viewers may experience predictive-looking errors and need to update their memories. In two experiments (*N*s = 38 and 98), we examined the consequences of mnemonic predictive-looking errors for comprehending and remembering event changes. University students watched movies of everyday activities with actions that were repeated exactly and actions that were repeated with changed features—for example, an actor reached for a paper towel on one occasion and a dish towel on the next. Memory guidance led to predictive-looking errors that were associated with better memory for subsequently changed event features. These results indicate that retrieving recent event features can guide predictions during unfolding events and that error signals derived from mismatches between mnemonic predictions and actual events contribute to new learning.

Keywords

action observation, event cognition, memory updating, mnemonic prediction error, predictive looking, open data, open materials

Received 1/14/21; Revision accepted 9/9/21

Predicting the actions of other people is essential for skilled social interactions. For example, to cooperate in preparing dinner, one may need to anticipate when one's partner will use the sink to wash vegetables. Prediction can improve action comprehension and facilitate one's own action planning (Gredebäck, 2018; Gredebäck & Falck-Ytter, 2015). Anticipating the future on the basis of the past is adaptive because people often repeat behaviors. But when situations change, people behave differently. If the partner in the example above purchased prewashed vegetables on a later occasion, they may bypass washing them in the sink, leading the observer to experience a mnemonic prediction error. Viewers' comprehension of such unexpected action changes may require registering such mismatches between predicted and actual events.

Action-observation studies have shown that adult viewers make on-line predictions about future actions, especially actions they have experienced. Viewers look ahead to contacted objects (Eisenberg et al., 2018; Flanagan & Johansson, 2003; Hayhoe et al., 2003; Land & McLeod, 2000) and learn to predict object trajectories from repeated sequences (Barnes et al., 2005; Diaz et al., 2013). Infants learn to predict manual actions that are repeatedly performed (Cannon et al., 2012; Falck-Ytter et al., 2006), show positive correlations between predictive looking and their experience performing the actions (Cannon et al., 2012; Gredebäck et al., 2018; Melzer et al., 2012), and look earlier when observing recently learned actions (Gerson & Woodward, 2014). Thus, memory for past actions can guide action prediction and comprehension.

When actions change across occasions, mnemonic predictions trigger surprise responses that may stimulate new learning and update predictions for future

Corresponding Author:
Christopher N. Wahlheim, University of North Carolina at Greensboro, Department of Psychology
Email: cnwahlhe@uncg.edu

actions (Gredebäck et al., 2018). Infant eye-tracking studies show longer looking times to events that violate their expectations about the world (Gredebäck et al., 2018; Juvrud et al., 2019; Stahl & Feigenson, 2015). Infants and children also show greater pupil dilation when actions end unexpectedly (Gredebäck et al., 2018; Juvrud et al., 2019). Such prediction errors increase exploration and new learning. For example, infants who watched a ball unexpectedly roll through a wall later explored the ball more than infants who had seen a ball stop. When infants were taught that the ball squeaked, infants whose expectations had been violated better remembered the new property of the ball (Stahl & Feigenson, 2015). These findings show that when an action violated infants' expectations about the world, infants actively updated their understanding of the object's properties.

Conditions associated with prediction errors are also associated with memory updating. When learned image sequences later include unexpected images that trigger prediction errors, memory is reduced for expected but no-longer-relevant images (Kim et al., 2014). In addition, after people learn to associate consistently valanced words with a scene category, pairing similar scenes with oppositely valanced words leads to better memory (Greve et al., 2017). Similarly, switching sequential contingencies such that symbol cues are first followed by objects from one semantic category and then another enhances recollection for unexpected objects (Kafkas & Montaldi, 2018). Also, prediction errors based on memory for virtual scenes enhance subsequent memory, and the benefit increases with the number of changed features (Bein et al., 2020). When people view action sequences, interrupting repetitions before the expected outcomes distorts their memory for prior events, indicating a form of error-driven memory updating (Sinclair & Barense, 2018). Relatedly, when watching basketball, people's prediction errors and subsequent memories are associated with increased pupil size and cortical-activity shifts (Antony et al., 2021).

When viewers see an actor perform everyday activities that change over time, prediction errors may stimulate better encoding. In one series of experiments, viewers watched an actor perform action sequences on two occasions on which some actions changed (Wahlheim & Zacks, 2019). For example, the actor initially retrieved a bath towel from the closet, but she later opened the closet in the same way and retrieved a hand towel. Memory for changed actions was better when changes were noticed and later recollected. Also, greater neural reinstatement of the original actions before viewing changed actions was associated with better subsequent memory (Stawarczyk et al., 2020).

Statement of Relevance

Everyday life is filled with changes: Friends change their food preferences, construction projects change roads and buildings, workers come and go from businesses. To remember accurately what happened recently and avoid confusing different events, people have to track these changes. In this research, we used a combination of eye tracking and memory tests to study how people update their memories when things change. We found that when watching a new activity, people used memories of recent events to look ahead to where an actor might reach next. However, when things changed, this led to prediction errors (i.e., looking to the wrong location). Both of these factors—using memory to help with comprehension of new information and making predictions about the near future—may be important for accurately encoding new events. In particular, our results suggest that prediction errors drive new learning. These findings also suggest that, outside the laboratory, directing viewers' attention to instances when events change may help them better track and remember changes.

This suggests that anticipating repeated actions triggered updating for changes. Although neural reinstatement may be necessary for mnemonic predictions (Bein et al., 2020), this measure alone does not definitively index predictive processing.

In the current study, we assayed mnemonic prediction error in this *everyday-changes* paradigm using anticipatory eye movements as a converging measure. In the first session, participants watched two movies of an actor performing everyday-activity sequences depicting two fictional days in her life (Day 1 and Day 2 movies). *Changed activities* started with the same actions in both movies but ended with different actions in each movie, whereas *repeated activities* started and ended with the same action in both movies. Experiment 2 also included control activities with actions that were performed only in the second movie. Eye movements were recorded while participants watched the movies. In the second session, 1 week later, participants recalled Day 2 activity features (i.e., contacted objects), indicated whether the contacted objects had changed between movies, and if so, recalled the Day 1 features.

Changed Day 2 features are better remembered when participants can recall that those features had changed and can recall the original Day 1 features (Garlitch & Wahlheim, 2021; Hermann et al., 2021; Stawarczyk et al., 2020; Wahlheim & Zacks, 2019). According to *event-memory-retrieval-and-comparison* theory (Wahlheim & Zacks, 2019), the memory benefit for changed Day 2 features associated with recall of Day 1 features occurs partly because mnemonic prediction errors stimulate encoding of new features. Event-memory-retrieval-and-comparison theory proposes a causal cascade in which viewers may (a) retrieve the activity's previous ending, (b) use the remembered action to predict the Day 2 ending, (c) experience a mnemonic prediction error when they view an unexpected action change, and (d) update their memory in response to that error. This view is consistent with the findings of error-driven associative learning (Rescorla & Wagner, 1972).

The present two experiments tested the hypothesis that while viewing the beginnings of changed activities, people's predictive looking based on memory for Day 1 endings should be associated with better subsequent memory for changes, recall of Day 1 features, and memory for the changed Day 2 features. We tested this against the hypothesis that predictive processing is not necessary for changed features to be better remembered when they are recollected as such. This could occur when viewers retrieve Day 1 activities after viewing changed Day 2 activity endings because postdictive processes (cf. Neely et al., 1989) may be used to compare activities and improve the encoding of changed activity endings.

Method

These experiments were approved by the institutional review board of Washington University in St. Louis. Participants were recruited from the Department of Psychological and Brain Sciences at Washington University in St. Louis. Participants self-enrolled in both of the following experiments through an online recruitment system. All participants received course credit or $10 per hour as compensation for their time.

Participants

In Experiment 1, we set the sample size on the basis of previous studies using the everyday-changes paradigm to measure memory effects (Wahlheim & Zacks, 2019; $N = 36$) because no relevant previous data were available to support a power analysis. We recruited 43 participants and excluded five, either because we could not track their eyes ($n = 4$) or because of attrition ($n = 1$). The final sample included 38 participants (13 women; age: range = 18–27 years, $M = 20.37$, $SD = 2.17$). In Experiment 2, we ran bootstrapping power analyses using custom codes in R software (Version 4.1.1; R Core Team, 2021) for each of the primary hypotheses using the data from Experiment 1. We sampled Experiment 1 data with replacement to obtain varying numbers of simulated participants. We then ran mixed-effects models for the results of interest and determined whether they were significant. We ran 1,000 iterations for the recall models and 500 iterations for the eye-tracking models. We determined the proportion of significant results, which provided the estimates for power. The results indicated that a sample of 90 participants would be enough to achieve 80% power. Consequently, we recruited 111 participants, anticipating that some would have to be excluded. We excluded 13 participants either because we could not track their eyes ($n = 12$) or because the equipment failed ($n = 1$). The final sample included 98 participants (56 women; age: range = 18–25 years, $M = 19.61$, $SD = 1.39$).

Materials and design

Both experiments included the same two movies showing a female actor performing sequences of everyday activities in or around her home and workplace on two fictional days in her life. There were two versions (A and B) of each activity that included the same peripheral features and a changed central feature that always showed her manipulating an object (e.g., inserting a key into a door lock). Figure 1a shows key moments from two versions of an example activity in which the actor unlocked her front door to enter her home. In Version A, she unlocked the doorknob, whereas in Version B, she unlocked the deadbolt. The different central features in each version are displayed at the point of contact.

Experiment 1 used a single-factor, two-level (activity type: repeated vs. changed) within-subjects design to manipulate the relationship between activities in each movie. Each movie included 44 critical activities (22 per condition) and 20 filler activities (10 per condition) that served to maintain narrative continuity (64 activities total). We counterbalanced the assignment of activity versions across activity types and movies in the following way. First, we created a Day 1 movie that included all critical activities assigned to either Version A or Version B. No more than three activities of the same version appeared sequentially. Then we created another Day 1 movie, switching all the activity versions (i.e., Version A activities became Version B activities, and vice versa). Finally, we created two Day 2 movies—one that included all Version A activities and one that included all Version B activities. The combinations of the Day 1 and Day 2 movies produced four experimental formats.

Experiment 2 included the same repeated and changed conditions as Experiment 1 but also included a

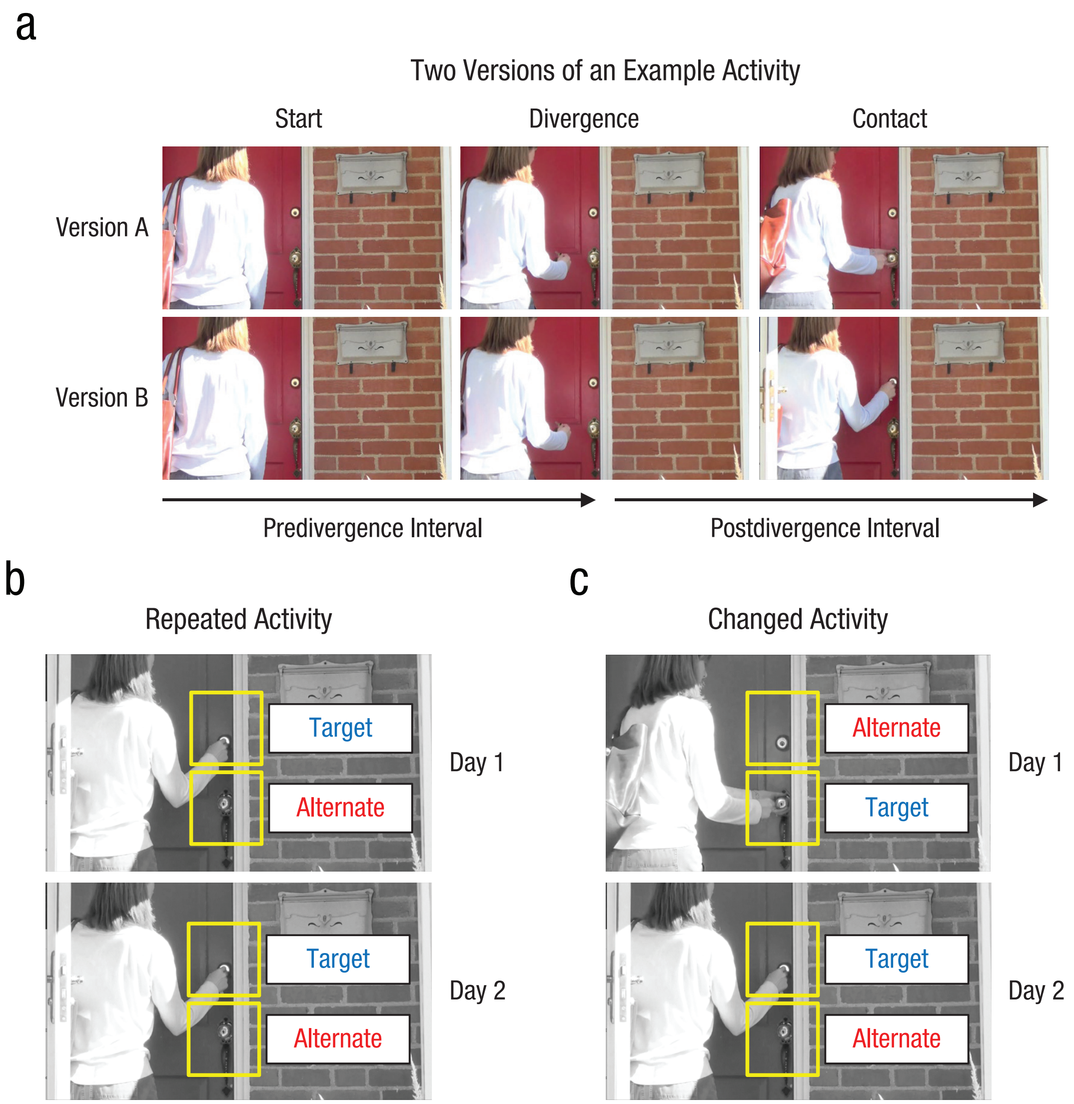

Fig. 1. Example activities and interest areas. Key moments from two versions of an example activity, the actor unlocking a door, are shown in (a). The interest period for eye-tracking analysis began as she approached the door and ended when she first contacted either the doorknob (Version A) or the deadbolt (Version B). The first images (Start) show the start of the interest period, the second images (Divergence) show the moment before the activity versions diverge, and the third images (Contact) show the two possible contacted objects. Example interest areas for a repeated activity and for a changed activity are shown in (b) and (c), respectively. Yellow boxes highlight interest areas for target (blue) and alternate (red) objects. The designation of target and alternate objects was the same in both movies for repeated activities and switched from the Day 1 to the Day 2 movie for changed activities. The images are shown in black and white here to enhance the visual contrast of the yellow boxes showing the interest areas. The complete movies were shown in color and were not altered to indicate interest areas. The movies can be viewed at OSF (https://osf.io/w8skp/).

control condition with activities that appeared only in the Day 2 movie. Thus, it used a three-level (activity type: repeated vs. control vs. changed) within-subjects design. We assigned twice the number of activities to the changed condition than to the other conditions to focus power on changed activities, which were of primary theoretical interest. The Day 1 movies included 33 critical activities (22 changed and 11 repeated) and 20 filler activities (10 changed and 10 repeated) for a total of 53 activities. The Day 2 movies included 44 critical activities (22 changed, 11 repeated, and 11 control) and 20 filler activities (10 changed and 10 repeated) for a total of 64 activities. The activities were counterbalanced in the same way as in Experiment 1.

Both experiments used the same cued-recall test items. The test included 64 questions probing memory

Fig. 2. General schematic of events in each experimental session. Participants watched two movies successively during the first session (a). They were instructed to pay attention to the actor's activities in both movies. After a delay, participants completed a second session that included a cued-recall test of features from both movies (b). For each activity, participants first tried to recall the Day 2 feature. Then they were asked to classify whether the activity had changed from Day 1 to Day 2. When participants indicated that an activity had changed, they tried to recall the Day 1 feature and then moved on to the next trial (blue arrows). When participants indicated that an activity had not changed, the program advanced to the next test trial (red arrow). The test instructions differed slightly between experiments (see the Supplemental Material available online).

for all activities from the Day 2 movie. Cues asked about the central features that could have changed. For example, the cue for the door-unlocking example in Figure 1 asked about the central lock feature (i.e., "Which lock did the actor unlock to enter her home?").

Procedure

Figure 2 displays a schematic of the two experimental sessions, which were separated by approximately 1 week (Experiment 1: M = 7.07 days, SD = 0.81 days, range = 7–11 days; Experiment 2: M = 7.07 days, SD = 0.50 days, range = 6–10 days). Both experiments used the same procedure, with one exception noted below. The exact instructions shown to participants are available in the Supplemental Material available online. Movies were shown on a 19-in. monitor (1,440- × 900-pixel resolution) at a 1,280- × 720-pixel aspect ratio using *Experiment Builder* software (Version 2.3.38; SR Research, 2020). Gaze location was recorded from the right eye using an infrared pupil-corneal eye tracker (EyeLink 1000; SR Research, Mississauga, Ontario, Canada) that sampled at 1000 Hz. Participants placed their heads against a chin and forehead rest to minimize motion. The camera was positioned 52 cm from the top of the rest. The viewing distance was 58 cm from the rest, and the viewing angle was 38.6°.

During Session 1, participants were told that they would watch movies of an actor performing everyday activities and to pay attention to those activities. Participants watched two different movies from ostensibly separate days in the actor's life. As described above, half of the Day 2 activities included changed endings. While watching each movie, participants took short breaks between scenes showing morning, work, afternoon, and evening activities.

During Session 2, participants completed the cued-recall test presented via E-Prime software (Version 2.0; Schneider et al., 2012) in another room. They were told that they would recall activity features and indicate which features had changed between movies. Before the test, participants viewed two example clips of a hair-styling activity with a changed feature. The actor styled her hair first with a comb and then with a brush. Test items then appeared individually (for a schematic of the test trial structure, see Fig. 2b). On each trial, participants first recalled the Day 2 feature by typing their response. Then they indicated whether the feature had changed by responding "yes" or "no" with the "1" or "2" key (Experiment 1) or by classifying the activity as repeated, changed, or shown only on Day 2 with the "1," "2," or "3" key (Experiment 2). When participants indicated that an activity had changed ("yes" or "changed"), they attempted to recall the Day 1 feature by typing their response. When participants indicated that an activity had not changed, the program advanced to the next trial.

Statistical analysis

All analyses were conducted using R software. We fitted linear and logistic mixed-effects models including experimental variables as fixed effects and participants and activities (items) as random effects using functions from the *lme4* package (Version 1.1.27.1; Bates et al., 2015). We tested for significant effects of predictor variables using the Wald test in the "Anova" function from the *car* package (Version 3.0.10; Fox & Weisberg, 2019). We conducted pairwise comparisons using the "emmeans" function from the *emmeans* package (Version 1.4.7; Lenth, 2020), controlling for multiple comparisons with the Tukey method. All statistical tests were two-sided. The level for significance was set at an α of .05. The probabilities and confidence intervals below were estimated from these models.

Results

Each activity included a pair of critical objects that the actor might manipulate, such as the doorknob and deadbolt in Figure 1a. For repeated activities, the actor manipulated the same target object on both days—for example, she could unlock the deadbolt on both days (Fig. 1b). In that case the alternate object on both days would be the doorknob. For changed activities, the target object in the Day 1 movie became the alternate object in the Day 2 movie and vice versa (Fig. 1c). Both repeated and changed activities appeared in both experiments. Control activities that the actor performed only in the Day 2 movie also appeared in Experiment 2. The primary eye-tracking measure of interest was the proportion of time participants looked to the target and alternate objects while watching each activity. We report results from analyses comparing these proportions across activity types and movies from both experiments together. We also conducted exploratory analyses of pupil areas, which we report in the Supplemental Material.

Proportions of looks to interest areas

To characterize looking patterns as activities unfolded, C. N. Wahlheim and M. L. Eisenberg first divided each activity into two intervals of interest (Fig. 1a). The raters watched each version of each activity together and jointly decided about interval placement. The *predivergence* interval started when the target and alternate object were both visible and ended the moment before the actor started to move more toward the target than the alternate object (e.g., the first moment that her reach trajectory indicated that she was more likely to contact the deadbolt than the doorknob). The *postdivergence* interval started the moment after this divergence point and ended when the actor contacted the target object. Total interest intervals ranged from 470 ms to 29,200 ms (M = 6,434 ms, SD = 5,032 ms). Predivergence intervals (range = 100–16,920 ms, M = 5,212 ms, SD = 3,597 ms) were significantly longer than postdivergence intervals (range = 140–13,250 ms, M = 1,222 ms, SD = 2,447 ms), $t(87) = 10.61$, $p < .001$.

Because the pre- and postdivergence intervals varied in length, we divided each interval into 10 isochronous bins for analysis (20 total bins). This allowed us to compare the time course of looking behavior in terms of the relative time elapsed from the onset of each activity to the divergence point and from the divergence point to object contact, controlling for differences in duration.

To analyze looking patterns as activities unfolded, we included time bin as a categorical predictor variable in the mixed-effects models. When testing hypotheses about the effects of prior viewing on predictive looking before and after divergence points, we used separate models that included only either the first 10 bins (predivergence interval) or the second 10 bins (postdivergence interval). For each activity, we defined spatial regions of interest by drawing polygons around the target and alternate objects (Figs. 1b and 1c). We then recorded, for each participant, whether they looked in one or both interest areas during each time bin. In the following analyses, we report looking proportions in each interest area aggregated across all activities. In addition to the analyses reported here, we examined the consistency of looking-proportion differences across activities based on pre- and postdivergence interval lengths (see the Supplemental Material).

Repeated activities. For repeated activities, the actor manipulated the same object on both days, so mnemonic predictions based on Day 1 actions should lead to more looking to Day 2 target objects. This was what we found (Fig. 3, top left graph for each experiment). In both experiments, logistic mixed-effects models were used to estimate proportions of looking to each object of interest (object) during both movies (day) at each time bin (bin). Models including object, day, and bin as fixed effects

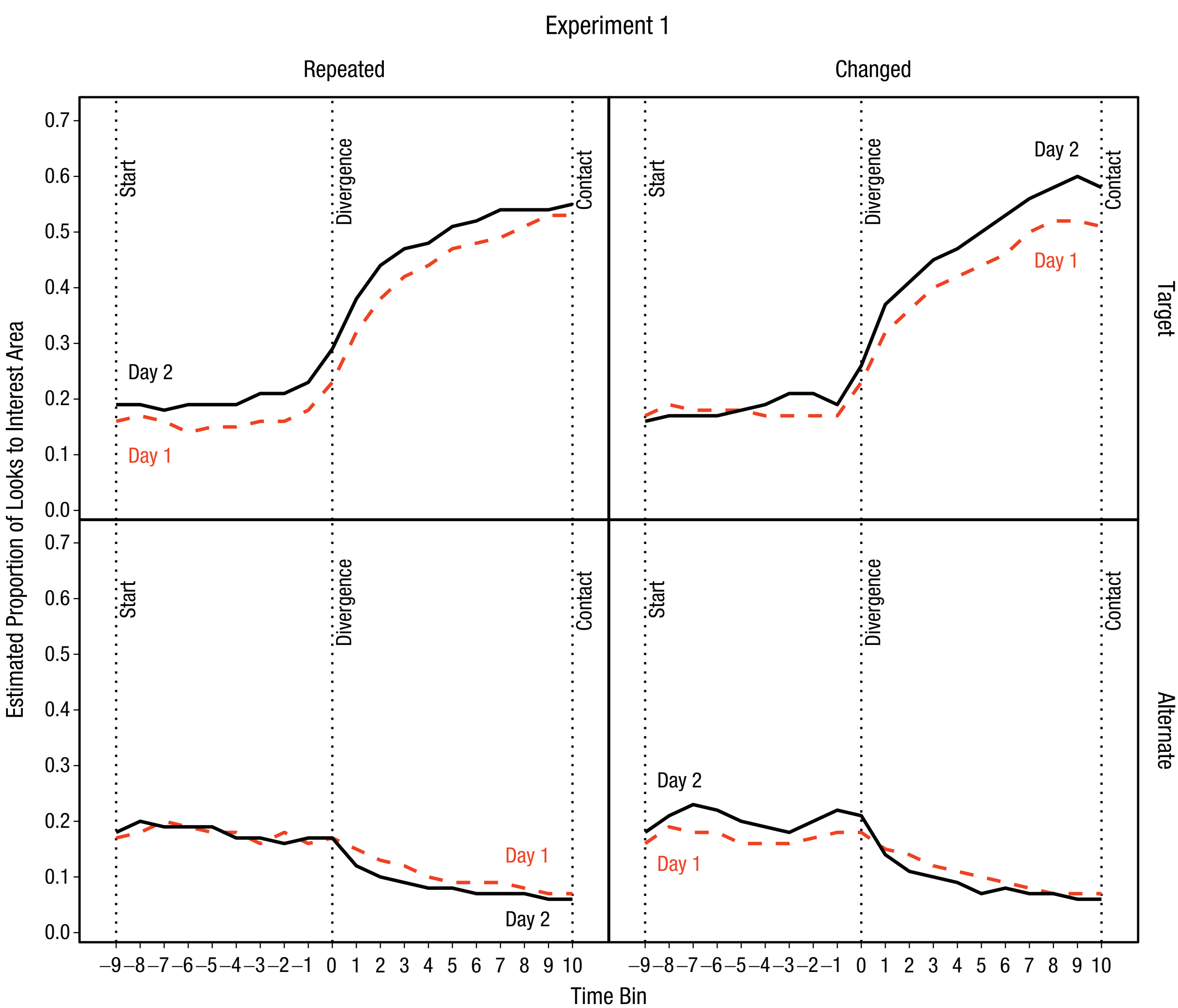

Fig. 3. *(continued on next page)*

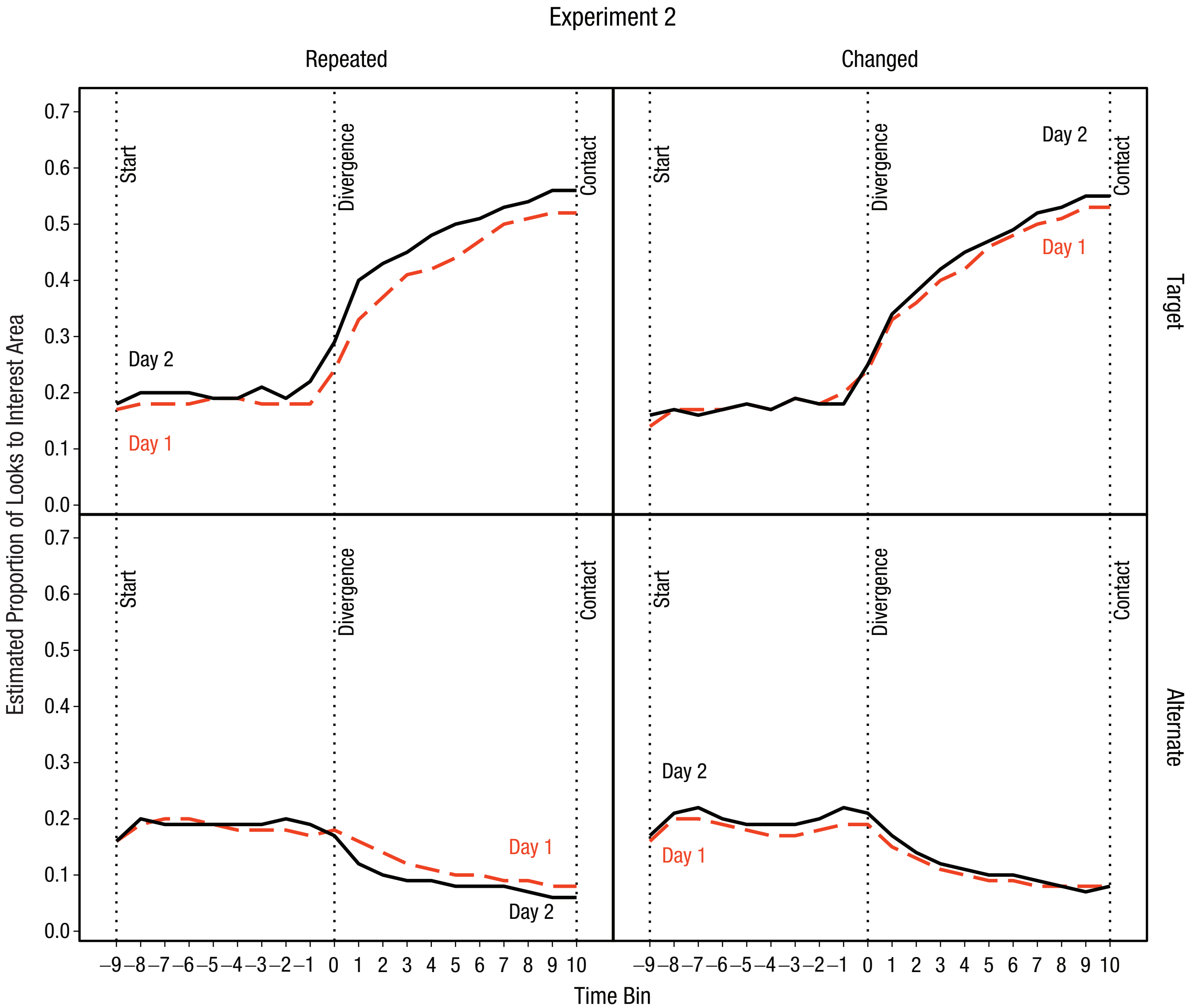

Fig. 3. Looks to target and alternate objects during movie viewing. Model-estimated proportions of looking to interest areas including target objects (top row) and alternate objects (bottom row) during interest periods are shown for critical repeated activities (left columns) and changed activities (right columns), separately for each time bin in Experiments 1 and 2. "Start" indicates the moment the interest period began. "Divergence" indicates the point of divergence when the two activity versions first became distinguishable. "Contact" indicates the moment the actor contacted the target or alternate object. Red dashed lines indicate looking proportions during Day 1 movie viewing. Black solid lines indicate looking proportions during Day 2 movie viewing.

(Table 1) indicated significant Object × Day interactions. Participants looked significantly more to target objects when watching the Day 2 than the Day 1 movie, smallest z ratio = 9.39, $p < .001$. In contrast, they looked significantly less to alternate objects when watching the Day 2 than the Day 1 movie, smallest z ratio = 2.03, $p = .043$.

The models in both experiments also indicated significant Object × Bin interactions showing the following patterns. In the predivergence intervals (time bins –9 to 0; from "start" to "divergence"), looking proportions were mostly consistent across adjacent bins. Pairwise comparisons indicated significant increases only for target objects in both experiments from bin –1 to 0, smallest z ratio = 4.83, $p < .001$. In the postdivergence intervals (time bins 1 to 10; from the bin after "divergence" to "contact"), looking patterns differed for target and alternate objects. For target objects, looking proportions increased consistently until Bin 6 (Experiment 1) and Bin 7 (Experiment 2) and then began to reach an asymptote. In contrast, for alternate objects in both experiments, there were overall decreases in looking proportions across the interest periods. Looking

Table 1. Looking-Model Results for Each Activity Type: Experiments 1 and 2

	Experiment 1			Experiment 2		
Activity and effect	χ^2	*df*	*p*	χ^2	*df*	*p*
Repeated activity						
Object	5,216.37	1	< .001	5,956.10	1	< .001
Day	31.31	1	< .001	23.44	1	< .001
Bin	1,387.58	19	< .001	1,406.82	19	< .001
Object × Day	73.41	1	< .001	73.88	1	< .001
Day × Bin	7.76	19	.989	7.53	19	.991
Object × Bin	4,885.20	19	< .001	5,353.73	19	< .001
Object × Day × Bin	17.88	19	.530	42.82	19	.001
Changed activity						
Object	4,760.08	1	< .001	9,733.85	1	< .001
Day	59.77	1	< .001	33.70	1	< .001
Bin	1,336.73	19	< .001	3,092.08	19	< .001
Object × Day	24.99	1	< .001	0.75	1	.386
Day × Bin	10.24	19	.947	3.26	19	.999
Object × Bin	5,100.31	19	< .001	11,544.90	19	< .001
Object × Day × Bin	79.34	19	< .001	28.79	19	.069

Note: The model results shown here correspond to the data displayed in Figure 3.

proportions were significantly greater in the first bin (1) than the last bin (10), smallest z ratio = 7.41, $p < .001$. However, the decrease in looking was sharpest across the three bins after the divergence point, as there were few significant differences beyond Bin 3.

There was also an Object × Day × Bin interaction in Experiment 2. To characterize this interaction, we fitted separate models with the same variables to data including only the time bins from pre- or postdivergence intervals. In the predivergence model (Table 2), an Object × Day interaction indicated significantly greater proportions of looking to target objects in the Day 2 than the Day 1 movie, z ratio = 4.76, $p < .001$, but no significant difference in the proportions of looking to alternate objects in each movie, z ratio = 1.19, $p = .234$. In the postdivergence interval, an Object × Day interaction indicated significantly greater proportions of looking to target objects in the Day 2 than the Day 1 movie, z ratio = 8.55, $p < .001$, and to alternate objects in the Day 1 than the Day 2 movie, z ratio = 4.33, $p < .001$. These results show that viewing experience increased predictive looking to objects the actor would contact early in each activity.

Changed activities. For changed activities, the actor manipulated one object in the Day 1 movie but then a different object in the Day 2 movie. The objects that were manipulated on Day 1 but not Day 2 are referred to as *alternate objects* for Day 2 (see Fig. 1c). During the early part of a changed activity on Day 2, mnemonic predictions based on actions from Day 1 movies should increase looking to the alternate object because that object was previously the action target. After the divergence point, as visual cues provide information that the activity will end differently, we expected looking to the target to increase. Figure 3 shows that this qualitative pattern was observed in both experiments.

For Experiment 1, a model with the same variables as in the looking-proportion model for repeated activities (Table 1) indicated a significant Object × Day × Bin interaction. This interaction was characterized using separate models for the pre- and postdivergence intervals. In the predivergence interval (Table 2), a significant Object × Day interaction showed greater proportions of looking to both objects when participants watched the Day 2 movie than the Day 1 movie, and this difference was greater for alternate objects, z ratio = 6.01, $p < .001$, than for target objects, z ratio = 1.99, $p = .047$. These results suggest that memory for actions in Day 1 movies guided participants' looking toward areas including objects that could be contacted in Day 2 movies, especially for areas including the earlier-contacted objects. In the postdivergence interval, a significant Object × Day interaction indicated greater proportions of looking to target objects in the Day 2 movie than the Day 1 movie, z ratio = 10.00, $p < .001$, and to alternate objects in the Day 1 movie than the Day 2 movie, z ratio = 2.37, $p = .018$. These results suggest that observed movement trajectories that contradicted mnemonic predictions stimulated more

Table 2. Looking-Model Results for Repeated Activities in Experiment 2 and Changed Activities in Experiment 1

Interval and effect	χ^2	*df*	*p*
Experiment 2: repeated activity			
Predivergence interval			
Object	19.93	1	< .001
Day	17.74	1	< .001
Bin	55.97	9	< .001
Object × Day	6.45	1	.011
Day × Bin	7.54	9	.581
Object × Bin	65.84	9	< .001
Object × Day × Bin	8.81	9	.455
Postdivergence interval			
Object	9,415.61	1	< .001
Day	8.93	1	.003
Bin	89.84	9	< .001
Object × Day	83.43	1	< .001
Day × Bin	1.27	9	.999
Object × Bin	401.91	9	< .001
Object × Day × Bin	7.67	9	.568
Experiment 1: changed activity			
Predivergence interval			
Object	0.20	1	.651
Day	31.78	1	< .001
Bin	43.95	9	< .001
Object × Day	8.09	1	.004
Day × Bin	6.43	9	.697
Object × Bin	35.33	9	< .001
Object × Day × Bin	10.85	9	.286
Postdivergence interval			
Object	8,035.43	1	< .001
Day	29.01	1	< .001
Bin	117.62	9	< .001
Object × Day	76.63	1	< .001
Day × Bin	3.35	9	.949
Object × Bin	517.59	9	< .001
Object × Day × Bin	0.92	9	.999

Note: The model results shown here correspond to the data displayed in Figure 3.

looking to objects that had not been contacted in the Day 1 movie.

In both experiments, there were also significant Object × Bin interactions indicating the following patterns. In the predivergence intervals, looking proportions were mostly consistent across adjacent bins, with some exceptions. For target objects in both experiments, there were significant increases from bin −1 to 0, smallest z ratio = 5.19, $p < .001$. For target objects in Experiment 2, the first bin (−9) was significantly lower than most other bins, smallest z ratio = 3.17, $p = .049$. For alternate objects in Experiment 2, there were significant increases in bins −8, −7, −1, and 0, smallest z ratio = 3.20, $p = .045$. In the postdivergence intervals, there were substantial differences in proportions of looking to target and alternate objects. For targets in both experiments, there were generally consistent increases in looking across adjacent bins until Bin 7. In contrast, for alternate objects in both experiments, looking proportions decreased as activities unfolded. Looking proportions were significantly greater in the first bin (1) than the last bin (10), smallest z ratio = 8.65, $p < .001$. However, the decrease in looking to alternate objects was sharpest in the four bins after the divergence point, as there were few significant differences beyond Bin 4.

Finally, in Experiment 2, there was a significant effect of day, replicating the finding from Experiment 1 that proportions of looking to both objects were greater during Day 2 than Day 1 movies. However, in contrast to Experiment 1, results showed that neither the Object × Day nor the Object × Day × Bin interaction were significant. In sum, when watching an activity for the second time, participants looked predictively to the target object that the actor had previously manipulated. Participants did this starting early in the activity, before the two potential endings diverged. For changed activities, this resulted in mnemonic predictive-looking errors that were subsequently corrected as the actor's hand approached the changed target object.

Cued-recall performance

When participants were asked about features of Day 2 activities on the cued-recall test, most responses were correct reports of the object manipulation seen on that day—Day 2 recalls: 57% (Experiment 1), 50% (Experiment 2)—or incorrect reports of the alternate-object manipulation—alternate intrusions: 20% (Experiment 1), 21% (Experiment 2). In the example changed activity shown in Figure 1c, reporting "she unlocked the deadbolt" would be a correct Day 2 recall and reporting that "she unlocked the doorknob" would be an alternate intrusion. Note that alternate intrusions are intrusions from episodic memory only for changed activities; for repeated and control activities, such responses reflect false remembering of an action that was not performed but fit within the semantic context of the activity. The remaining responses included ambiguous responses (4% for Experiment 1, 9% for Experiment 2) that were correct but did not identify a target or alternate features (e.g., "she unlocked the house") and other errors such as recalling other activity features or omissions (19% for Experiment 1, 20% for Experiment 2).

Event changes can lead to proactive facilitation in memory. Both experiments showed that repeating an activity led to overall better correct recall of central features and fewer alternate intrusions than did repeating an initial action sequence and including a changed ending (Fig. 4). In Experiment 2, we also replicated the somewhat surprising finding that repeating an activity with a changed feature can lead to facilitation rather than interference in overall correct recall (Wahlheim & Zacks, 2019): Recall of central features for changed activities was better than recall of such features for control activities seen only on Day 2 (Fig. 4, bottom left). Separate models estimating correct Day 2 recall and alternate intrusion probabilities including only activity type as a fixed effect (repeated and changed activities in Experiment 1; repeated, control, and changed activities in Experiment 2) indicated significant effects for Day 2 recall—Experiment 1: $\chi^2(1) = 6.59$, $p = .010$; Experiment 2: $\chi^2\ (2) = 60.32$, $p < .001$—and for alternate intrusions—Experiment 1: $\chi^2(1) = 28.21$, $p < .001$; Experiment 2: $\chi^2(2) = 33.58$, $p < .001$. For Experiment 2, all pairwise comparisons among the three levels of activity type were significantly different for Day 2 recall (smallest z ratio = 3.43, $p = .002$), whereas the only nonsignificant difference for alternate intrusions was between repeated and control activities, z ratio = 1.86, $p = .150$.

Recollecting that activities had changed is associated with memory facilitation. To test whether successful updating was associated with recollecting that an activity had changed, we analyzed cued-recall performance for changed activities conditionalized on classifications indicating whether activities had changed, made just after each cued-recall attempt. We created three response-classification types for changed activities: *Change recollected* responses were accurate classifications with recall of Day 1 features, *change remembered* responses were accurate classifications without correct recall of Day 1 features, and *change-not-remembered* responses were inaccurate classifications. Change-classification rates were comparable in both experiments: recollected (Experiment 1 = 38%, Experiment 2 = 37%), remembered (Experiment 1 = 19%, Experiment 2 = 18%), and not remembered (Experiment 1 = 43%, Experiment 2 = 45%).

Correct recall of changed Day 2 features varied depending on how participants classified changed activities at test (Fig. 4). In both experiments, models fitted to correct Day 2 recall and including classification as the fixed effect indicated significant effects, smallest $\chi^2(2) = 188.29$, $p < .001$. Pairwise comparisons indicated that recall was significantly greater when change was recollected than when it was remembered but not recollected or not remembered, smallest z ratio = 11.13, $p < .001$, and that the latter two conditions were not significantly different, largest z ratio = 0.84, $p = .678$. These results replicate findings showing that enhanced event-memory updating for changed activities was associated with change recollection (Garlitch & Wahlheim, 2021; Hermann et al., 2021; Stawarczyk et al., 2020; Wahlheim & Zacks, 2019).

In both experiments, alternate intrusions were most likely when change was remembered but not recollected and least likely when change was recollected. The classification effect was significant in both experiments, smallest $\chi^2(2) = 90.96$, $p < .001$. All pairwise comparisons were significantly different, smallest z ratio = 3.54, $p = .001$.

Predictive-looking errors are associated with better event-memory updating. The previous analyses established that participants sometimes make mnemonic predictive-looking errors when watching Day 2 activities unfold and that watching a changed activity ending can lead to facilitation, rather than interference, in memory for changed activity features. We hypothesized that these two results are related such that making a predictive-looking error can induce memory updating (cf. Stahl & Feigenson, 2015). To test this hypothesis, we assessed proportions of looking while participants watched changed activities on Day 2 back-sorted on whether changes were later recollected (Fig. 5). We compared activities for which changes were recollected, that is, that were correctly classified and for which the Day 1 activity feature was recalled (*recollected*), with activities for which change was not recollected (*not recollected*). The latter category included activities for which Day 1 features were not recalled, regardless of whether activities were classified as having changed, because recall of Day 2 features in those cells was not significantly different (see Fig. 4, left).

For changed activities in the Day 2 movie, mnemonic predictive-looking errors were looks to alternate objects that the actor had manipulated in the Day 1 movie. Consistent with our hypothesis, Figure 5 (left) shows that in both experiments, proportions of looking to alternate objects for changed activities were greater during activities for which participants subsequently recollected that features had changed. This was especially true when participants watched the actor repeat actions from the Day 1 movie in the predivergence intervals.

In both experiments, models fitted to looks to alternate objects with change recollection (recollected vs. not recollected, and vs. control in Experiment 2 only) and bin as fixed effects showed significant main effects (Table 3). For Experiment 1, there was also a significant interaction that was characterized using separate models for pre- and postdivergence intervals (Table 3). The proportion of looking to alternate objects was significantly greater for changed activities that were subsequently recollected as such than for those not recollected

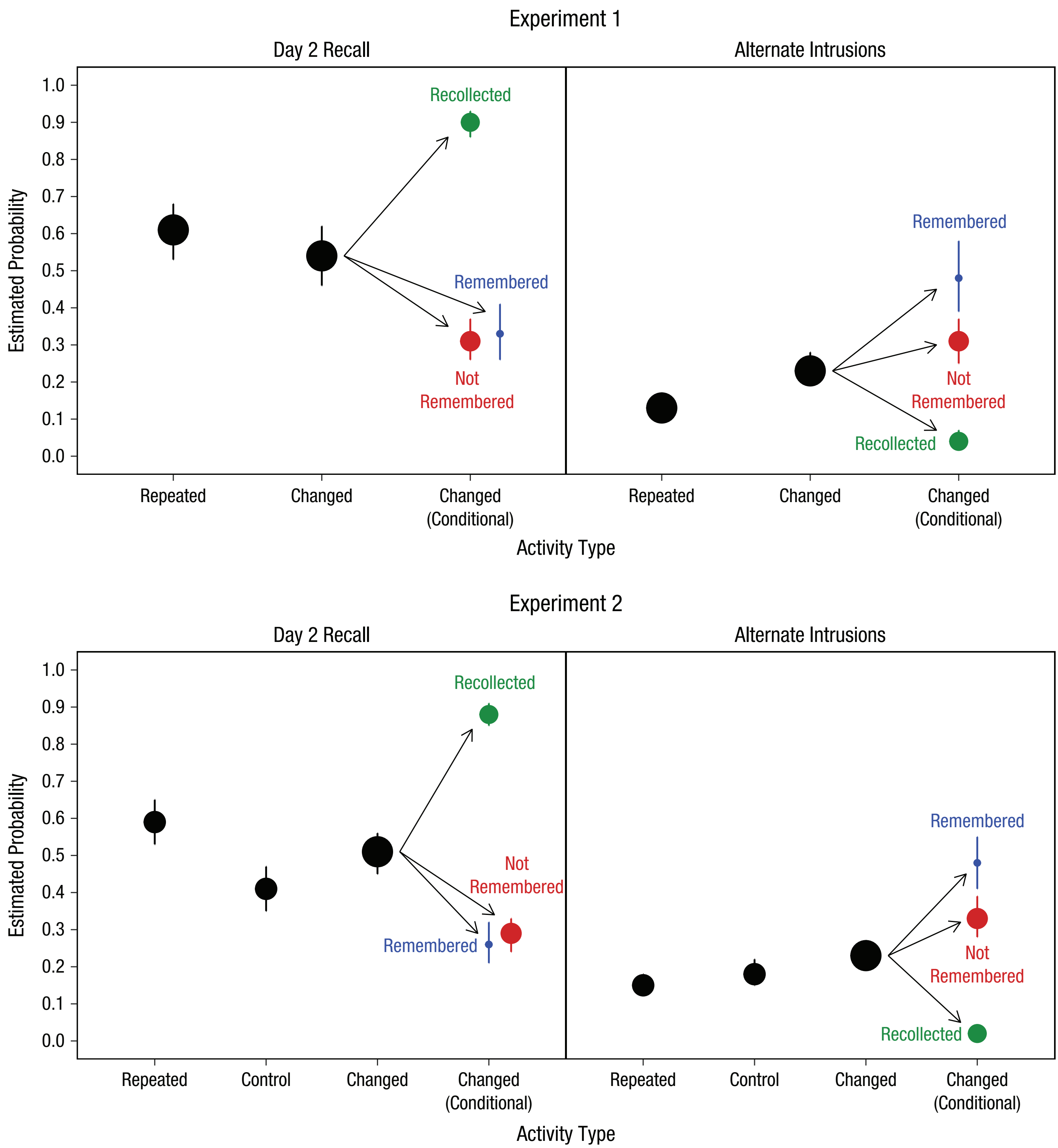

Fig. 4. Cued recall of activity features. The estimated probability of Day 2 recall (left column) and alternate intrusions (right column) is shown separately for each activity type in each experiment. Black points indicate model-estimated probabilities for overall responses. Colored points for conditional cells indicate model-estimated response probabilities for changed activities when changes were recollected (green points), remembered (blue points), or not remembered (red points). Point-area differences indicate the relative observation frequencies for those cells. Error bars represent 95% confidence intervals. Error bars are obscured when the intervals are smaller than the point areas.

as such. This difference was greater in the prediver-gence interval than in the postdivergence interval. For Experiment 2, the additional control condition prompted pairwise comparisons to characterize the effect of change recollection. Participants looked to alternate objects more during activities for which changes were

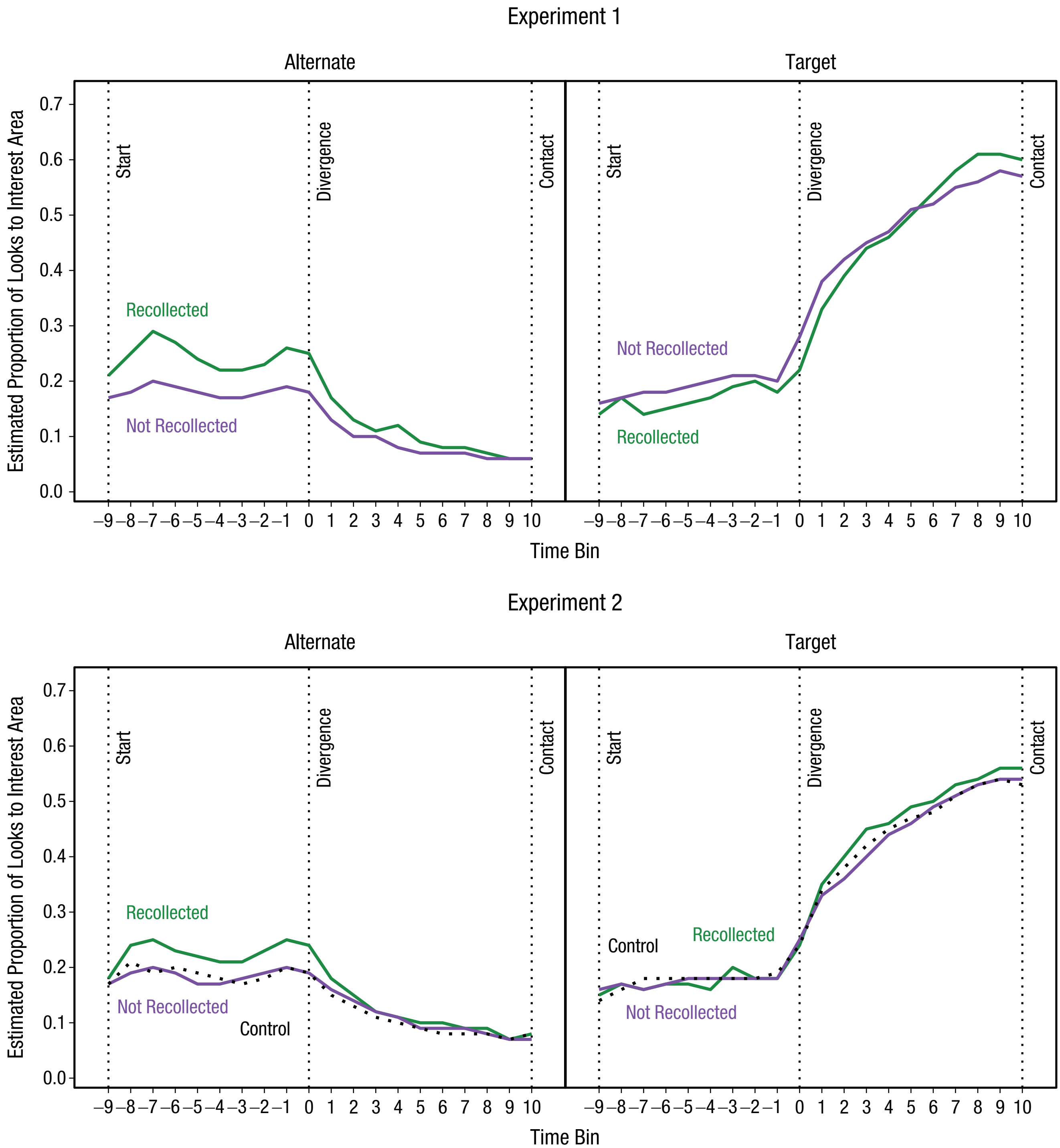

Fig. 5. Looks to interest areas during changed activities on Day 2. Model-estimated proportions of looking to interest areas for changed activities during Day 2 movies are shown for alternate objects (left column) and target objects (right column), separately for each time bin in Experiments 1 and 2. "Start" indicates the moment the interest period began. "Divergence" indicates the point of divergence when the two activity versions first became distinguishable. "Contact" indicates the moment the actor contacted the target or alternate object. Green solid lines indicate looking proportions for activities recollected as changed, purple solid lines indicate looking proportions for activities not recollected as changed, and black dotted lines (Experiment 2 only) indicate looking proportions for control activities.

subsequently recollected than during activities for which changes were not recollected, z ratio = 6.84, $p <$ .001, and during control activities, z ratio = 7.45, $p <$ .001. This pattern was highly consistent across activities varying in length of the predivergence interval (see the Supplemental Material). These results suggest that

Table 3. Model Results for Day 2 Looking to Alternate Objects Conditionalized on Change Recollection: Experiments 1 and 2

Time bin and effect	χ^2	*df*	*p*
Experiment 1			
All			
Change recollection	49.15	1	< .001
Bin	652.95	19	< .001
Change Recollection × Bin	34.91	19	.014
Predivergence intervals			
Change recollection	52.27	1	< .001
Bin	15.34	9	.082
Change Recollection × Bin	3.40	9	.946
Postdivergence intervals			
Change recollection	7.57	1	.006
Bin	73.12	9	< .001
Change Recollection × Bin	5.27	9	.810
Experiment 2			
All			
Change recollection	60.87	2	< .001
Bin	1,611.21	19	< .001
Change Recollection × Bin	41.22	38	.332

Note: The model results shown here correspond to the data displayed in Figure 5 (left). For Experiment 2, note that the change-recollection variable includes both changed activities conditionalized on change recollection and control activities that were not conditionalized on change recollection.

mnemonic predictive-looking errors played a critical role in the facilitation in memory updating that occurred when changes were recollected.

Proportions of looking to target objects in the Day 2 movie (Fig. 5, right) were also compared. These analyses tested the hypothesis that participants should not show greater predictive looking associated with subsequent change recollection because target objects had not been contacted in the Day 1 movie. Thus, mnemonic predictions should not guide the eyes to look earlier and more often to objects that the actor had not manipulated. Models comparable with those for alternate objects confirmed this prediction (Table 4). In Experiment 1, neither the effect of change recollection nor the Change Recollection × Bin interaction was significant. In Experiment 2, there was a significant effect of change recollection, but pairwise comparisons indicated no significant differences among conditions, smallest z ratio = 2.31, p = .054. Overall, these results are consistent with the hypothesis that memory predictively guides the eyes only to previous action targets. Finally, we also found that overall correct recall of changed activity features from the Day 2 movie was associated with greater proportions of looking to alternate objects in the predivergence interval (see the Supplemental Material).

Discussion

Two experiments examined the role of mnemonic prediction error in event-memory updating by assessing the association between predictive looking and memory for changed activities. Predictive looking to objects an actor had contacted was associated with facilitated memory for changed activity features. This was shown by an association between memory for earlier-contacted objects and looking to those objects during repeated actions as well as better memory for changed objects and recollection of the change. These results suggest that mnemonic prediction errors partly contribute to event-memory updating for dynamic everyday actions.

The present results converge with those of action-observation studies showing that viewers look ahead when performing and watching everyday actions (Eisenberg et al., 2018; Hayhoe et al., 2003; Land & McLeod, 2000). These results replicated findings of predictive looking during movies of everyday activities (Eisenberg et al., 2018), as viewers looked increasingly to target objects when watching Day 1 movies. These findings support the perspective that viewers use motor knowledge to predict action goals (Flanagan & Johansson, 2003). The present results also converged with findings from simplified trial-based designs showing that viewers look ahead to earlier-watched actions (Cannon et al., 2012; Falck-Ytter et al., 2006; Gredebäck et al., 2018). Specifically, when watching the Day 2 movie, participants looked more to objects that had been contacted in the Day 1 movie. This was true for activities that repeated completely and those that included only repetitions of the initial action sequence in the predivergence interval, suggesting that memory guided action predictions during activity viewing.

Table 4. Model Results for Day 2 Looking to Targets Conditionalized on Change Recollection: Experiments 1 and 2

Experiment and effect	χ^2	*df*	*p*
Experiment 1			
Change recollection	2.04	1	.153
Bin	2,940.44	19	< .001
Change Recollection × Bin	19.17	19	.446
Experiment 2			
Change recollection	6.40	2	.041
Bin	9,022.16	19	< .001
Change Recollection × Bin	24.37	38	.958

Note: The model results shown here correspond to the data displayed in Figure 5 (right). For Experiment 2, note that the change-recollection variable includes both changed activities conditionalized on change recollection and control activities that were not conditionalized on change recollection.

The present study also showed that mnemonic prediction errors enhanced adults' new learning. These results extend those of action-observation studies with infants and children showing that prediction errors invite active learning (Gredebäck et al., 2018; Juvrud et al., 2019; Stahl & Feigenson, 2015). Participants in the present study looked earlier and more often to repeated predivergence actions and changed postdivergence actions, suggesting that mnemonic prediction errors stimulated encoding of changed features. This view was also supported by the finding that change recollection, which was strongly associated with facilitated memory updating, was also associated with predictive looking to earlier-contacted objects. The present results therefore show that mnemonic prediction errors facilitate encoding and retrieval of everyday events.

The role of prediction error in event-memory updating has been studied in various ways, leading to different outcomes and theoretical perspectives. Some theories assume that prediction errors impair memory when unfulfilled expectations lead to weakened representations of expected events (Kim et al., 2014) or increases in intrusion errors (Sinclair & Barense, 2018). These views are similar to models proposing that reactivating existing memories prior to new events destabilizes those memories (Exton-McGuinness et al., 2015; Lee et al., 2017). Other theories posit that repeating event features triggers retrieval of existing memories, leading to predictions that associated stimuli will follow. However, when changes occur, prediction errors stimulate encoding (Antony et al., 2021; Bein et al., 2020; Chen et al., 2015; Greve et al., 2017; Wahlheim & Zacks, 2019). The present findings are consistent with the latter perspective. Evidence from looking proportions associated with error-driven updating converges with results showing that neural reactivation of dynamic naturalistic events is associated with improved encoding of changed activity features (Stawarczyk et al., 2020).

One potential mechanism for the error-driven updating observed here is integrative encoding of event representations. Interference theories predict that event changes should impair memory (Anderson & Neely, 1996). However, there is mounting evidence that interference is mitigated when prior events are retrieved and integrated with new events (Chanales et al., 2019; Wahlheim & Jacoby, 2013). This integration may be mediated partly by medial temporal structures that are biased toward encode states by mnemonic prediction errors (Bein et al., 2020). Another mechanism that could operate simultaneously is integrative encoding of eye-movement shifts from first- to second-contacted objects after prediction errors. This is consistent with the view that eye movements can become embedded in and support subsequent memory (Ryan et al., 2020; Wynn et al., 2019). Such gaze-enhanced encoding could be examined by restricting eye movements (e.g., Henderson et al., 2005) to determine whether preventing shifts between objects impairs memory for recently contacted objects. Finally, although we have interpreted the present looking patterns as showing that updating benefited from retrieval of activities before changes appeared, it may have also benefited from retrospective comparisons of activity features occurring after changes appeared. Research is needed to understand how the timing of retrieval-and-comparison processes affects subsequent memory updating.

Although the current paradigm is more naturalistic than many other memory-updating paradigms, some features of the task limit its generalizability. Participants watched the actor perform everyday actions in a distraction-limited environment, which differs from everyday life in which viewing goals are dynamic and susceptible to distractions. Also, the present movies included an unfamiliar actor with specific demographic characteristics, which differs from everyday life because viewers also observe actions of actors of varying familiarity from various backgrounds. Future studies should manipulate these variables to better assess the conditions under which mnemonic prediction errors are associated with facilitation in memory updating. Finally, the present samples included only young adults from a selective university. It is thus unclear how the results would generalize to people of different ages whose predictive-processing abilities vary and to people from other cultures for whom the everyday activities are less familiar.

In summary, while watching an actor perform naturalistic everyday actions, participants looked ahead to earlier-contacted objects before action changes were perceived. This predictive looking occurred more for actions subsequently recollected to have changed, and this was associated with better memory updating for changed actions. These results support the view that mnemonic prediction errors play a critical role in event-memory updating by promoting more effective encoding and retrieval of changed activity features.

Transparency

Action Editor: Karen Rodrigue
Editor: Patricia J. Bauer
Author Contributions

C. N. Wahlheim and M. L. Eisenberg conceptualized the study and contributed equally to it. J. M. Zacks provided substantial guidance about the execution of both experiments. C. N. Wahlheim and J. M. Zacks developed the stimuli used in both experiments. M. L. Eisenberg and C. N. Wahlheim prepared the stimuli, designed the program, and coordinated data collection for Experiment 1. D. Stawarczyk prepared the stimuli, designed the program, and coordinated data collection for Experiment 2. J. M. Zacks, M. L. Eisenberg, and C. N. Wahlheim cleaned and analyzed the data for both experiments. C. N. Wahlheim prepared the initial draft of the manuscript. All authors provided critical comments and revisions and approved the final manuscript for submission.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This project was funded by National Institute on Aging Grants (Nos. T32AG000030 and R21AG05231401), a National Science Foundation Graduate Research Fellowship Grant (No. DGE-1143954), the Mr. and Mrs. Spencer T. Olin Fellowship, and the European Union's Horizon 2020 Research and Innovation Programme (under Marie Skłodowska-Curie Grant Agreement No. 798109).

Open Practices

Deidentified data, analysis code, and experimental stimuli have been made publicly available via OSF and can be accessed at https://osf.io/w8skp/. The design and analysis plans for the experiments were not preregistered. This article has received the badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

ORCID iDs

Christopher N. Wahlheim https://orcid.org/0000-0002-2381-1493

Michelle L. Eisenberg https://orcid.org/0000-0001-7276-2176

Acknowledgments

For their assistance with data collection and coding, we thank Shelly Gupta, Katie Munson, Connor Perkins, and Keshav Jayaraman. For her assistance with experiment development and data collection, we thank Hannah Aaron. Portions of the findings reported here were presented at the 57th, 60th, and 61st annual meetings of the Psychonomic Society and the 2017 North Carolina Cognition Conference.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/09567976211053596

References

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

Antony, J. W., Hartshorne, T. H., Pomeroy, K., Gureckis, T. M., Hasson, U., McDougle, S. D., & Norman, K. A. (2021). Behavioral, physiological, and neural signatures of surprise during naturalistic sports viewing. *Neuron*, *109*(2), 377–390. https://doi.org/10.1016/j.neuron.2020.10.029

Barnes, G. R., Collins, C. J. S., & Arnold, L. R. (2005). Predicting the duration of ocular pursuit in humans. *Experimental Brain Research*, *160*(1), 10–21. https://doi.org/10.1007/s00221-004-1981-3

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

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

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental Science*, *15*(1), 35–42. https://doi.org/10.1111/j.1467-7687.2011.01095.x

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

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

Diaz, G., Cooper, J., Rothkopf, C., & Hayhoe, M. (2013). Saccades to future ball location reveal memory-based prediction in a virtual-reality interception task. *Journal of Vision*, *13*(1), Article 20. https://doi.org/10.1167/13.1.20

Eisenberg, M. L., Zacks, J. M., & Flores, S. (2018). Dynamic prediction during perception of everyday events. *Cognitive Research: Principles and Implications*, *3*(1), Article 53. https://doi.org/10.1186/s41235-018-0146-z

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

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, *9*(7), 878–879. https://doi.org/10.1038/nn1729

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*(6950), 769–771. https://doi.org/10.1038/nature01861

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). SAGE. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

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

Gerson, S. A., & Woodward, A. L. (2014). Learning from their own actions: The unique effect of producing actions on infants' action understanding. *Child Development*, *85*(1), 264–277. https://doi.org/10.1111/cdev.12115

Gredebäck, G. (2018). How visual and motor experience shapes the development of infants' perception of actions performed by social partners. *Journal of Motor Learning and Development*, *6*(Suppl. 1), S89–S104. https://doi.org/10.1123/jmld.2016-0074

Gredebäck, G., & Falck-Ytter, T. (2015). Eye movements during action observation. *Perspectives on Psychological Science*, *10*(5), 591–598. https://doi.org/10.1177/1745691615589103

Gredebäck, G., Lindskog, M., Juvrud, J. C., Green, D., & Marciszko, C. (2018). Action prediction allows hypothesis testing via internal forward models at 6 months of age. *Frontiers in Psychology*, *9*, Article 290. https://doi.org/10.3389/fpsyg.2018.00290

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

Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, *3*(1), Article 6. https://doi.org/10.1167/3.1.6

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition*, *33*(1), 98–106.

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*, 1387–1404. https://doi.org/10.3758/s13421-021-01173-2

Juvrud, J., Bakker, M., Kaduk, K., DeValk, J. M., Gredebäck, G., & Kenward, B. (2019). Longitudinal continuity in understanding and production of giving-related behavior from infancy to childhood. *Child Development*, *90*(2), e182–e191. https://doi.org/10.1111/cdev.13131

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, USA*, *111*(24), 8997–9002. https://doi.org/10.1073/pnas.1319438111

Land, M. F., & McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball. *Nature Neuroscience*, *3*(12), 1340–1345. https://doi.org/10.1038/81887

Lee, J. L. C., Nader, K., & Schiller, D. (2017). An update on memory reconsolidation updating. *Trends in Cognitive Sciences*, *21*(7), 531–545. https://doi.org/10.1016/j.tics.2017.04.006

Lenth, R. (2020). *emmeans: Estimated marginal means, aka least-squares means* (R Package Version 1.4.7) [Computer software]. https://CRAN.R-project.org/package=emmeans

Melzer, A., Prinz, W., & Daum, M. M. (2012). Production and perception of contralateral reaching: A close link by 12 months of age. *Infant Behavior and Development*, *35*(3), 570–579. https://doi.org/10.1016/j.infbeh.2012.05.003

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1003–1019. https://doi.org/10.1037/0278-7393.15.6.1003

R Core Team. (2021). *R: A language and environment for statistical computing* (Version 4.1.1) [Computer software]. https://www.R-project.org/

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement learning and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–88). Appleton-Century-Crofts.

Ryan, J. D., Shen, K., & Liu, Z. (2020). The intersection between the oculomotor and hippocampal memory systems: Empirical developments and clinical implications. *Annals of the New York Academy of Sciences*, *1464*(1), 115–141. https://doi.org/10.1111/nyas.14256

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime 2 reference guide*. Psychology Software Tools.

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

SR Research. (2020). *Experiment Builder* (Version 2.3.38) [Computer software]. https://www.sr-research.com/experiment-builder/

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, *348*(6230), 91–94. https://doi.org/10.1126/science.aaa3799

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, USA*, *117*(47), 29346–29353.

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

Wahlheim, C. N., & 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

Wynn, J. S., Shen, K., & Ryan, J. D. (2019). Eye movements actively reinstate spatiotemporal mnemonic content. *Vision*, *3*(2), Article 21. https://doi.org/10.3390/vision3020021