Journal of Experimental Psychology: General

Cognitive Mechanisms of Aversive Prediction Error-Induced Memory Enhancements

Kaja Loock, Felix Kalbe, and Lars Schwabe

Online First Publication, January 23, 2025. https://dx.doi.org/10.1037/xge0001712

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

https://doi.org/10.1037/xge0001712

Cognitive Mechanisms of Aversive Prediction Error-Induced Memory Enhancements

Kaja Loock, Felix Kalbe, and Lars Schwabe
Department of Cognitive Psychology, Institute of Psychology, Universität Hamburg

While prediction errors (PEs) have long been recognized as critical in associative learning, emerging evidence indicates their significant role in episodic memory formation. This series of four experiments sought to elucidate the cognitive mechanisms underlying the enhancing effects of PEs related to aversive events on memory for surrounding neutral events. Specifically, we aimed to determine whether these PE effects are specific to predictive stimuli preceding the PE or if PEs create a transient window of enhanced, unselective memory formation. In a combined incidental encoding-fear learning task, participants ($n = 355$) estimated aversive shock probabilities after trial-unique stimuli. Physiological arousal and explicit PEs were measured during encoding to predict recognition memory tested either immediately after encoding (Experiment 3) or 24 hr later (Experiments 1–4). Our results show that the retroactive memory enhancement induced by PEs may extend back longer than previously assumed, impacting stimuli presented 10 s before the PE. Furthermore, PE-driven memory enhancement extends beyond predictive stimuli preceding the PE event to those encountered afterward. Importantly, our findings reveal that PE-related memory enhancement for stimuli preceding the PE event is specific to predictive stimuli, with uninformative stimuli not benefiting from PEs and even interfering with the PE-driven memory enhancement. This pattern demonstrates that PE effects are not unspecific but that PEs enhance memory for predictive stimuli encountered around a PE event. Notably, memory-enhancing effects of PEs persisted even when controlling for changes in arousal. These findings provide insights into the cognitive mechanisms of PE-induced enhancements of memory, with potential implications for understanding aberrant emotional memory in fear-related disorders.

Public Significance Statement

In our daily lives, we encounter a myriad of stimuli, yet our brain selects only a fraction of these for long-term storage in memory. While emotional arousal is well established to facilitate memory storage, recent research revealed an additional mechanism contributing to memory formation for inherently neutral stimuli that includes PEs, that is, discrepancies between anticipated and actual outcomes. This study demonstrates that PEs related to aversive events enhance memory specifically for predictive stimuli surrounding the PE, but not for proximal uninformative stimuli, suggesting that primarily stimuli that contain predictive value for emotional events are stored in memory, with significant implications for understanding aberrant emotional memory in fear-related mental disorders.

Keywords: episodic memory, emotional memory, aversive learning, prediction error, arousal

Supplemental materials: https://doi.org/10.1037/xge0001712.supp

In our daily lives, we are constantly exposed to a continuous stream of information. However, only a fraction of this information is retained in our memory for extended periods. This raises a fundamental question in memory research: what determines the selection of information stored in memory and why? Predicting which information is subsequently remembered has tremendous implications for several applied contexts, including eyewitness testimony, educational contexts, working contexts, or clinical

Ashleigh Maxcey served as action editor.

Lars Schwabe https://orcid.org/0000-0003-4429-4373

Data, code, and materials are publicly available at https://www.fdr.uni-hamburg.de/record/14147. Part of the data of Experiments 1 and 2 were analyzed to address a research question unrelated to the present article and has been reported in Kalbe and Schwabe (2022a). Part of the data of Experiment 3 have been presented at the "Psychology and Brain" conference in Hamburg in 2024 under the title: The Specificity of Prediction Error-Related Memory Enhancement.

The authors gratefully acknowledge the assistance of Leah Just, Tim Dreβler, Konstantin Döhr, Anna Behnck, Fatih Yücel, Christoph Frühlinger, Hazem Abuhanya, Constantin Zier, Philipp Strauβ, Lilli Buch, Klara Kaczmarek, Olga Ehmling, and Sabrina Franke during data collection.

Kaja Loock played a lead role in data curation, project administration,

continued

contexts. Decades of research demonstrated the impact of emotional relevance on memory formation: compared to neutral events, emotionally arousing events are more vividly and more accurately remembered (Cahill & McGaugh, 1996; Christianson, 2014; Reisberg & Hertel, 2005). This emotional memory enhancement is attributed to the noradrenergic arousal-related activation of the amygdala, which then modulates memory storage processes in other brain areas such as the hippocampus (Buchanan, 2007; Cahill & McGaugh, 1998; LaBar & Cabeza, 2006; McGaugh, 2018; Phelps, 2004). In line with this arousal view, recent research showed that short bursts of arousal, reflected in pupil dilation, during encoding can predict subsequent memory of individual stimuli (Bergt et al., 2018; Clewett et al., 2018; Clewett & McClay, 2024; Mather et al., 2016).

While the influence of emotional arousal on memory formation is well established, recent evidence from appetitive and aversive learning domains points to an additional mechanism contributing to memory formation: prediction errors (PEs; Antony et al., 2021; Kalbe & Schwabe, 2020; Rouhani et al., 2018). Emotional events are often unpredictable and deviate from expectation, resulting in PEs. PEs drive memory updating and act as a teaching signal to the brain enabling incremental learning to optimize future behavior (Bar, 2007; Bein et al., 2020; Clark, 2013; Ergo et al., 2020; Rouhani et al., 2020; Trapp et al., 2018). Recent research further showed that PEs contribute to memory formation (for a review, see Rouhani et al., 2023). In particular, it has been repeatedly shown that PEs associated with aversive or appetitive events can boost the memory for preceding neutral events (Greve et al., 2017; Jang et al., 2019; Kalbe & Schwabe, 2020, 2022b; Pine et al., 2018; Rouhani et al., 2018; Rouhani & Niv, 2019, 2021); although the effect can differ between unsigned PEs (uPEs; i.e., absolute magnitude of deviation between prediction and outcome, ranging between 0 and 1) and signed PEs (sPEs; that is, direction of deviation containing information about the value of the outcome, positive vs. negative, ranging from −1 to 1; Rouhani et al., 2023), suggesting that uPEs range between 0 and 1, whereas sPEs can take any value between −1 and 1. While there is some overlap between the arousal-related route and the PE-related route to enhanced memory formation (Rouhani et al., 2023), initial evidence suggests that PE-related effects on episodic memory formation go beyond the effects of physiological arousal (Kalbe & Schwabe, 2020) and rely on a distinct brain mechanism (Kalbe & Schwabe, 2022b). Importantly, the arousal- and PE-related mechanisms of memory enhancement imply different approaches to modulate memory, for instance, in clinical contexts. Whereas the arousal mechanism suggests (e.g., pharmacological) manipulations of arousal to modify memory, the PE mechanism suggests targeting individual expectations. Despite the growing evidence for PE effects on memory and their potential implications, the cognitive mechanisms that are involved in PE-induced memory enhancements are not well understood.

At first glance, PE effects on memory might be considered comparable to the well-known memory boost for novel or salient events that are due to increased attention (Schlüter et al., 2019; Sinclair & Barense, 2019). However, PE effects differ fundamentally from these "oddball" effects, referring not to the enhanced memory for the PE event itself but to a memory enhancement for inherently neutral stimuli preceding the PE (Kalbe & Schwabe, 2020, 2022b; Rouhani et al., 2018; Rouhani & Niv, 2021). These retrospective effects of a PE could hardly be explained by mere attentional mechanisms. A major question related to PE effects on memory concerns, however, whether the PE-related memory boost is selective to the predictive stimulus preceding the PE event. If so, then there should be no memory enhancement for either stimuli following the PE or for uninformative stimuli preceding the PE. Alternatively, it could be hypothesized that PEs open a transient window of enhanced mnemonic processing that may also enable better memory for nonpredictive stimuli. Whereas the selective memory enhancement would involve mnemonic efficiency, the latter would reflect a "better-safe-than-sorry" mechanism making sure that all stimuli that occurred in the surrounding of an unexpected emotional event are preferentially stored in memory.

Irrespective of whether the PE-related memory enhancement is specific to the predictive stimulus or whether there is a PE-related window of enhanced memory formation, another open question is for how long the demonstrated (retrospective) effect of PEs lasts. Extant studies implemented a constant and relative short interval between predictive stimulus and outcome (i.e., PE) and it remains unknown whether the PE effect on memory is reduced, or even abolished, once the stimulus-outcome interval is extended. Similarly, it is unknown whether a predictive stimulus needs to be actively maintained, presumably in working memory, until the occurrence of PE to get preferentially stored in memory. If there is a selective PE-related enhancement for predictive stimuli and predictive stimuli need to be actively maintained, then it can be predicted that this enhancement should be sensitive to the stimulus-outcome interval and to interference by other (uninformative) stimuli.

In the present series of experiments, we aimed to elucidate the cognitive mechanisms underlying aversive PE-driven memory enhancements for inherently neutral events. In all of these experiments, participants completed a combined incidental encoding-fear conditioning task in which they saw a stream of initially neutral stimuli from different categories that were associated with a differential probability of electric shocks. We asked participants to predict the occurrence of an electric shock which then allowed us to calculate PEs based on their expectation and its deviation from the actual outcome (shock vs. no-shock). Memory for the presented pictures was tested in a surprise recognition test either immediately after encoding (Experiment 3) or 24 hr later (Experiments 1–4). In Experiment 1, we aimed to replicate the previously reported memory enhancing

formal analysis, validation, writing–original draft, and writing–review and editing and a supporting role in conceptualization. Felix Kalbe played a lead role in data curation and project administration and a supporting role in formal analysis, conceptualization, and writing–review and editing. Lars Schwabe played a lead role in conceptualization, funding acquisition, supervision, writing–original draft, and writing–review and editing.

Kaja Loock played a lead role in data curation, formal analysis, methodology, and visualization, a supporting role in conceptualization, and an equal role in project administration, writing–original draft, and writing–review and editing. Felix Kalbe played a supporting role in conceptualization, data curation, formal analysis, and methodology and an equal role in writing–review and editing. Lars Schwabe played a lead role in conceptualization, funding acquisition, resources, and supervision and an equal role in project administration, writing–original draft, and writing–review and editing.

Correspondence concerning this article should be addressed to Lars Schwabe, Department of Cognitive Psychology, Institute of Psychology, Universität Hamburg, Hamburg 20146, Germany. Email: Lars.Schwabe@uni-hamburg.de

effect of aversive PEs on the subsequent memory for predictive stimuli presented before the PE (retrospective PE effects; Kalbe & Schwabe, 2020, 2022b). Furthermore, we tested in this first experiment whether the memory-enhancing PE effect extends to stimuli presented after the PE (prospective PE effects), in line with the assumption of a PE-induced window of enhanced memory formation. Experiment 2 served to replicate the findings of Experiment 1. In Experiment 3, we varied the stimulus-outcome interval to probe the time dependency of the PE effects on memory. Moreover, we varied in this experiment the encoding-test interval (immediate vs. 24 hr delayed) to assess whether aversive PE effects on memory develop already during encoding or whether they require a period of consolidation. In this experiment, we further modified the assessment of PEs to allow a distinction between signed and uPE effects, which have previously been shown to exert differential effects on memory (Kalbe & Schwabe, 2022b; Rouhani & Niv, 2021). The final Experiment 4 tested whether PEs enhance memory also for uninformative stimuli presented between the predictive stimulus and the PE, thus probing the specificity of the PE-related memory enhancement, and whether uninformative stimuli can even interfere with the PE-driven memory boost. In all of the four experiments, we measured autonomic arousal to test whether the observed PE effects go beyond the known effects of arousal on memory formation.

Experiment 1: Examining the Prospective and Retrospective Effects of PEs Related to Aversive Events on Episodic Memory Formation

The objectives of Experiment 1 were two-fold: first, this experiment aimed to replicate the enhancing effects of PEs related to aversive events on the memory for (predictive) stimuli that preceded the PE. Second, we aimed to test whether the memory-enhancing effects of PEs are limited to stimuli that immediately preceded the PE or whether PEs may enhance memory also for stimuli encountered shortly after a PE. In other words, we asked whether PEs may, in addition to the previously reported retrospective effects (Kalbe & Schwabe, 2020), exert prospective effects on memory formation. To this end, participants completed a combined incidental encoding-fear learning task in which they predicted whether a stimulus would be followed by an electric shock. During this task, we measured skin conductance responses (SCRs) as an indicator of physiological arousal, enabling us to probe potential arousal effects on memory formation. Twenty-four hours after encoding, memory was tested in a recognition test.

Method

Participants

Eighty-four healthy volunteers with normal or corrected to normal vision participated in this experiment (age: $M = 25.11$ years, $SD = 3.57$ years, range = 18–33 years). Participants were fluent German speakers, had no current illnesses, no life-time history of any mental or neurological disorders and did not take any prescriptive medication as assessed in a standardized telephone interview. Furthermore, women being pregnant were excluded from participation. Six participants were excluded from the analyses because they did not return for the second experimental day or due to technical failure during the experiment, resulting in a final sample of $n = 78$. This sample was part of a larger study on emotional learning processes (Kalbe & Schwabe, 2022a).

The sample size was based on an a priori power calculation using G*Power (3.1.9.6; Faul et al., 2009). Based on previous research by Kalbe and Schwabe (Kalbe & Schwabe, 2022a), we assumed $d_z = .45$ as a point estimate for the expected PE effects. The power calculation showed that a sample of at least 67 participants is required to detect an effect of the expected size in a two-tailed paired *t*-test with a statistical power of 0.95. In line with these assumptions, a post hoc power simulation using the R-package simR (Green & MacLeod, 2016) for the observed effects of subsequent PEs and our final sample size of 78 participants yielded a power of 0.99 based on 1,000 simulations. All participants provided written informed consent before participation and received a monetary reimbursement of 20€ at the end of the study. The study was approved by the ethics committee of the Faculty of Psychology and Human Movement Science at the Universität Hamburg and carried out in line with the Declaration of Helsinki.

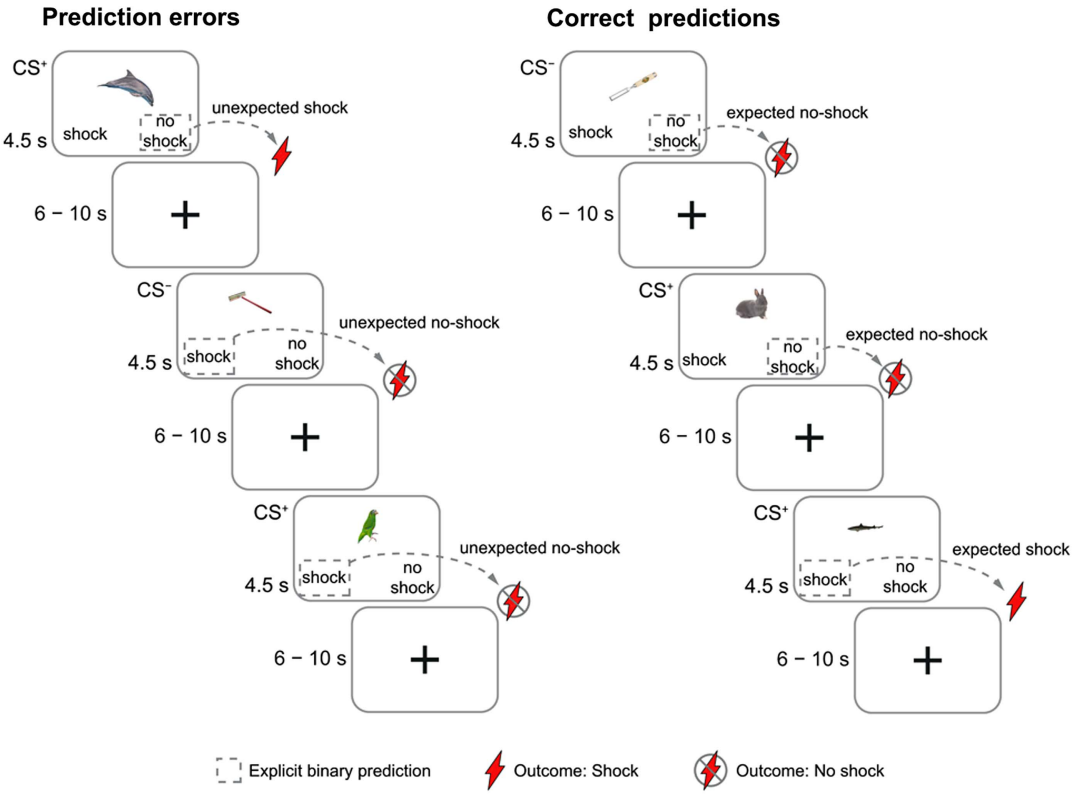
Materials

For this experiment, we used the same stimulus set as in Kalbe and Schwabe (2022a). It consisted of 180 color photographs of animals and 180 color photographs of tools isolated on white backgrounds. These photographs were taken from already existing databases (Bank of Standardized Stimuli; Brodeur et al., 2010, 2014), McGill Calibrated Color Image Database (Olmos & Kingdom, 2004), SUN database (Xiao et al., 2010), Konklab (Konkle et al., 2010), which were developed for nonemotion research on cognition, vision, and psycholinguistics. All stimuli were assumed to be of neutral valence, and each object or animal represented a unique exemplar of its category. Importantly, each photograph was only presented once and there were not two different photographs of, for example, cats or screwdrivers. From this pool, 30 photographs of animals and 30 photographs of tools were randomly drawn and used on the first experimental day and 120 photographs for encoding tasks unrelated to the purpose of the present study. These unrelated tasks took place before and after the relevant learning paradigm and importantly, did not contain any predictions or aversive events, thus making it highly unlikely that these tasks interfered with the memory paradigm of interest here. The remaining 180 photographs served as lures for the surprise recognition test on the second experimental day. The order in which individual items were presented was randomized across participants.

Procedure

The experiment consisted of 2 days, with an encoding session on the first experimental day and a recognition test on the second experimental day, 24 hr later (see Figure 1). Upon arrival on the first experimental day, participants provided written informed consent and received written instructions indicating that they were going to see a series of photographs of animals and tools and that some of them might be followed by a brief electric shock. Participants were instructed to try to predict whether a shock would follow the current photograph ("shock" vs. "no shock" response). Importantly, participants were neither informed about the shock contingencies, nor that a subsequent memory test would follow on the second day.

Figure 1
Experimental Procedure of Experiments 1 and 2



Note. Participants completed a combined Pavlovian fear-conditioning and incidental memory paradigm. They saw initially neutral pictures from two different categories, one of which was associated with receiving an electric shock with a shock-contingency of 67% (CS^+). In each trial, they were asked to make their binary prediction about the occurrence of an electric shock. Critically, in Experiment 1, the 200 ms shock occurred after stimulus offset while the shock coterminated with the predictive stimulus in Experiment 2 (4.3 s after stimulus onset). On a second experimental day, memory was tested for these in items in a surprise recognition test. Pictures taken and reprinted from "The Bank of Standardized Stimuli (BOSS), a New Set of 480 Normative Photos of Objects to Be Used as Visual Stimuli in Cognitive Research," by M. B. Brodeur, E. Dionne-Dostie, T. Montreuil, and M. Lepage, 2010, *PLOS ONE*, *5*(5), Article e10773 (https://doi.org/10.1371/journal.pone.0010773). CC BY 4.0 and from "Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos," by M. B. Brodeur, K. Guérard,and M. Bouras, 2014, *PLOS ONE*, *9*(9), Article e106953 (https://doi.org/10.1371/journal.pone.0106953). CC BY 4.0. CS = conditioned stimulus. See the online article for the color version of this figure.

To record SCRs as indicator of physiological arousal and conditioned fear, electrodes were placed on the distal phalanx of the second and third finger of the left hand. Skin conductance was measured using the MP-160 data acquisition and analysis BIOPAC system (BIOPAC systems, Goleta, California, United States). For electrical stimulation, we used the STM-200 stimulator module connected to the MP-160. A stimulation electrode was placed on the back of the right hand near the wrist. Stimulation intensity was adjusted individually to be unpleasant but not painful using a standardized procedure. More specifically, a total of twelve 200-ms single pulse shocks were administered, with an initial intensity of 10 V. After each trial, participants rated whether the received shock had been painful in a forced choice fashion using the "left" ("not painful") and "right" ("painful") keys. Whenever a shock was rated as not painful, its intensity for the next trial was increased slightly. Analogous, when participants rated the shock as painful, it was decreased slightly. The aim was to choose an intensity that was unpleasant but not painful to the participants.

During the encoding session, 30 photographs of animals and 30 photographs of tools were presented in a pseudorandomized order, so that no more than three pictures of the same category appeared in a row. In each trial, a photograph was shown in the center of a computer screen for 4.5 s, during which participants were asked to make their binary prediction about the occurrence of an electric shock using the "1" and "2" buttons on the keyboard, corresponding to *no shock* and *shock*, respectively (see Figure 1). Critically, shock contingencies were linked to the item category, such that one image category served as excitatory conditioned stimulus (CS^+) and the other one served as inhibitory conditioned stimulus (CS^-). The assignment of tools or animals as CS^+ or CS^- was counterbalanced across participants. In CS^+ trials the shock contingency was two-thirds, resulting in 20 out of 30 trials that included a shock. In

CS^- trials, no shocks were administered. Each trial was followed by a black fixation cross centered on white background for 8 ± 2 s, which enabled measuring the relatively slow SCRs elicited by the photographs and the shocks. After the experimental task which lasted approximately 12 min, electrodes were removed and participants were asked to rate the intensity of the shocks on a scale from 1 (*not unpleasant at all*) to 10 (*extremely unpleasant*).

On Experimental Day 2, 22–26 hr after the encoding session, participants returned for a surprise recognition test. First, they completed a short questionnaire to assess whether they anticipated a memory test and then rated how surprised they were about the recognition test on a scale from 1 (*not surprised at all*) to 5 (*very surprised*). Next, they received written instructions explaining details of the recognition test. During the recognition test, participants were presented all pictures they had seen on Experimental Day 1 (90 pictures of animals and 90 pictures of tools) as well as 180 "new" pictures (90 pictures of animals and 90 pictures of tools) that had not been presented on the previous day. Each trial started with a central black fixation cross on a white background for 1.5 ± 0.5 s, followed by an "old" or "new" picture presented centrally on the computer screen. For each item, participants had to indicate whether the currently presented picture was *definitely old*, *maybe old*, *maybe new*, or *definitely new* by pressing the "1," "2," "3," or "4" button on the keyboard, respectively. There were no time restrictions for participants' responses.

Data Analysis

For each trial, we derived binary PEs which were calculated as the absolute value of the difference between participants' explicit binary shock expectancy ratings (coded 0 when no shock was expected and coded 1 when a shock was expected) and the actual outcome of the trial (coded 0 when no shock occurred and 1 when a shock occurred in the current trial). Therefore, the resulting unsigned PE is also binary, attaining 0 for any correct prediction (i.e., either an expected shock or an expected shock omission) and 1 for any incorrect prediction (i.e., either an unexpected shock or an unexpected shock omission).

SCRs were analyzed using Continuous Decomposition Analysis in Ledalab Version 3.4.9 (Benedek & Kaernbach, 2010). For a detailed description of the SCR analysis, see Supplemental Material.

To investigate how PEs impacted the ability to recognize pictures presented during incidental encoding on the next day, we fitted generalized linear mixed models (GLMMs) with a logit link function using the lme4 R package (Bates et al., 2015). Compared with a "classic" analysis of proportions of binary recognition per condition and per participant, GLMMs have several advantages, such as increased statistical power and being less prone to spurious results (Dixon, 2008; Jaeger, 2008). Following guidelines to maximize the generalizability of these models, we included the maximal random effects structure, treating subjects as random effects for both the intercept and all slopes of the fixed effects included in the model (Barr et al., 2013). The recognition of an individual item was treated as the binary dependent variable, coded "0" for misses and "1" for hits. In line with previous research on episodic memory (Bartlett et al., 1980; Kalbe & Schwabe, 2022b), our analysis focused on high-confidence responses, that is, only trials in which participants indicated that they were very sure were considered. Such high-confidence recognitions have been linked to a hippocampus-based recollection rather than only familiarity with an item, which is assumed to depend on the perirhinal cortex (Eichenbaum et al., 2007). Accordingly, we computed hit rates (i.e., recognizing an item as "surely old") and category-based false alarm rates referring to conditioning on category level. (CS^+ vs. CS^-; please note that new items of the category that had been a CS^+ during encoding have never been paired with the shock.) We fitted models using different sets of independent variables, including subsequent PEs, previous PEs, anticipatory and outcome-related arousal, and the explicit shock prediction. To further elucidate the effects of PEs on episodic memory, we computed *subsequent PEs* and *previous PEs* to investigate retro- and prospective effects of PEs on subsequent memory. In the analysis of retrospective PE effects, subsequent PEs were linked to the memory of the preceding stimulus in the same trial, that is, PE of Trial 3 referred to the predictive item of Trial 3. In the analysis of potential prospective PE effects, previous PEs were linked to the memory of the stimulus in the following trial, that is, the PE in Trial 3 was used as predictor for subsequent memory of the item presented in Trial 4.

Transparency and Openness

The materials, data, and R analysis scripts are publicly available on the Research Data Management System of University of Hamburg and can be accessed at https://www.fdr.uni-hamburg.de/record/14147 (Loock et al., 2024). This experiment was not preregistered.

Results and Discussion

Successful Fear Conditioning

An analysis of SCR data confirmed that fear was successfully induced for CS^+ items (see Figure 2A). On average, participants showed significantly higher anticipatory SCRs to CS^+ items (M = 0.21, SD = 0.01) compared to CS^- items (M = 0.17, SD = 0.01); $t(76) = 3.87$, $p < .001$, $d = 0.42$. Furthermore, outcome-related SCRs were significantly higher for shocked items (M = 0.42, SD = 0.01) compared to unshocked items (M = 0.10, SD = 0.01), $t(76)$ = 21.29, $p < .001$, $d = 3.15$.

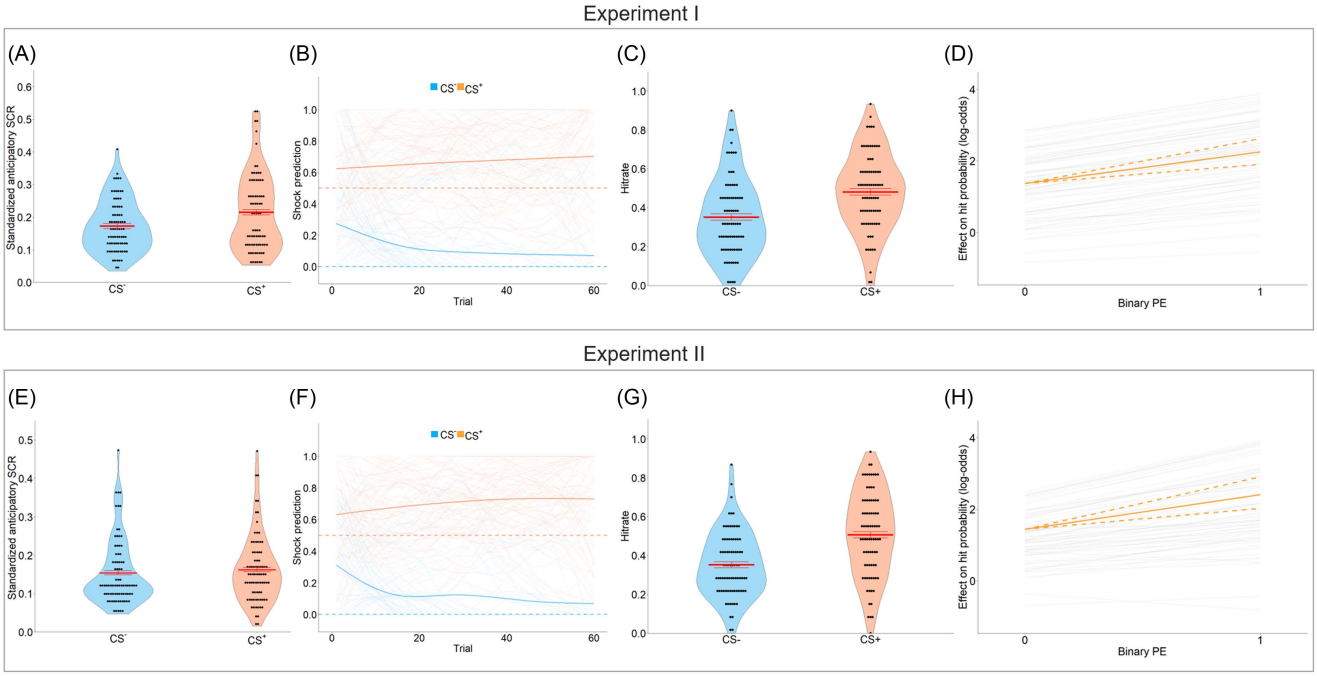
Explicit shock ratings showed that participants learned the shock contingencies very well. On average, incorrect predictions were made in 29% (SD = 0.11) of all trials with substantially more PEs for CS^+ (M = 0.47, SD = 0.10) compared to CS^- items (M = 0.11, SD = 0.16), $t(77) = 19.05$, $p < .001$, $d = 2.71$ (see Figure 2B). As expected, PEs decreased as the task progressed, $r(58) = -0.53$, $p < .001$.

General Memory Performance

As expected, participants on average were moderately surprised by the recognition test (see Supplemental Material).

Overall, participants performed very well in the recognition task, as indicated by significantly higher hit rates (M = 0.42, SD = 0.17) than category-based false alarm rates (M = 0.23, SD = 0.10), $t(77)$ = −7.54, $p < .001$, $d = 1.37$ (see Figure 2C). Importantly, the category-based false alarm rate for the CS^+ items (M = 0.23, SD = 0.12) was comparable to the false alarm rate for items from the CS^- category (M = 0.23, SD = 0.13), $t(77) = 0.51$, $p = .611$, $d = 0.07$. Note that new items have never been paired with a shock and false alarms relate to the whole category. As expected, the average hit rate for items from the CS^+ category (M = 0.48, SD = 0.20) was significantly higher than for items from the CS^- category (M = 0.35,

Figure 2
Results of Experiments 1 and 2



Note. Anticipatory skin conductance responses and hit rates were always significantly higher for items from the CS^+ category compared to items from the CS^- category in both Experiments 1 (A, C) and 2 (E, G), confirming that the fear conditioning procedure was successful. In addition, shock contingencies were learned well in Experiments 1 (B) and 2 (F; thick lines) approaching the underlying shock contingencies (dotted lines). Generalized linear mixed models revealed that subsequent prediction errors (PEs) significantly enhanced recognition memory for predictive stimuli in both experiments (D, G). Black dots show individual data. Thick-red bar represents group mean, whereas thin-red bars show ±1 standard error of the mean. SCR = skin conductance response; CS = conditioned stimulus; PE = prediction errors. See the online article for the color version of this figure.

$SD = 0.20$), $t(77) = 5.59$, $p < .001$, $d = 0.65$. These finding dovetails with the assumption that memory advantages for CS^+ items are attributed to increased physiological arousal. The pattern of results was comparable when using the signal detection theory-based parameter d' (see Supplemental Material).

Modeling Recognition Performance at Item Level

So far, we showed that CS^+ items were better remembered after 24 hr than CS^- items. To test whether PEs drive the emotional memory enhancement, we computed GLMMs at item level treating the binary recognition of an item as the dependent variable.

In a first minimal model, we tested whether uPEs that followed a CS contribute to subsequent recognition of this CS item. Therefore, we treated the binary subsequent PE as the sole independent variable to predict the binary recognition of an item. This revealed that episodic memory was indeed enhanced for trials in which an incorrect shock prediction for the predictive target has been made ($z = 4.90$, $p < .001$, $\beta = 0.90$; see Figure 2D). To rule out that those PE effects were due to the explicit shock prediction, we also computed a model that included the explicit shock prediction and the binary subsequent PE as independent variables to predict the binary item recognition. Critically, even after controlling for the shock prediction we found a memory enhancing effect of subsequent PEs in this model ($z = 4.72$, $p < .001$, $\beta = 0.76$). To investigate the possibility that the effects of physiological arousal and the effects of PEs on memory might be confounded, we added both measures of arousal (i.e., anticipatory and outcome related SCRs) to the minimal model that featured only the binary subsequent PE as the sole independent variable. This model revealed no significant effect of anticipatory SCRs on item recognition ($z = -1.47$, $p = .142$, $\beta = -0.54$). Larger outcome-related SCRs, on the other hand, were associated with better item recognition ($z = 4.20$, $p < .001$, $\beta = 1.45$). For subsequent PEs, our results showed that recognition was improved significantly ($z = 4.36$, $p < .001$, $\beta = 0.76$), suggesting that subsequent PEs enhanced memory even after controlling for arousal.

In a next step, we tested whether uPEs that preceded an item and were actually related to the previous item may exert also prospective effects, contributing to the recognition of the item following the PE. To this end, we treated the binary previous PE as the sole independent variable to predict the binary recognition of the following item. This revealed that episodic memory was indeed enhanced for items following a PE, $z = 3.48$, $p < .001$, $\beta = 0.57$. In addition, we also computed a model that included the explicit shock prediction and the previous PE as independent variables to predict the binary item recognition. Critically, even after controlling for the shock prediction the memory enhancing effect of previous PEs remained in this model ($z = 2.68$, $p = .007$, $\beta = 0.37$). Same as in the analysis of subsequent PE effects, we added both measures of arousal (i.e., anticipatory and outcome related SCRs) to the minimal model that featured only the binary previous PE as the sole

independent variable to rule out confounds with physiological arousal. This revealed no significant effect of anticipatory SCRs on item recognition ($z = -1.48$, $p = .140$, $\beta = -0.53$). Larger outcome-related SCRs, on the other hand, were associated with better item recognition ($z = 4.19$, $p < .001$, $\beta = 1.50$). For previous PEs, there was a strong trend in the direction of memory enhancement, which did, however, not reach significance anymore ($z = 1.95$, $p = .051$, $\beta = 0.29$).

For models determining whether subsequent PEs and previous PEs reflect distinct mechanisms see Supplemental Material.

Conclusion

The findings of this experiment replicate the previously reported enhancing effects of PEs associated with aversive events on subsequent memory for the stimulus encountered before the PE. In line with these previous reports (Kalbe & Schwabe, 2020, 2022b), these PE effects could not be explained by mere increases of arousal. Interestingly, beyond these retrospective memory enhancements of PEs, we obtained also first evidence that PE enhance memory not only for items preceding the PE but also for items that followed a PE. In contrast to the retrospective effects of PEs, however, these prospective PE effects appeared to be at least partly driven by (outcome-related) arousal.

Experiment 2: Replicating the Retrospective and Prospective Effects of PEs Related to Aversive Events on Subsequent Memory

Experiment 1 provided initial evidence for a prospective effect of PEs on subsequent memory for stimuli encountered shortly after the PE. Experiment 2 served to replicate this prospective PE effect on memory, as well as the retrospective PE effect that was shown to go beyond the well-established effects of arousal on memory.

Method

Participants

Eighty-four healthy volunteers participated in this study (age: $M =$ 25.17 years, $SD = 4.26$ years, range = 18–34 years). Exclusion criteria were the same as those in Experiment 1. Three participants were excluded from the analyses due to technical failure during the experiment, resulting in a final sample of $n = 81$. None of the participants had participated in Experiment 1. This sample is part of a larger study on emotional learning processes (Kalbe & Schwabe, 2022a).

The target sample size was based on an a priori power calculation with identical parameters as in Experiment 1, showing that a sample of 67 participants is sufficient to detect a medium-sized effect ($d_z = .45$) of subsequent PEs with a power of .95. In addition, we performed a post hoc power simulation using the R-package simR (Green & MacLeod, 2016). For the observed subsequent PE effects and our final sample size of 81 participants, it yielded a power of 0.99 based on 1,000 simulations.

All participants provided written informed consent before participation and received a monetary reimbursement of 30€ at the end of the study. The study was approved by the ethics committee of the Faculty of Psychology and Human Movement Science at the University of Hamburg and carried out in line with the Declaration of Helsinki.

Materials

We used the same stimulus set as in Experiment 1. The stimulus set consisted of 60 photographs (i.e., 30 animals and 30 tools) on the first experimental day and 180 photographs that were used as lures in the recognition test on the second experimental day. Same as in Experiment 1, the order of stimulus presentation was randomized across participants. The assignment of photograph category (i.e., tools or animals) as CS^+ or CS^- was counterbalanced across participants.

Procedure

The procedure of Experiment 2 was largely identical to the procedure of Experiment 1, except that we changed the timing of the electric shock during the incidental encoding-fear learning session to make our procedure more comparable with previous learning paradigms (Dunsmoor et al., 2015; see Kalbe & Schwabe, 2022a). Specifically, in Experiment 2, a 200-ms-electric shock occurred 4.3 s after stimulus onset and thus coterminated with the predictive stimulus during the learning task.

Data Analysis

The statistical analysis was identical to Experiment 1.

Transparency and Openness

The materials, data, and R analysis scripts are publicly available on the Research Data Management System of University of Hamburg and can be accessed at https://www.fdr.uni-hamburg.de/record/14147 (Loock et al., 2024). This experiment was not preregistered.

Results and Discussion

Successful Fear Conditioning

Descriptively, participants showed higher anticipatory SCRs to CS^+ items ($M = 0.16$, $SD = 0.01$) compared to CS^- items ($M = 0.15$, $SD = 0.01$). However, this descriptive difference was not statistically significant, $t(80) = 1.03$, $p = .308$, $d = 0.09$ (see Figure 2E). For an additional through-to-peak analysis (Boucsein, 1992; Kalbe & Schwabe, 2022a, see Supplemental Material).

Outcome-related SCRs were significantly higher for shocked items ($M = 0.34$, $SD = 0.01$) compared to unshocked items ($M = 0.10$, $SD = 0.01$), $t(80) = 16.46$, $p < .001$, $d = 2.21$.

Explicit shock ratings showed that participants learned the shock contingencies very well. On average, incorrect predictions were made in 29% ($SD = 0.12$) of all trials with substantially more PEs for CS^+ items ($M = 0.45$, $SD = 0.09$) compared to CS^- items ($M = 0.12$, $SD = 0.18$), $t(80) = 17.40$, $p < .001$, $d = 2.41$ (see Figure 2F). Similarly to Experiment 1, PEs decreased as the task progressed, $r(58) = -0.48$, $p < .001$.

General Memory Performance

As expected, Participants on average were moderately surprised by the recognition test (see Supplemental Material).

Again, participants performed very well in the recognition task, as indicated by significantly higher hit rates ($M = 0.43$, $SD = 0.16$) than category-based false alarm rates ($M = 0.26$, $SD = 0.10$), $t(80) = -7.77$, $p < .001$, $d = 1.25$ (see Figure 2G). Importantly, the category-based false alarm rate for the CS^+ items ($M = 0.27$, $SD = 0.13$) was comparable to the false alarm rate for items from the CS^- category ($M = 0.26$, $SD = 0.14$), $t(80) = 0.69$, $p = .494$, $d = 0.10$. As expected, the average hit rate for items from the CS^+ category ($M = 0.51$, $SD = 0.22$) was significantly higher than for items from the CS^- category ($M = 0.35$, $SD = 0.17$), $t(80) = 6.84$, $p < .001$, $d = 0.81$, in line with the findings of Experiment 1. The pattern of results was comparable when using d' (see Supplemental Material).

Modeling Recognition Performance at Item Level

To further elucidate whether PEs enhance memory for preceding and subsequent stimuli, we computed the same GLMMs as in Experiment 1 to predict the binary recognition of an item.

We started with a minimal model, in which we tested whether uPEs that followed a CS contribute to subsequent recognition of this CS item. Therefore, we treated the binary subsequent PE as the sole independent variable to predict the binary recognition of an item. Again, this analysis revealed that memory was enhanced for trials in which a PE occurred ($z = 4.58$, $p < .001$, $\beta = 1.02$; see Figure 2H). To rule out that those PE effects were only due to the explicit shock prediction, we also computed a model that included the explicit shock prediction and the binary subsequent PE as independent variables to predict the binary item recognition. Critically, even after controlling for the shock prediction we were able to replicate the memory enhancing effect of subsequent PEs in this model ($z = 4.49$, $p < .001$, $\beta = 0.78$). To investigate the possibility that the effects of physiological arousal and the effects of PEs on memory are confounded, we added both measures of arousal (i.e., anticipatory and outcome related SCRs) to the minimal model that featured only the binary subsequent PE as the sole independent variable. This analysis showed no significant effect of anticipatory SCRs on item recognition ($z = 1.26$, $p = .209$, $\beta = 0.42$, nor of outcome-related SCRs, $z = 1.53$, $p = .127$, $\beta = 0.49$). Most importantly and in line with Experiment 1, our results showed for subsequent PEs that recognition was significantly enhanced ($z = 4.95$, $p < .001$, $\beta = 0.87$), suggesting that subsequent PEs enhanced memory even after controlling for arousal.

Experiment 1 provided initial evidence for prospective effects of PEs, that is, memory enhancing effects of PEs for stimuli encoded *after* the PE. To test whether we can replicate this effect, we treated the binary previous PE as the sole independent variable to predict the binary recognition of the following item. This analysis revealed that episodic memory was indeed enhanced for items following an incorrect prediction (i.e., a PE; $z = 5.03$, $p < .001$, $\beta = 0.74$). In addition, we also computed a model that included the explicit shock prediction and the previous PE as independent variables to predict the binary item recognition. Critically, even after controlling for the shock prediction, we obtained a memory enhancing effect of previous PEs in this model ($z = 3.46$, $p < .001$, $\beta = 0.50$). Same as in Experiment 1, we added both measures of arousal (i.e., anticipatory and outcome-related SCRs) to the minimal model that featured only the binary previous PE as the sole independent variable to rule out confounds with physiological arousal. Interestingly, this model revealed neither a significant effect of anticipatory SCR ($z = 1.57$, $p = .116$, $\beta = 0.61$) nor of outcome-related SCR on item recognition ($z = 1.19$, $p = .234$, $\beta = 0.39$). Importantly, however, we obtained a significant effect of previous PEs on item recognition in this model ($z = 4.73$, $p < .001$, $\beta = 0.69$).

For models determining whether subsequent PEs and previous PEs reflect distinct mechanisms, see Supplemental Material.

Conclusion

The findings of Experiment 2 replicate the previously reported enhancing effects of (unsigned) PEs associated with aversive events on subsequent memory for the stimulus encountered before the PE. Moreover, our findings further replicate the prospective memory enhancement induced by PEs. In contrast to Experiment 1, which suggested that the prospective PE effects on memory may be due to arousal, the findings of Experiment 2 show that both the retrospective PE effects and the prospective PE effects on memory were independent of physiological arousal, as measured by SCR. The failure to replicate the effects of arousal on memory that we observed in Experiment 1 suggests that arousal effects on memory might be less robust than those of PEs on memory.

Experiment 3: Is the PE-Induced Memory Enhancement Time Dependent?

In line with previous studies (Kalbe & Schwabe, 2020; Kalbe & Schwabe, 2022b; for a review, see Rouhani et al., 2023), the results of Experiments 1 and 2 show consistently that PEs associated with aversive events enhance subsequent memory. Furthermore, the existing evidence suggests that these PE effects on memory go beyond the well-established arousal effects on memory. Notably, the results of Experiments 1 and 2 indicated further that PE effects may extend to stimuli following the PE. In Experiment 3, we aimed to determine to what extent PE effects on memory are time dependent. More specifically, we tested (a) to what extent PE effects on memory depend on the time interval between the stimulus and the PE and (b) whether PE effects on memory are dependent on the interval between encoding and testing, that is, whether these PE effects emerge during memory encoding or consolidation. To these ends, we varied in Experiment 3 the CS-outcome delay between 0 and 10 s and tested recognition memory either 24 hr after encoding, as in Experiments 1 and 2 (and all previous studies on aversive PEs and memory), or immediately after encoding.

In addition to these timing-related modifications, we modified our experimental paradigm in two ways: first, we presented three CSs, with one CS^- and two CS^+ with different contingencies. Second, participants were asked to rate their shock expectancy on a continuous scale from 0 to 100. Both of these modifications served to achieve an adequate distribution of both positive and negative PEs, which have been suggested to exert distinct effects on memory formation (Kalbe & Schwabe, 2022b; Rouhani et al., 2023; Rouhani & Niv, 2021).

Method

Participants

One hundred twenty-three healthy volunteers participated in this study (age: $M = 25.95$ years, $SD = 4.33$ years, range = 18–35 years).

Exclusion criteria were identical to those in Experiments 1 and 2. Five participants were excluded from the analyses due to technical failure during the experiment or because they did not return for the second experimental day, resulting in a final sample of $n = 118$. Importantly, none of the participants had participated in Experiment 1 or 2.

The target sample size was based on previous findings of signed aversive PE effects on episodic memory formation (Kalbe & Schwabe, 2022b). Because we modified the experimental design, in particular by adding a variable time interval between CS and outcome and by adding the between-factor retention interval (immediate vs. 24 hr delay), we doubled the reported sample size of Kalbe and Schwabe (2022b). Thus, we expected a sample of 120 participants to be sufficient to detect a power of at least 0.90. In line with these assumptions, a post hoc power simulation using the R-package simR (Green & MacLeod, 2016) for the observed effects of the subsequent sPE and our final sample size of 118 participants yielded a power of 0.92 based on 1,000 simulations.

All participants provided written informed consent before participation and received a moderate monetary reimbursement (up to 50 €) at the end of the study. The study was approved by the ethics committee of the Faculty of Psychology and Human Movement Science at the Universität Hamburg and carried out in line with the Declaration of Helsinki.

Materials

In Experiment 3, we used stimuli from Kalbe and Schwabe (2022b), but added more stimuli due to the increased trial number. Stimuli were taken from existing image databases, that is, Bank of Standardized Stimuli (Brodeur et al., 2010, 2014), McGill Calibrated Color Image Database (Olmos & Kingdom, 2004), SUN database (Xiao et al., 2010), Konklab (Konkle et al., 2010), and open-online sources. In total, the stimulus set consisted of 810 pictures of vehicles, tools, and clothes, isolated on white background. All stimuli were assumed to be emotionally neutral and represented a unique exemplar of its category. From this pool, 120 pictures of vehicles, 120 pictures of tools, and 120 pictures of clothes were randomly drawn and used during encoding on the first experimental day. From the remaining 450 pictures, 180 randomly chosen pictures (60 pictures per category) served as lures for the surprise recognition test on the second experimental day. The allocation of images as encoding items or lures was randomized across participants and thus unique per participant. In addition, the order in which individual items were presented was randomized across participants.

Procedure

The experiment consisted of two parts, with an encoding session on the first experimental day and a recognition test on the second day (see Figure 3A). Depending on the experimental group, the recognition test took place either immediately after the encoding session or 22–26 hr later. Participants were pseudorandomly assigned to the two groups, to achieve a comparable number of men and women per group, immediate group: 20 men, 36 women, M_{age}: 25.93 ($SD = 4.13$); 24-hr-delay group: 23 men, 37 women, M_{age}: 25.97 ($SD = 4.54$). Upon arrival on the first experimental day, participants provided written informed consent and received written instructions that they were going to see a series of photographs of vehicles, tools, and clothes and that some of them might be followed by a brief electric shock. They were then instructed to predict how likely a shock would be to follow the current picture (see Figure 3A). Therefore, they were requested to adjust the slider on the screen to a value corresponding to their prediction of the shock probability (ranging from 0% to 100%). Importantly, participants were neither told about the shock contingencies, nor that their memory would be tested later on. They were informed that their predictions would not affect the probability that a shock would occur, but that they should aim at improving their predictions over the task. Unbeknownst to the participants, the probabilities of a shock were linked to the image categories. One category served as CS^{a+} (67% shock probability), one as CS^{b+} (33% shock probability), and one as CS^{-} (0% shock probability). The assignment of image categories (i.e., vehicles, tools, and clothes) to the CS categories (i.e., CS^{a+}, CS^{b+}, and CS^{-}) was counterbalanced across participants and groups.

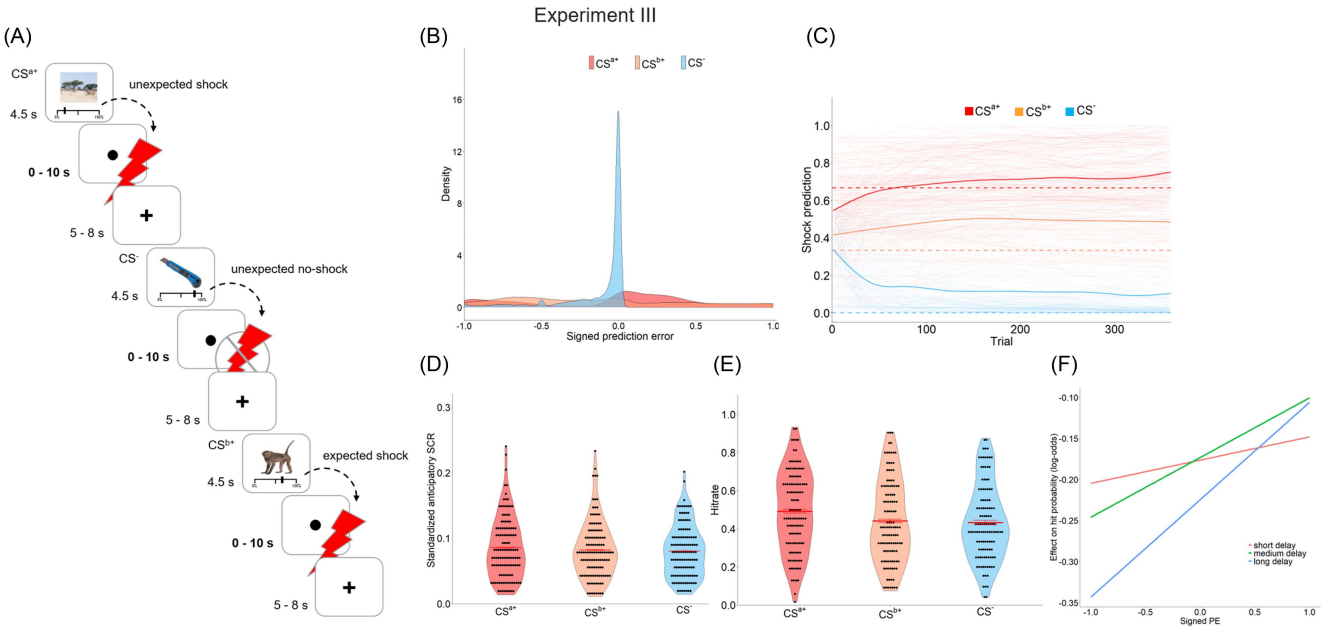
To measure SCRs as indicator of physiological arousal through the incidental encoding-fear learning task, we placed disposable, pregelled snap-electrodes on the thenar and hypothenar eminence of the left hand (see Kalbe & Schwabe, 2022b). Skin conductance was measured using the MP-160 BIOPAC system (BIOPAC systems, Goleta, California, United States). For electrical stimulation, we used the STM-100C module connected to the MP-160. A stimulation electrode was placed on the back of participants' right lower leg, approximately 20 cm above the ankle. Before the learning task, stimulation intensity was adjusted individually to be unpleasant but not painful as described in Experiment 1.

The encoding session on the first experimental day consisted of four blocks with 90 trials each. In each block, 30 pictures of vehicles, 30 pictures of tools, and 30 pictures of clothes were presented in a pseudorandomized order, so that no more than three pictures of the same category appeared in a row. On each trial, a picture was shown in the center of a computer screen for 4.5 s, during which participants were asked to make their prediction about the probability of an electric shock (Figure 3A). Therefore, a slider was presented beneath each image which could be adjusted to any integer value between 0% and 100% using the computer mouse. After stimulus offset, a black dot was presented centrally on the screen which coterminated with the 200 ms-outcome (shock vs. no-shock). Critically, the duration of the dot's presentation ranged randomly between 0 and 10 s per trial to vary the critical CS-outcome delay. Each trial was followed by a black fixation cross centered on gray background for 6.5 ± 1.5 s, which enabled us to measure the relatively slow SCRs elicited by the pictures and the shocks. Between blocks there were short breaks (1–2 min) during which participants had the chance to recalibrate the shock intensity, if required. Each encoding block lasted approximately 25 min, resulting in a total duration of 100 min for the entire incidental encoding-fear learning session.

The surprise recognition test took place either immediately after the encoding session or on the next day, 22–26 hr later, depending on the experimental group. Same as in Experiments 1 and 2, participants completed a short questionnaire to assess whether they anticipated a memory test and then rated how surprised they were about the recognition test on a scale from 1 (*not surprised at all*) to 5 (*very surprised*). In the recognition test, participants saw all pictures they had seen during the encoding session (120 pictures of vehicles, 120 pictures of tools and 120 pictures of clothes) as well as 180 "new" pictures (60 pictures of vehicles, 60 pictures of tools, and

Figure 3

Experimental Procedure and Results of Experiment 3



Note. In the encoding task (A), participants saw a series of unique pictures from three different categories (animals, scenes, and tools) linked to fixed probabilities to receive an electric shock (CS^{a+} = 67%, CS^{b+} = 33%, and CS^{-} = 0%). On each trial, participants indicated their shock expectation on a continuous scale from 0% to 100%. The delay with which the outcome (shock vs. no-shock) occurred after stimulus-offset was varied between 0 and 10 s. (B) PEs were equally distributed around zero. (C) Participants' mean shock expectancy ratings (thick lines) approached the true shock probabilities (dotted lines) relatively fast. (D, E) Mean standardized anticipatory SCR and hit rates confirmed successful fear conditioning, as reflected in significantly elevated SCR and increased hit rates of CS^{a+} compared with CS^{-} items. Black dots show data from individual participants. Thick-red bar represents group mean, whereas thin-red bars show ±1 standard error of the mean. (F) The CS-outcome delay did not interact significantly with the subsequent sPE-effect on item memory suggesting that PEs seem to be independent of the time between stimulus and outcome. Pictures taken and reprinted from "Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos," by M. B. Brodeur, K. Guérard,and M. Bouras, 2014, *PLOS ONE, 9*(9), Article e106953 (https://doi.org/10.1371/journal.pone.0106953). CC BY 4.0 and from "SUN Database: Large-Scale Scene Recognition From Abbey to Zoo," by J. Xiao, J. Hays, K. A. Ehinger, A. Oliva,and A. Torralba, *2010 IEEE computer society conference on computer vision and pattern recognition* (pp. 3485–3492), 2010 (https://doi.org/10.1109/CVPR.2010.5539970). CC BY 4.0. CS = conditioned stimulus; SCR = skin conductance response; PE = prediction errors. See the online article for the color version of this figure.

60 pictures of clothes) that had not been presented before in a randomized order. Each trial started with a central white fixation cross on a white background for 1.5 ± 0.5 s, followed by an "old" or "new" picture presented centrally on the computer screen for 6 s. For each item, participants were instructed to indicate whether the currently presented picture was *definitely old*, *maybe old*, *maybe new*, or *definitely new* by pressing the "1," "2," "3," or "4" button on the keyboard, respectively.

Data Analysis

For each trial, we derived uPEs which were calculated as the absolute value of the difference between participants' continuous explicit shock expectancy ratings (ranging from 0, corresponding to full confidence that no shock would occur, to 1, corresponding to full confidence that a shock would occur) and the actual binary outcome of the trial (coded 0 when no shock occurred and 1 when a shock occurred in the current trial). The resulting uPE is, therefore, ranging between 0 and 1. Because the modified paradigm of Experiment 3 allowed us to measure continuous PEs we derived also an sPE. We focused our analyses on sPE effects because these represent a more accurate measure of the PE and allow a distinction between positive and negative PEs. Analyses using the uPE as predictor are presented in the Supplemental Material.

The sPE was calculated as the relative value of the difference between the binary outcome of a trial (1 for shock and 0 for nonshock) and the explicit shock prediction and could take any integer value between −1 and 1. Importantly, the sign of the sPE contained information about the outcome's value: A negative sPE (sPE < 0) could only occur in unshocked trials corresponding to unexpected shock omissions, whereas positive sPEs (sPE > 0) could only occur in shocked trials corresponding to unexpected shock occurrence.

Again, SCRs were analyzed using continuous decomposition analysis in Ledalab Version 3.4.9 (Benedek & Kaernbach, 2010) and calculated in the same way as in Experiments 1 and 2. Deviations were due to altered response windows as a consequence of the variable CS-outcome delay (see Supplemental Material).

Again, we fitted GLMMs with a logit link function using the lme4 R package (Bates et al., 2015) and treated subjects as random effects for both the intercept and all slopes of the fixed effects included in the model (Barr et al., 2013). The recognition of an individual

item was treated as the binary dependent variable, coded "0" for misses and "1" for confident hits. We fitted models using different sets of independent variables, including subsequent PEs, previous PEs, anticipatory and outcome-related arousal, the explicit shock prediction, the retention interval, and the CS-outcome delay. In line with Experiments 1 and 2, we also distinguished between subsequent and previous PEs with the former referring to PE effects on the recognition of the preceding stimulus and the latter referring to PE effects on the recognition of the following stimulus. If not indicated otherwise, all analyses were collapsed across both retention intervals.

Transparency and Openness

The materials, data, and R analysis scripts are publicly available on the Research Data Management System of University of Hamburg and can be accessed at https://www.fdr.uni-hamburg.de/record/14147 (Loock et al., 2024). This experiment was not preregistered.

Results and Discussion

Successful Fear Conditioning

SCR data confirmed the expected fear learning process. Specifically, anticipatory SCR differed significantly between conditioning categories, $F(2, 234) = 4.18$, $p = .016$, partial $\eta^2 = 0.002$ (Figure 3D). Post hoc paired *t*-tests showed that participants showed higher anticipatory SCRs to CS^{a+} items ($M = 0.09$, $SD = 0.01$) compared to CS^{-} items ($M = 0.08$, $SD = 0.02$), $t(117) = 2.53$, $p = .012$, $d = 0.11$. Anticipatory SCRs did not differ significantly between CS^{b+} items ($M = 0.08$, $SD = 0.01$) and CS^{-} items, $t(115) = 0.79$, $p = .434$, $d = 0.04$. Outcome-related SCRs were significantly higher for shocked trials ($M = 0.19$, $SD = 0.10$) compared to unshocked trials ($M = 0.05$, $SD = 0.10$), $t(115) = 10.83$, $p < .001$, $d = 1.40$.

Explicit shock ratings further showed that participants learned the shock contingencies over the task very well (see Figure 3C). Participants had a significantly higher shock expectancy for CS^{a+} ($M = 0.71$, $SD = 0.17$) compared to CS^{b+} ($M = 0.47$, $SD = 0.12$), $t(117) = 13.27$, $p < .001$, $d = 1.52$, and for CS^{b+} compared to CS^{-} ($M = 0.13$, $SD = 0.17$); $t(117) = 18.16$, $p < .001$, $d = 2.11$. In addition, PEs were equally distributed around zero (Figure 3B) suggesting a sufficient number of positive and negative PEs that could be analyzed.

General Memory Performance

Again, participants were moderately surprised by the recognition test (see Supplemental Material).

Overall, participants performed very well in the recognition task, as indicated by significantly higher hit rates ($M = 0.46$, $SD = 0.21$) than category-based false alarm rates ($M = 0.21$, $SD = 0.16$), $t(353) = 17.02$, $p < .001$, $d = 1.29$.

Importantly, while false alarm rates were comparable between conditioning categories, $F(2, 234) = 2.34$, $p = .098$, partial $\eta^2 = 0.004$, hit rates differed significantly between conditioning categories, $F(2, 234) = 9.61$, $p < .001$, partial $\eta^2 = 0.016$ (see Figure 3E) suggesting that memory but not the response bias differed between categories. Post hoc paired *t*-tests revealed that the average hit rate for CS^{a+} items ($M = 0.49$, $SD = 0.11$) was significantly higher than for CS^{b+} items ($M = 0.44$, $SD = 0.10$), $t(117) = 3.97$, $p < .001$, $d = 0.24$, and CS^{-} items ($M = 0.43$, $SD = 0.12$); $t(117) = 3.75$, $p < .001$, $d = 0.29$. The average hit rate for CS^{b+} items did not differ from the hit rate for CS^{-} items, $t(117) = 0.52$, $p = .605$, $d = 0.04$. The pattern of results was comparable when using d' (see Supplemental Material).

Moreover, when taking the retention interval into account, recognition memory performance differed significantly between the immediate and 24 hr delayed groups, as expected. Participants who underwent the recognition test immediately after the encoding session (hit rate: $M = 0.53$, $SD = 0.13$; d': $M = 1.53$, $SD = 0.71$) had a significantly better recognition memory than participants who performed the recognition test about 24 hr after encoding, hitrate: ($M = 0.38$, $SD = 0.14$), $t(348.6) = 7.05$, $p < .001$, $d = 0.74$; d': ($M = 1.24$, $SD = 0.57$), $t(329.11) = 4.24$, $p < .001$, $d = 0.43$, reflecting the well-known decline in memory over time. There was no interaction between CS type and retention interval, hit rate: $F(2, 232) = 0.19$, $p = .830$, partial $\eta^2 = 0.000$; d': $F(2, 222) = 1.43$, $p = .241$, partial $\eta^2 = 0.004$, suggesting that the differential memory performance for stimuli from the three CS categories did not differ between the immediate and 24-hr-delayed groups.

Modeling Recognition Performance at Item Level

To elucidate the mechanisms of episodic memory formation, we again fitted GLMMs with recognition of an item as the binary dependent variable and added certain independent predictors in a step-wise manner, similarly to Experiments 1 and 2.

We started with a first minimal model, in which we tested whether trial-wise subsequent sPEs contribute to later recognition. Therefore, we treated the sPE (ranging from −1 to 1) following on a CS as the sole independent variable to predict the binary recognition of this CS item. Estimates obtained revealed that sPEs ($z = 2.53$, $p = .012$, $\beta = 0.08$) showed a positive relationship with item recognition. To rule out that the PE effects were confounded with the shock prediction, we also computed a model where we added the explicit shock prediction as a predictor to the previous model. When controlling for the explicit shock prediction, the memory enhancing effect of the subsequent sPEs remained significant ($z = 4.20$, $p < .001$, $\beta = 0.15$).

In a follow-up model, we added anticipatory arousal and outcome-related arousal as predictors. While anticipatory SCRs were associated with decreased memory ($z = -3.95$, $p < .001$, $\beta = -0.42$), outcome-related SCRs did not influence item recognition significantly ($z = 1.05$, $p = .295$, $\beta = 0.12$). After controlling for arousal effects on memory, the sPE effect did not reach statistical significance anymore ($z = 1.44$, $p = .150$, $\beta = 0.05$).

To examine whether sPE effects on memory are dependent on the retention interval and the CS-outcome delay, we included the retention interval and the CS-outcome delay as well as their interaction as predictors in an additional set of models.

First, we set up a model that treated the subsequent sPE and the retention interval and their interaction as independent variables to predict the binary recognition of an item. When controlling for the retention interval, we obtained a significant effect of subsequent sPEs on memory ($z = 2.04$, $p = .041$, $\beta = 0.09$) and a nonsignificant *Retention Interval* × *Subsequent sPE* interaction ($z = -0.32$, $p = .753$, $\beta = -0.02$), suggesting that these "retrospective" sPE effects are not dependent on the interval between encoding and test.

In a next step, we set up a model that treated the subsequent sPE, the CS-outcome delay and their interaction as independent variables to predict the binary recognition of an item. This revealed that memory was neither influenced by the CS-outcome delay ($z = -1.57$, $p = .117$, $\beta = -0.01$) nor by the *CS-Outcome Delay* × *Subsequent sPE* interaction ($z = 1.56$, $p = .120$, $\beta = 0.01$). These findings suggest that the CS-outcome delay does not influence memory and does not modulate the subsequent sPE effects on memory (see Figure 3F).

In an additional model, we treated the CS-outcome delay, the retention interval, the sPE and their interaction as independent variables. Estimates obtained showed no significant *CS-Outcome Delay* × *Retention Interval* × *sPE* interaction ($z = 0.70$, $p = .481$, $\beta = 0.01$).

Next, we performed additional models in which we treated the previous sPE as the sole independent variable to predict the binary recognition of the following item. This revealed no significant effect of previous sPEs on memory for items following the PE ($z = 0.21$, $p = .835$, $\beta = 0.00$). Again, we added anticipatory and outcome-related SCRs to the former model to investigate confounds with physiological arousal. This revealed a significant impairing effect of anticipatory SCRs on item recognition ($z = -4.54$, $p < .001$, $\beta = -0.47$), whereas we obtained no effect of outcome-related SCRs on memory formation ($z = 1.67$, $p = .090$, $\beta = 0.16$). The previous sPE effect remained nonsignificant ($z = 0.57$, $p = .570$, $\beta = 0.01$). To rule out that the previous sPE effects were confounded with the shock prediction, we also computed a model where we added the explicit shock prediction and the previous sPE as predictors. When controlling for the explicit shock prediction, the effect of the previous sPE remained nonsignificant ($z = 0.58$, $p = .563$, $\beta = 0.01$).

In a next step, we tested whether the effect of sPEs on the memory for items following the sPE are dependent on the retention interval. Estimates obtained revealed that neither the (previous) sPE influence recognition of the following item significantly ($z = 0.26$, $p = .795$, $\beta = 0.01$) nor the *Retention Interval* × *Previous sPE* interaction ($z = -0.34$, $p = .737$, $\beta = -0.01$). Notably, there was a significant effect of the retention interval on CS recognition ($z = 4.35$, $p < .001$, $\beta = 0.33$).

For models investigating whether subsequent PEs and previous PEs reflect distinct mechanisms, see Supplemental Material.

Effects of uPEs

We found no significant effects of subsequent uPEs on episodic memory, suggesting that these are weaker when the CS-outcome delay is extended. For previous uPEs, we found a memory enhancing effect even when controlling for arousal. These prospective effects seemed to be unaffected by the retention interval, but dependent on the CS-outcome delay. For a detailed analysis of the uPE effects in Experiment 3, see Supplemental Material.

Conclusion

The findings of Experiment 3 replicate previously reported enhancing effects of (signed) PEs associated with aversive events on memory for the stimulus encountered before the PE. Our findings, however, show no prospective (signed) PE effects on memory, whereas such prospective effects are observed for uPE, in line with Experiments 1 and 2. Interestingly, we found that the retrospective sPE effects on episodic memory seem to be (a) independent of the time interval between the stimulus and the PE and (b) emerge already when memory is tested shortly after encoding, suggesting that they are not consolidation dependent. Thus, we obtain first evidence for retrospective PE effects to be time resistant and that they rather emerge during encoding than during consolidation processes. Notably, the results of Experiment 3 showed that the sPE effects were at least partly related to arousal because these effects disappeared when we controlled for arousal. The latter might be due to the extended CS-outcome interval.

Experiment 4: Is the PE-Related Memory Enhancement Exclusive to Predictive Stimuli?

Experiments 1–3 showed consistently that PEs following inherently neutral events enhance the subsequent memory of these events. A key question is whether this PE-related memory enhancement is specific to the (predictive) stimuli that evoked the PE or whether the PE opens a transient window of enhanced memory encoding for all events that occur in this time window. The findings of Experiment 3 revealed that the PE-related memory enhancement is still observed even when the CS-outcome delay is up to 10 s long, which might point to a window of enhanced encoding around a PE that is at least 10 s long or that a PE can be specifically linked to a predictive stimulus that occurred up to 10 s before. The prospective PE effects on the encoding of stimulus encountered after the PE might point to the possibility of a transient window of enhanced memory encoding around a PE. However, this prospective PE effect was not consistently found across studies. In Experiment 4, we aimed to determine whether PEs are exclusively linked to the predictive stimulus or whether PEs trigger a window in which memory formation is (retrospectively) strengthened for all stimuli, including entirely uninformative ones. To this end, we used a modified paradigm, in which we presented in some of the encoding blocks entirely uninformative stimuli between the predictive stimulus and the outcome (i.e., potential PE). If the PE induces a transient window of enhanced encoding, then memory should also be enhanced for uninformative stimuli presented shortly after the predictive stimulus. However, if the PE affects specifically the encoding of the predictive stimulus, then memory should not be enhanced for the uninformative stimulus presented between the predictive stimulus and the PE. Moreover, because previous research (Kalbe & Schwabe, 2020, 2022b; and our Experiments 1 and 2) suggested that the PE effects might be at least partly independent of physiological arousal but included only SCR as the only arousal measure, we included here, in addition to SCR, respiratory responses and heart rate as further measures of arousal.

Method

Participants

Eighty-two healthy volunteers participated in this study (age: $M = 25.45$ years, $SD = 3.98$ years, range = 18–35 years). Exclusion criteria were identical to those in Experiments 1–3. Four participants were excluded from the analyses due to technical failure during the experiment or because they did not return for the second experimental day, resulting in a final sample of $n = 78$. Importantly, none of the participants had participated in Experiments 1–3.

The target sample size was based on an a priori power calculation using G*Power (3.1.9.6; Faul et al., 2009). Based on previous

research by Kalbe and Schwabe (2020), we assumed $d_z = .39$ as a point estimate for the expected effects. A two-tailed paired *t*-test with $\alpha = .05$ required at least 72 participants to achieve a statistical power of 0.90. In line with this calculation, a post hoc power simulation using the R-package simR (Green & MacLeod, 2016) for the observed effects of subsequent PEs and our final sample size of 78 participants yielded a power of 0.84 based on 1,000 simulations. All participants provided written informed consent before participation and received a moderate monetary reimbursement (up to 50 €) at the end of the study. The study was approved by the ethics committee of the Faculty of Psychology and Human Movement Science at the University of Hamburg and carried out in line with the Declaration of Helsinki.

Materials

We used the same stimulus set as in Experiment 3, that is, 810 pictures of vehicles, tools, and clothes, isolated on white background, but added 270 pictures of outdoor scenes, resulting in a total set of 1,080 stimuli. Outdoor scenes were taken from image databases, that is, Bank of Standardized Stimuli (Brodeur et al., 2010, 2014), McGill Calibrated Color Image Database (Olmos & Kingdom, 2004), SUN database (Xiao et al., 2010), Konklab (Konkle et al., 2010), and open-online sources. Again, 120 randomly drawn pictures from each of the categories vehicles, tools, and clothes were used during encoding on the first experimental day. In addition, 180 randomly drawn pictures of scenes were shown during these encoding sessions. From the remaining 540 pictures, 270 randomly chosen pictures, that is, 60 pictures of tools, 60 pictures of vehicles, 60 pictures of clothes, and 90 pictures of scenes, served as lures for the surprise recognition test on the second experimental day. All images were assumed to be emotionally neutral. The allocation of images to CS categories or lures as well as the order in which individual items were presented was randomized across participants and unique per participant.

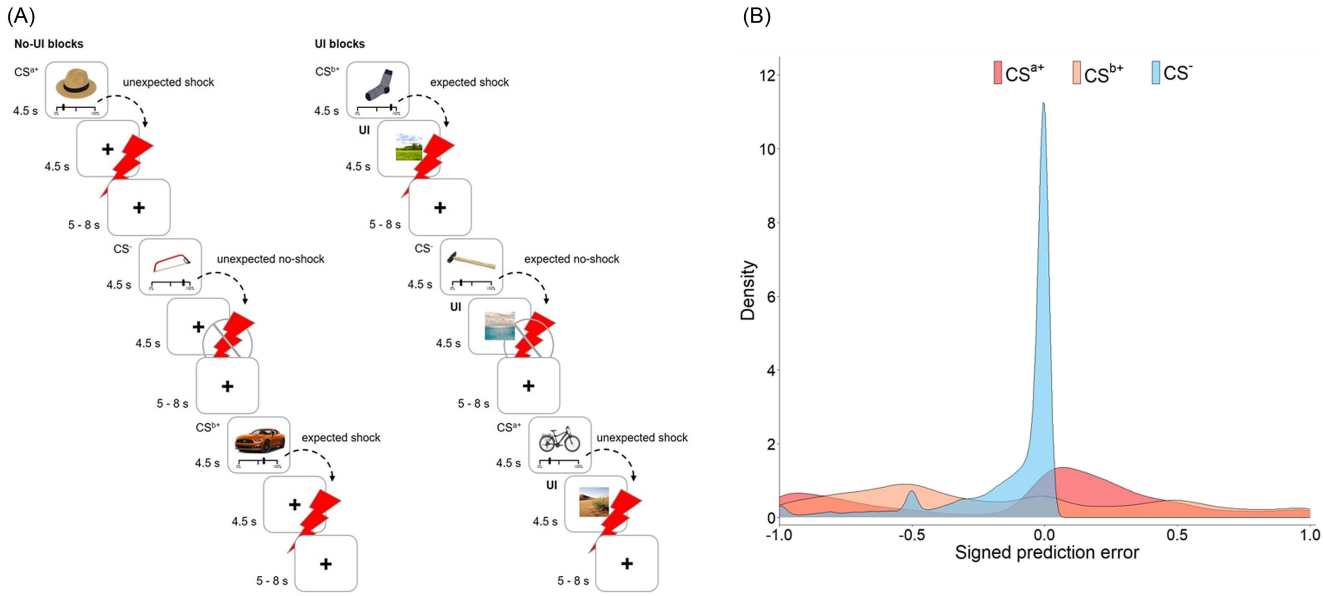
Procedure

The procedure was largely comparable to Experiment 3 but contained some important differences, in particular the inclusion of encoding blocks in which an uninformative stimulus (UI) was presented between CS and outcome (see Figure 4A). Furthermore, the CS-outcome interval was kept constant in this experiment. Same as in Experiments 1 and 2, this experiment consisted of two sessions, with an encoding session on the first experimental day and a recognition test on the following day, about 22–26 hr after encoding.

Upon arrival on the first experimental day, participants provided written informed consent and received written instructions that they were going to see a series of photographs of vehicles, tools, clothes, and scenes and that some of them might be followed by a brief electric shock. Again, they were then instructed to predict how likely a shock would be to follow a picture by adjusting a slider on the

Figure 4

Experimental Procedure of Experiment 4



Note. In the encoding task (A), participants saw a series of unique pictures from three different categories (clothes, vehicles, and tools) linked to fixed probabilities to receive an electric shock (CS^{a+} = 67%, CS^{b+} = 33%, and CS^{-} = 0%). On each trial, participants indicated their shock expectation on a continuous scale from 0% to 100%. Critically, in UI-blocks an UI-stimulus appeared between CS and outcome, whereas in no-UI blocks a black fixation cross was presented on the screen. (B) PEs were equally distributed around zero. Pictures taken and reprinted from "Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos," by M. B. Brodeur, K. Guérard, and M. Bouras, 2014, *PLOS ONE*, *9*(9), Article e106953 (https://doi.org/10.1371/journal.pone.0106953). CC BY 4.0 and from "SUN Database: Large-Scale Scene Recognition From Abbey to Zoo," by J. Xiao, J. Hays, K. A. Ehinger, A. Oliva, and A. Torralba, *2010 IEEE computer society conference on computer vision and pattern recognition* (pp. 3485–3492), 2010 (https://doi.org/10.1109/CVPR.2010.5539970). CC BY 4.0. CS = conditioned stimulus; PE = prediction errors; UI = uninformative stimulus. See the online article for the color version of this figure.

screen to a value corresponding to their prediction of the shock probability (ranging from 0% to 100%). Importantly, participants were neither told about the shock contingencies, nor that their memory would be tested later on. They were informed that their predictions would not affect the probability that a shock would occur, but that they should aim at improving their predictions over the task. In line with Experiment 3 and unbeknownst to the participants, the probabilities of a shock were linked to the image categories. One category served as CS^{a+} (67% shock probability), one as CS^{b+} (33% shock probability), and one as CS^{-} (0% shock probability). The assignment of image categories (i.e., vehicles, tools, and clothes) to the CS categories (i.e., CS^{a+}, CS^{b+}, and CS^{-}) was counterbalanced across participants and groups. We also added a new, stimulus category (UI), which was uninformative with respect to the occurrence of an electric shock. Scene images were always used as UI to make sure that these sufficiently distinct from the CS categories (i.e., CS^{a+}, CS^{b+}, and CS^{-}).

To measure SCRs and to apply the electric shocks, we used the same equipment as in Experiment 3 and followed an identical procedure. Before the learning task, stimulation intensity was adjusted individually to be unpleasant but not painful as described in Experiment 1.

The encoding session on the first experimental day consisted of four blocks with 90 trials each, with the critical difference to Experiment 3 that we used two different types of blocks: UI-blocks versus no UI-blocks (see Figure 4A). In blocks 1 and 3, referred to as *no-UI blocks*, participants saw pictures from the three CS categories only, whereas blocks 2 and 4 additionally contained UI stimuli between the CS and outcome (*UI-blocks*). We presented UI stimuli only in blocks 2 and 4 to rule out conditioning to the UI stimuli. More specifically, participants were assumed to learn the specific associations to the CS in the first encoding block (in which no UI stimuli were presented). The addition of the UI after the CS in the second block should not lead to conditioning to the UI stimulus, according to the classic blocking effect (Fanselow, 1998; Kamin, 1968, 1969). The third block, in which the CS was again presented without the UI stimulus, was further supposed to refresh the specific CS-outcome association, underlining that CS contingencies were independent of the UI. The inclusion of two blocks including UI stimuli between CS and outcome and two blocks not containing these UI stimuli, which were apart from the UI stimulus identical, allowed us further to directly assess the effect of the UI stimulus within Experiment 4 and to link the findings of Experiment 4 to those of the other three experiments.

In Blocks 1 and 3, 30 pictures of vehicles, 30 pictures of tools, and 30 pictures of clothes were presented in a pseudorandomized order. On each trial, a picture was shown in the center of the computer screen for 4.5 s, during which participants were asked to make their prediction about the probability of an electric shock. Therefore, a slider which could be adjusted to any integer value between 0% and 100% using the computer mouse was presented beneath each image. After stimulus offset, a black fixation cross was presented centrally for 4.5 s on the screen and which was immediately followed by the 200 ms outcome (shock vs. no-shock). Between trials, the fixation cross was presented on the screen for 6.5 ± 1.5 s, which again enabled us to measure the relatively slow (anticipatory) SCRs.

Blocks 2 and 4 differed from blocks 1 and 3 in the inclusion of UI stimuli. During the delay of 4.5 s between CS and outcome participants did not see a fixation cross, but a picture from the UI category. Importantly, the UI was presented centrally on the screen without a slider and had no influence on the CS-shock contingencies, leaving the UI completely uninformative for the shock predictions.

Again, there were short breaks (1–2 min) between blocks during which participants had the chance to recalibrate the shock intensity, if required. Each encoding block lasted approximately 25 min, resulting in a total duration of 100 min for the entire incidental encoding-fear learning session.

Data of respiratory frequency were collected continuously during the encoding session using a BioNomadix Respiratory Transducer (BIOPAC Systems, Goleta, California, United States) that was wrapped around the participants' upper torso approximately 5 cm below the arm pit at the point of maximum respiratory expansion and connected to the BioNomadix Respiratory Transmitter (BIOPAC Systems, Goleta, California, United States).

For the measurement of heart rate, we used a NIBP100D noninvasive blood pressure monitoring system (BIOPAC Systems, Goleta, California, United States) connected to the MP160 module. A blood-pressure cuff was placed on the participants' left arm and a double finger cuff sensor was placed on the index and middle finger of the left hand to measure heart rate continuously.

The surprise recognition test took place on the next day, 22–26 hr after encoding. Same as in Experiments 1–3, participants completed a short questionnaire to assess whether they anticipated a memory test and then rated how surprised they were about the recognition test on a scale from 1 (*not surprised at all*) to 5 (*very surprised*). In the recognition test, participants saw all 540 pictures they had seen during the encoding session (120 pictures of vehicles, 120 pictures of tools, 120 pictures of clothes, and 180 pictures of scenes) as well as 270 "new" pictures (60 pictures of vehicles, 60 pictures of tools, 60 pictures of clothes, and 90 pictures of scenes) that had not been presented before in a randomized order. Each trial started with a central white fixation cross on a white background for 1.5 ± 0.5 s, followed by an "old" or "new" picture presented centrally on the computer screen for 6 s. Again for each item, participants were asked to indicate whether the currently presented picture was *definitely old*, *maybe old*, *maybe new*, or *definitely new* by pressing the "1," "2," "3," or "4" button on the keyboard, respectively.

Data Analysis

Same as Experiment 3, the paradigm used in Experiment 4 enabled us to measure continuous PEs. Again, we calculated both types of PEs, that is, uPEs and sPEs. Our main analyses focus on sPEs. The detailed analyses of uPE effects are presented in the Supplemental Material.

Again, SCRs were analyzed using continuous decomposition analysis in Ledalab Version 3.4.9 (Benedek & Kaernbach, 2010) and derived in line with Experiments 1–3 (see Supplemental Material for detailed information).

Heart rate data were analyzed using the PsychoPhysiological Modelling Toolbox in MATLAB 4.2.1 (Bach et al., 2016). Detailed information is provided in the Supplemental Material.

For the analysis of respiration data, we also used the Psycho-Physiological Modelling Toolbox (Bach et al., 2016).

Again, we fitted GLMMs with a logit link function using the lme4 R package (Bates et al., 2015) and treated subjects as random effects for both the intercept and all slopes of the fixed effects included in the model (Barr et al., 2013). The recognition of an individual item

was treated as the binary dependent variable, coded "0" for misses and "1" for confident hits. We fitted models using different sets of independent variables including PEs, anticipatory as well as UI and outcome-related arousal, explicit shock prediction and block. In line with Experiments 1–3, we also distinguished between subsequent and previous PEs with the former referring to PE effects on the recognition of the preceding stimulus and the latter referring to PE effects on the recognition of the following stimulus.

Transparency and Openness

The materials, data, and R analysis scripts are publicly available on the Research Data Management System of University of Hamburg and can be accessed at https://www.fdr.uni-hamburg.de/record/14147 (Loock et al., 2024). This experiment was not preregistered.

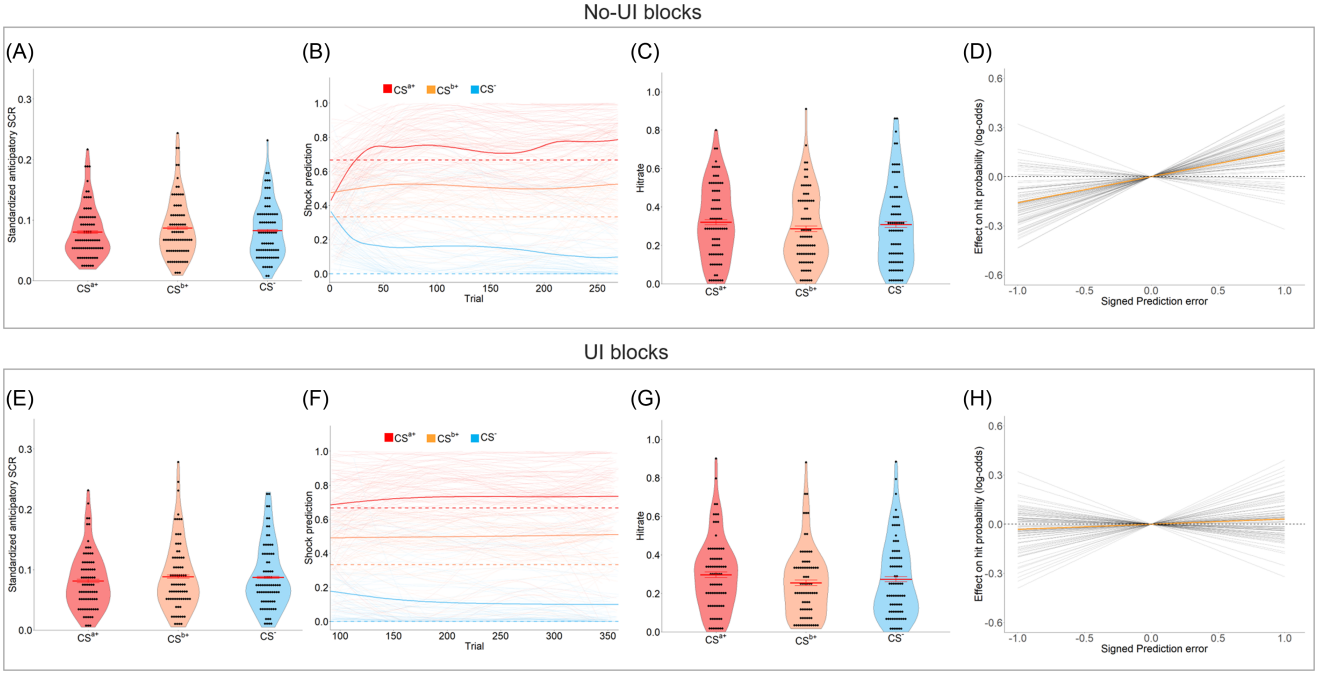
Results and Discussion

Successful Fear Conditioning

One participant had to be excluded from the SCR analysis due to technical failure, resulting in sample of $n = 77$ for this analysis. Again, SCR data confirmed the expected fear learning process. Specifically, anticipatory SCR differed significantly between CS categories, $F(2, 152) = 4.08$, $p = .019$, partial $\eta^2 = 0.004$. Post hoc paired t-tests showed that participants showed higher anticipatory SCRs to CS^{a+} items ($M = 0.09$, $SD = 0.02$) compared to CS^{-} items ($M = 0.08$, $SD = 0.02$), $t(76) = 2.32$, $p = .023$, $d = 0.15$ (Figure 5A and 5E). Anticipatory SCRs did not differ significantly between CS^{b+} items ($M = 0.08$, $SD = 0.01$) and CS^{-} items, $t(76) = 1.74$, $p = .085$, $d = 0.04$. Notably, the SCR to the different CS types did not differ between the UI und no-UI blocks, $F(2, 152) = 1.46$, $p = .234$, partial $\eta^2 = 0.003$. In addition, the SCR in response to the UI stimuli ($M = 0.08$, $SD = 0.04$) did not differ from the anticipatory SCR for the CS^{-} items, $t(76) = -0.76$, $p = .452$, $d = 0.06$. Outcome-related SCRs were significantly higher for shocked items ($M = 0.47$, $SD = 0.27$) compared to unshocked items ($M = 0.36$, $SD = 0.27$), $t(76) = 2.44$, $p = .017$, $d = 0.27$.

Explicit shock ratings further showed that participants learned the shock contingencies over the task very well. Overall, participants had a significantly higher shock expectancy for CS^{a+} ($M = 0.72$, $SD = 0.17$) compared to CS^{b+} ($M = 0.50$, $SD = 0.12$), $t(76) = 10.10$, $p < .001$, $d = 1.31$, and for CS^{b+} compared to CS^{-} ($M = 0.14$, $SD = 0.19$); $t(76) = 14.87$, $p < .001$, $d = 2.20$. Importantly, shock expectancies did not differ between UI and no-UI blocks, $F(1, 76) = 3.29$, $p = .074$, partial $\eta^2 = 0.00$ (see Figure 5B and 5F). In addition, PEs were equally distributed around zero (Figure 4B) suggesting a sufficient number of positive and negative PEs that could be analyzed.

Figure 5
Results of Experiment 4



Note. Mean standardized anticipatory SCR (A, E) and hit rates (C, G) confirmed successful fear conditioning, as reflected in significantly elevated SCR and increased hit rates of CS^{a+} compared with CS^{-} items in no-UI and UI blocks. Black dots show data from individual participants. Thick-red bar represents group mean, whereas thin-red bars show ±1 standard error of the mean. (B, F) Participants' mean shock expectancy ratings (thick lines) approached the true shock probabilities (dotted lines) relatively fast in no-UI and UI blocks. Subsequent sPEs boosted item memory only in no-UI blocks (G), whereas the effect was abolished in UI-blocks (H). SCR = skin conductance response; CS = conditioned stimulus; PE = prediction errors; UI = uninformative stimulus. See the online article for the color version of this figure.

General Memory Performance

On average, participants were moderately surprised by the recognition test (see Supplemental Material).

Overall, participants showed a lower recognition performance than in the other experiments ($M = 0.27$, $SD = 0.19$), which might be due to the higher number of encoded stimuli in total and the UI stimuli in particular, for which memory was poor (see below). The hit rate for CS items ($M = 0.29$, $SD = 0.19$) was significantly higher than the category-based false alarm rate ($M = 0.25$, $SD = 0.18$) for these items, $t(311) = 1.18$, $p = .024$, $d = 0.20$, thus demonstrating intact memory for the CS items.

Importantly, while false alarm rates did not differ between all four stimulus categories, $F(3, 231) = 0.59$, $p = .623$, partial $\eta^2 = 0.003$, hit rates differed significantly between categories, $F(3, 231) = 12.71$, $p < .001$, partial $\eta^2 = 0.05$ (see Figures 5C and 4G), suggesting that memory but not the response bias differed between categories. Post hoc paired t-tests revealed that the average hit rate for CS^{a+} items ($M = 0.31$, $SD = 0.12$) was significantly higher than for CS^{b+} items ($M = 0.27$, $SD = 0.12$), $t(77) = 2.07$, $p = .042$, $d = 0.21$, but did not differ significantly from CS^{-} items ($M = 0.29$, $SD = 0.11$); $t(77) = 1.23$, $p = .223$, $d = 0.11$. The average hit rate for CS^{b+} items did not differ from the hit rate for CS^{-} items, $t(77) = 0.87$, $p = .388$, $d = 0.09$. As expected, recognition memory performance was significantly lower for the UI items ($M = 0.20$, $SD = 0.12$) compared to all CS items, all: $t(77) > 3.97$, $p < .001$, $d > 0.43$, indicating that UI stimuli were considered irrelevant by participants and that memory was overall significantly enhanced for predictive stimuli.

This pattern of results was comparable when using d' (see Supplemental Material).

Moreover, recognition memory performance differed significantly between blocks with UI and without UI items. Recognition memory for CS items was significantly better in no-UI blocks (hit rate: $M = 0.30$, $SD = 0.07$) compared to UI blocks, hit rate: ($M = 0.27$, $SD = 0.07$), $t(155) = 4.95$, $p < .001$, $d = 0.17$, suggesting that the appearance of an UI stimulus affected memory formation for the predictive stimuli.

Modeling Recognition Performance at Item Level

To elucidate the mechanisms of episodic memory formation, we again fitted GLMMs with the recognition of an item as the binary-dependent variable and added the relevant independent predictors in a step-wise manner, in line with the previous experiments.

We started with a first minimal model in which we tested whether sPEs contribute to later recognition. Therefore, we treated the subsequent sPE (ranging from −1 to 1) following a CS as the sole independent variable to predict the binary recognition of this CS item, irrespective of the appearance of an UI stimulus. Estimates obtained revealed that (subsequent) sPEs ($z = 2.21$, $p = .027$, $\beta = 0.09$) showed a positive relationship with later memory. To rule out that the PE-effects were confounded with the shock prediction, we also computed a model where we added the explicit shock prediction as a predictor to the previous model. When controlling for the explicit shock prediction, the memory enhancing effect of the subsequent sPEs remained significant ($z = 2.52$, $p = .011$, $\beta = 0.12$).

In addition, we also set up a model that tested whether subsequent sPEs contribute to the recognition of UI stimuli treating subsequent sPEs as the sole independent variable to predict the binary recognition of an UI item. Crucially, estimates revealed that subsequent sPEs ($z = -0.66$, $p = .510$, $\beta = -0.04$) showed no significant relationship with later UI recognition, suggesting that the effect of subsequent sPEs is specific to the predictive stimulus and not found for UI stimuli presented between CS and outcome.

In a follow-up model, we added anticipatory arousal, UI-related arousal and outcome-related arousal as predictors to the minimal subsequent sPE-model for the binary recognition of a CS item. Anticipatory SCRs ($z = 0.63$, $p = .532$, $\beta = 0.09$), UI-related SCRs ($z = 0.72$, $p = .474$, $\beta = 0.10$) and outcome-related SCRs ($z = -0.03$, $p = .974$, $\beta = -0.00$) did not influence item recognition significantly. Even after controlling for arousal effects on memory, we still obtained a significant effect of subsequent sPE on CS memory ($z = 2.25$, $p = .024$, $\beta = 0.10$). To further elucidate whether heart rate and respiration amplitude as additional arousal metrics contribute to recognition memory, we also set up additional models separately for both predictors. Estimated showed that neither respiration amplitude ($z = 0.34$, $p = .732$, $\beta = 0.01$) nor heart rate ($z = 0.03$, $p = .977$, $\beta = 0.00$), predicted CS memory significantly. Following up on that, we added the subsequent sPE to each model as a predictor separately. When controlling for respiration amplitude, the subsequent sPE effect on memory of the predictive item remained significant ($z = 2.21$, $p = .027$, $\beta = 0.09$). A model including heart rate and subsequent sPE also yielded a significant subsequent sPE effect on item recognition ($z = 2.31$, $p = .021$, $\beta = 0.10$), when controlling for heart rate.

To examine whether sPE effects on memory interfere with the appearance of UI stimuli, we included block (no-UI block vs. UI block) as a predictor in an additional set of models. First, we set up a model that treated the subsequent sPE and block including their interaction as independent variables to predict the binary recognition of an item. We obtained a significant *Subsequent sPE* × *Block* interaction effect on memory ($z = -2.45$, $p = .014$, $\beta = -0.15$), suggesting that the "retrospective" PE effect on memory is influenced by the appearance of an UI stimulus. Accordingly, we set up separate models for blocks that contained UI items and for blocks that did not contain UI items. We treated the subsequent sPE as the sole independent variable to predict the binary recognition of an item. While this revealed that episodic memory was significantly increased by subsequent sPEs ($z = 2.90$, $p = .004$, $\beta = 0.16$) in no-UI blocks (see Figure 5D), we obtained a nonsignificant effect of subsequent sPEs on memory in UI blocks ($z = 0.63$, $p = .53$, $\beta = 0.13$; see Figure 5H), suggesting that those retrospective PE effects on memory disappear when an UI stimulus is presented between CS and outcome (i.e., PE). Even when controlling for anticipatory, UI- and outcome-related arousal, the pattern of results remained unchanged indicating a memory boost induced by subsequent sPE in no-UI blocks ($z = 2.87$, $p = .004$, $\beta = 0.16$), whereas there was no effect in UI blocks ($z = 0.58$, $p = .564$, $\beta = 0.03$).

Next, we performed additional models in which we treated the previous sPE as the sole independent variable to predict the binary recognition of the following item. This revealed a significant negative effect of sPEs on memory for items following the PE ($z = -2.05$, $p = .040$, $\beta = -0.07$), that is, previous sPEs appeared to be associated with a memory impairment. To rule out that the previous PE effects were confounded with the shock prediction, we also computed a model where we added the explicit shock prediction as a predictor to the previous model. When controlling for the explicit

shock prediction, the effect of the previous sPEs remained significant ($z = -2.20$, $p = .028$, $\beta = -0.07$).

Again, we added anticipatory, UI- and outcome-related SCRs to the former model to investigate confounds with physiological arousal. This revealed nonsignificant effects of anticipatory SCRs ($z = 0.73$, $p = .467$, $\beta = 0.10$), UI-related SCRs ($z = 0.82$, $p = .413$, $\beta = 0.12$), and outcome-related SCRs on memory formation ($z = 0.01$, $p = .992$, $\beta = 0.00$). The previous negative sPE effect on later memory remained significant ($z = -2.05$, $p = .041$, $\beta = -0.07$). In addition, we set up a model including respiration amplitude and previous sPE as variables to predict recognition of the following item. When controlling for respiration amplitude, the previous sPE effect on memory of the following item remained significant ($z = -1.99$, $p = .047$, $\beta = -0.07$). A model including heart rate and previous sPE yielded a trending previous sPE effect on recognition of the following item ($z = -1.76$, $p = .078$, $\beta = -0.06$).

In addition, we also set up a model that tested whether previous sPEs contribute to the recognition of UI stimuli treating sPEs as the sole independent variable to predict the binary recognition of the following UI item. Again, the model estimates revealed that previous sPEs ($z = -1.12$, $p = .261$, $\beta = -0.07$), showed no significant relationship with item recognition suggesting that the recognition of UI is independent of previous PEs.

In a next step, we tested whether the previously observed effect of (previous) PEs on the memory for items following the PE would be influenced by the appearance of an uninformative stimulus. Estimates obtained showed no significant *Previous sPE* × *Block* interaction effect on recognition of the following item ($z = -0.11$, $p = .915$, $\beta = -0.01$), suggesting that previous sPE effects on memory might be irrespective of the appearance of uninfomative stimuli. Even though the critical interaction effect was nonsignificant, in an explorative analysis, we set up separate models for blocks that contained UI items and for blocks that did not contain UI items. We treated the previous sPE as the sole independent variable to predict the binary recognition of the following item. Notably, this revealed that episodic memory was not influenced by previous sPEs in no-UI blocks ($z = -1.35$, $p = .177$, $\beta = -0.06$) nor in UI blocks ($z = -1.42$, $p = .155$, $\beta = -0.07$). When controlling for anticipatory arousal, UI-related, and outcome-related arousal, the previous sPE effect on memory of the following item remained nonsignificant in no-UI blocks ($z = -1.36$, $p = .174$, $\beta = -0.07$) and UI blocks ($z = -1.42$, $p = .155$, $\beta = -0.07$).

For models investigating whether subsequent PEs and previous PEs reflect distinct mechanisms, see Supplemental Material.

Effects of uPEs

We found no significant effects of subsequent or previous uPEs on episodic memory. For a detailed analysis of the uPE effects in Experiment 4, see Supplemental Material.

Conclusion

The findings of Experiment 4 replicate again the enhancing effects of (signed) PEs associated with aversive events on memory for the stimulus encountered before the PE. Critically, the findings of Experiment 4 show that the enhancing effect of (subsequent) PEs on memory is specific to predictive stimuli and not observed for UI stimuli presented between CS and outcome. Moreover, our results reveal that the presence of an UI stimulus between predictive stimulus and outcome (PE) can even abolish the PE-driven memory enhancement. Interestingly and in contrast to Experiments 1 and 2, we found that prospective (signed) PEs were associated with a memory impairment that was unaffected by the appearance of uninformative information.

General Discussion

Accumulating evidence shows that PEs related to aversive or rewarding events may enhance memory for preceding neutral events (Kalbe & Schwabe, 2020, 2022b; Rouhani et al., 2023). However, the cognitive underpinnings of such PE-induced memory enhancements are not well understood. In the present series of experiments, we aimed to elucidate the cognitive mechanisms underlying aversive PE-driven memory enhancements for inherently neutral events. Our results demonstrate that the memory-enhancing effect of PEs on stimuli preceding the PE, known as the retrospective PE effect, persists for at least 10 s. Importantly, this effect is evident immediately after encoding, suggesting that it does not necessitate a consolidation process. While we obtained evidence indicating that the PE-induced memory boost extends to stimuli presented after the PE (prospective PE effects), our results emphasize the specificity of this memory enhancement. It is exclusively linked to predictive stimuli, with uninformative stimuli even shown to interfere with the PE-driven memory enhancement.

Across all four experiments, our results replicate the previously reported beneficial PE effects on memory for preceding neutral stimuli. Consistent with previous research (Kalbe & Schwabe, 2020), Experiments 1 and 2 show memory-enhancing effects of uPEs, suggesting that the PE magnitude affected memory formation. In Experiments 3 and 4, we assessed PEs on a continuous scale, allowing us to differentiate between positive and negative sPEs which have been shown to exert differential effects on memory formation (Kalbe & Schwabe, 2022b; Rouhani & Niv, 2021). Our results of Experiments 3 and 4 show consistently that negative PEs were associated with impairing effects on subsequent memory, whereas positive PEs were linked to enhanced memory. This pattern of results is in stark contrast to our previous findings (Kalbe & Schwabe, 2022b), which showed the exact opposite pattern. However, a notable distinction between these studies is the testing environment and the number of experimental trials. Our previous study took place in a magnetic resonance imaging scanner, which may have resulted in higher state anxiety levels known to modulate PE processing (Hein & Herrojo Ruiz, 2022). Indeed, participants that volunteer for magnetic resonance imaging studies have been shown to be characterized by reduced trait anxiety levels compared to participants in behavioral experiments (Charpentier et al., 2021). Trait anxiety is typically correlated with depressive mood known to affect PE effects on memory formation (Rouhani & Niv, 2019). However, the potential modulation of PE effects on memory by trait anxiety remains speculative and needs to be tested explicitly in future studies. In addition, we increased the trial number significantly (360 trials vs. 120 trials in our previous study), underlining the high validity of our empirical findings. Because of the increased trial number, participants also received substantially more electric shocks, potentially resulting in higher sensitivity to the aversive outcome (Chen et al., 2000; Lonsdorf et al., 2017), which may have rendered the shock experience even more aversive and hence positive PEs more intense.

Previous research has demonstrated that reward-related PEs act as event boundaries, disrupting the sequential encoding of events into memory (Rouhani & Niv, 2019). Despite this evidence, the cognitive mechanisms underlying PE-driven memory enhancements have remained elusive. Here, we present initial evidence for a prospective effect of PEs on subsequent memory: in three out of four experiments, uPEs enhanced not only the memory for events preceding the PE but also for events following the PE. These prospective PE effects appeared generally weaker and less consistent than the retrospective PE effects. While we observed no significant effect of sPEs on memory for events following the PE in Experiment 3, there was even a memory impairment associated with prospective sPEs in Experiment 4. Even though more research is needed to elucidate the nature of potential prospective PE effects on memory for stimuli following the PE event, the repeatedly obtained prospective PE effects on memory may provide valuable insights by suggesting PE effects on stimuli not directly linked to the PE event, pointing to a potential nonspecific window of enhanced memory formation induced by a PE. Our findings from Experiment 3 demonstrate that this window of enhanced memory formation lasts for at least 10 s, significantly longer than previously considered in studies (see Kalbe & Schwabe, 2020, 2022b). While the present findings suggest that the PE may capture events encoded 10 s before, it remains unclear for how long exactly the PE effects last.

Although the observed prospective effects of PEs on subsequent memory may be considered as evidence for the idea that PEs open a transient window of enhanced memory formation for all stimuli, irrespective of their relevance for the PE event, the findings of Experiment 4 contradict such an unselective mechanism. Specifically, Experiment 4 demonstrates no memory enhancement for uninformative stimuli presented between the predictive stimulus and PE. How can we reconcile the absence of memory enhancement for these uninformative stimuli with the prospective PE effects, which imply memory enhancement for subsequent stimuli not informative for the current PE event? The answer to this question may relate to the predictive value of the stimuli per se. Participants presumably quickly learned the irrelevance of the uninformative stimuli, resulting in shallow processing as also reflected in the overall low-memory performance for uninformative stimuli. In contrast to these uninformative stimuli, the CS following the PE does carry informative value, namely for the subsequent PE event. Thus, PEs enhanced memory for predictive stimuli around the time of the PE event but not for entirely unpredictive stimuli, refuting the idea of a PE-induced window of unselective memory enhancement.

Intriguingly, the presentation of an uninformative stimulus between the CS and outcome (i.e., PE) even abolished the PE-induced memory enhancement for the predictive CS. This finding is remarkable, suggesting that uninformative information interferes with the association between the predictive stimulus and PE. Specifically, it highlights the necessary active maintenance of the predictive stimulus until the PE, which the UI stimulus interfered with. The predictive stimulus may be stored in working memory (Baddeley, 1992; Oberauer et al., 2003) and thus be highly vulnerable to competing stimuli appearing during the maintenance phase. If the predictive stimulus needs to be maintained in working memory, known to be sensitive to interference and delay (Dosher, 1999; Gresch et al., 2021), it is surprising that the PE-induced memory enhancement was not sensitive to the delay between the CS and outcome in Experiment 3. A potential explanation for the lack of a delay-related decrease in the PE effect may relate to the fact that only one item (the CS) needed to be maintained, implicating a rather limited working memory load.

While the PE-driven memory enhancement overlaps to some extent with the well-established arousal-induced enhancement of memory (Rouhani et al., 2023), previous research suggested that PE effects go beyond these of arousal (Kalbe & Schwabe, 2020, 2022b). In line with the idea that PE effects are, at least partly, independent of arousal effects, we show here that the effects of subsequent PE on memory remain even when controlling for measures of arousal. This pattern was confirmed in all of the experiments, except in Experiment 3, in which the PE effect was only a nonsignificant trend after including arousal parameters. The latter might suggest that the influence of arousal becomes more prominent, potentially superimposing PE effects, as the CS-outcome delay increases, related to the involvement of working memory processes as discussed above. Because previous research used SCR as the only measure of arousal, we added heart rate and respiration amplitude as additional arousal measures in Experiment 4, to further disentangle arousal- and PE-related effects on memory. When controlling for these measures, the memory-enhancing effect of subsequent PEs remained. These findings suggest that physiological arousal (beyond SCR) cannot account for the memory boost alone indicating that retrospective PE-induced memory enhancements presumably go at least partly beyond the mere effects of arousal on memory formation. Notably, when considering the association of prospective effects of PEs and arousal on memory, our results remain more inconclusive. While we show that the effects of previous PEs on memory of the following item remain when controlling for arousal in Experiments 2 and 4, this pattern was not significant in Experiments 1 and 3. Given that the prospective effects of PEs on memory cover a time window that is relatively long compared to the retrospective PE effect because it even spans to the next trial, we assume that the effect of arousal becomes more pronounced as the association between the CS and the outcome which has to be maintained mitigates over time. Thus, prospective PE effects appear to be more strongly driven by arousal effects.

In sum, our findings provide insights into the cognitive mechanisms involved in aversive PE-driven enhancements of memory for surrounding neutral events. Specifically, our data show that these PE effects persist for a longer duration than previously thought and extend to stimuli encoded after the PE event. Nevertheless, these PE effects are not unspecific, as reflected in the absence of any memory boost for uninformative stimuli presented between CS and PE. Rather, PEs appear to enhance memory for predictive stimuli encountered around a PE event. Importantly, these PE effects are sensitive to interference, pointing to an involvement of working memory maintenance. Moreover, the memory-enhancing effects of PEs are already present shortly after encoding, suggesting that consolidation processes are not required for these aversive PE effects on memory to unfold. Our findings provide insights into the fundamental question of which of the many stimuli that we are continuously presented with are stored in long-term memory: those that bear predictive value for unexpected emotional events. Beyond their theoretical relevance, these findings may have relevant implications, for instance, for eyewitness testimony or for clinical contexts, in which the effectiveness of therapeutic interventions is known to rely on expectancy violation (i.e., PEs; Pittig et al., 2023).

Constraints on Generality

Participants of each all four experiments consisted of healthy, young adults ranging between 18 and 35 years of age, recruited from a community sample. Thus, we believe that our results are reproducible with people from the general population. However, we do not have evidence that the findings will occur outside laboratory settings, in an older population that might suffer from memory impairments or in a population that suffers from mental disorders. We have no reason to believe that the results depend on other characteristics of the participants, materials, or context.

Transparency and Openness

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. All data, materials, and scripts have been made publicly available via the Research Data Repository of the University of Hamburg and can be accessed at (https://www.fdr.uni-hamburg.de/record/14147; Loock et al., 2024). Data were analyzed using R, Version 4.0.0 (R Core Team, 2021). The study design and its analysis were not preregistered.

References

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.e7. https://doi.org/10.1016/j.neuron.2020.10.029

Bach, D. R., Gerster, S., Tzovara, A., & Castegnetti, G. (2016). A linear model for event-related respiration responses. *Journal of Neuroscience Methods*, *270*, 147–155. https://doi.org/10.1016/j.jneumeth.2016.06.001

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280–289. https://doi.org/10.1016/j.tics.2007.05.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bartlett, J. C., Till, R. E., & Levy, J. C. (1980). Retrieval characteristics of complex pictures: Effects of verbal encoding. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 430–449. https://doi.org/10.1016/S0022-5371(80)90303-5

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

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

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Bergt, A., Urai, A. E., Donner, T. H., & Schwabe, L. (2018). Reading memory formation from the eyes. *European Journal of Neuroscience*, *47*(12), 1525–1533. https://doi.org/10.1111/ejn.13984

Boucsein, W. (1992). *Electrodermal activity*. Springer. https://doi.org/10.1007/978-1-4757-5093-5

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLOS ONE*, *5*(5), Article e10773. https://doi.org/10.1371/journal.pone.0010773

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PLOS ONE*, *9*(9), Article e106953. https://doi.org/10.1371/journal.pone.0106953

Buchanan, T. W. (2007). Retrieval of emotional memories. *Psychological Bulletin*, *133*(5), 761–779. https://doi.org/10.1037/0033-2909.133.5.761

Cahill, L., & McGaugh, J. L. (1996). Modulation of memory storage. *Current Opinion in Neurobiology*, *6*(2), 237–242. https://doi.org/10.1016/S0959-4388(96)80078-X

Cahill, L., & McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neurosciences*, *21*(7), 294–299. https://doi.org/10.1016/S0166-2236(97)01214-9

Charpentier, C. J., Faulkner, P., Pool, E. R., Ly, V., Tollenaar, M. S., Kluen, L. M., Fransen, A., Yamamori, Y., Lally, N., Mkrtchian, A., Valton, V., Huys, Q. J. M., Sarigiannidis, I., Morrow, K. A., Krenz, V., Kalbe, F., Cremer, A., Zerbes, G., Kausche, F. M., … O'Doherty, J. P. (2021). How representative are neuroimaging samples? Large-scale evidence for trait anxiety differences between fMRI and behaviour-only research participants. *Social Cognitive and Affective Neuroscience*, *16*(10), 1057–1070. https://doi.org/10.1093/scan/nsab057

Chen, D.-Y., Ho, S.-H., & Liang, K. C. (2000). Startle responses to electric shocks: Measurement of shock sensitivity and effects of morphine, buspirone and brain lesions. *The Chinese Journal of Physiology*, *43*(1), 35–47.

Christianson, S.-A. (2014). *The handbook of emotion and memory: Research and theory*. Psychology Press. https://doi.org/10.4324/9781315807454

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

Clewett, D., & McClay, M. (2024). Emotional arousal lingers in time to bind discrete episodes in memory. *Cognition and Emotion*. Advance online publication. https://doi.org/10.1080/02699931.2023.2295853

Clewett, D. V., Huang, R., Velasco, R., Lee, T.-H., & Mather, M. (2018). Locus Coeruleus activity strengthens prioritized memories under arousal. *The Journal of Neuroscience*, *38*(6), 1558–1574. https://doi.org/10.1523/JNEUROSCI.2097-17.2017

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language*, *59*(4), 447–456. https://doi.org/10.1016/j.jml.2007.11.004

Dosher, B. A. (1999). Item interference and time delays in working memory: Immediate serial recall. *International Journal of Psychology*, *34*(5–6), 276–284. https://doi.org/10.1080/002075999399576

Dunsmoor, J. E., Murty, V. P., Davachi, L., & Phelps, E. A. (2015). Emotional learning selectively and retroactively strengthens memories for related events. *Nature*, *520*(7547), 345–348. https://doi.org/10.1038/nature14106

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*(1), 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Ergo, K., De Loof, E., & Verguts, T. (2020). Reward prediction error and declarative memory. *Trends in Cognitive Sciences*, *24*(5), 388–397. https://doi.org/10.1016/j.tics.2020.02.009

Fanselow, M. S. (1998). Pavlovian conditioning, negative feedback, and blocking: Mechanisms that regulate association formation. *Neuron*, *20*(4), 625–627. https://doi.org/10.1016/S0896-6273(00)81002-8

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

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

Gresch, D., Boettcher, S. E. P., van Ede, F., & Nobre, A. C. (2021). Shielding working-memory representations from temporally predictable external interference. *Cognition*, *217*, Article 104915. https://doi.org/10.1016/j.cognition.2021.104915

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

Hein, T. P., & Herrojo Ruiz, M. (2022). State anxiety alters the neural oscillatory correlates of predictions and prediction errors during reward-based learning. *NeuroImage*, *249*, Article 118895. https://doi.org/10.1016/j.neuroimage.2022.118895

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Jang, A. I., Nassar, M. R., Dillon, D. G., & Frank, M. J. (2019). Positive reward prediction errors during decision-making strengthen memory encoding. *Nature Human Behaviour*, *3*(7), 719–732. https://doi.org/10.1038/s41562-019-0597-3

Kalbe, F., & Schwabe, L. (2020). Beyond arousal: Prediction error related to aversive events promotes episodic memory formation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(2), 234–246. https://doi.org/10.1037/xlm0000728

Kalbe, F., & Schwabe, L. (2022a). On the search for a selective and retroactive strengthening of memory: Is there evidence for category-specific behavioral tagging? *Journal of Experimental Psychology: General*, *151*(1), 263–284. https://doi.org/10.1037/xge0001075

Kalbe, F., & Schwabe, L. (2022b). Prediction errors for aversive events shape long-term memory formation through a distinct neural mechanism. *Cerebral Cortex*, *32*(14), 3081–3097. https://doi.org/10.1093/cercor/bhab402

Kamin, L. J. (1968). Attention-like processes in classical conditioning. *Symposium on aversive motivation Miami* (Bd. No. TR-5) (pp. 9–31). University of Miami Press.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296). Appleton-Century-Crofts.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, *139*(3), 558–578. https://doi.org/10.1037/a0019165

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience*, *7*(1), 54–64. https://doi.org/10.1038/nrn1825

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., Heitland, I., Hermann, A., Kuhn, M., Kruse, O., Meir Drexler, S., Meulders, A., Nees, F., Pittig, A., Richter, J., Römer, S., Shiban, Y., Schmitz, A., Straube, B., … Merz, C. J. (2017). Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience and Biobehavioral Reviews*, *77*, 247–285. https://doi.org/10.1016/j.neubiorev.2017.02.026

Loock, K., Kalbe, F., & Schwabe, L. (2024). *Cognitive mechanisms of aversive prediction error-induced memory enhancements* (Version 2) [Data set]. https://www.fdr.uni-hamburg.de/record/14147

Mather, M., Clewett, D., Sakaki, M., & Harley, C. W. (2016). Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behavioral and Brain Sciences*, *39*, Article e200. https://doi.org/10.1017/S0140525X15000667

McGaugh, J. L. (2018). Emotional arousal regulation of memory consolidation. *Current Opinion in Behavioral Sciences*, *19*, 55–60. https://doi.org/10.1016/j.cobeha.2017.10.003

Oberauer, K., Süβ, H.-M., Wilhelm, O., & Wittman, W. W. (2003). The multiple faces of working memory. *Intelligence*, *31*(2), 167–193. https://doi.org/10.1016/S0160-2896(02)00115-0

Olmos, A., & Kingdom, F. A. A. (2004). A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception*, *33*(12), 1463–1473. https://doi.org/10.1068/p5321

Phelps, E. A. (2004). Human emotion and memory: Interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology*, *14*(2), 198–202. https://doi.org/10.1016/j.conb.2004.03.015

Pine, A., Sadeh, N., Ben-Yakov, A., Dudai, Y., & Mendelsohn, A. (2018). Knowledge acquisition is governed by striatal prediction errors. *Nature Communications*, *9*(1), Article 1673. https://doi.org/10.1038/s41467-018-03992-5

Pittig, A., Heinig, I., Goerigk, S., Richter, J., Hollandt, M., Lueken, U., Pauli, P., Deckert, J., Kircher, T., Straube, B., Neudeck, P., Koelkebeck, K., Dannlowski, U., Arolt, V., Fydrich, T., Fehm, L., Ströhle, A., Totzeck, C., Margraf, J., … Wittchen, H.-U. (2023). Change of threat expectancy as mechanism of exposure-based psychotherapy for anxiety disorders: Evidence from 8,484 exposure exercises of 605 patients. *Clinical Psychological Science*, *11*(2), 199–217. https://doi.org/10.1177/21677026221101379

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

Reisberg, D., & Hertel, P. (2005). *Memory and emotion*. Oxford University Press.

Rouhani, N., & Niv, Y. (2019). Depressive symptoms bias the prediction-error enhancement of memory towards negative events in reinforcement learning. *Psychopharmacology*, *236*(8), 2425–2435. https://doi.org/10.1007/s00213-019-05322-z

Rouhani, N., & Niv, Y. (2021). Signed and unsigned reward prediction errors dynamically enhance learning and memory. *eLife*, *10*, Article e61077. https://doi.org/10.7554/eLife.61077

Rouhani, N., Niv, Y., Frank, M. J., & Schwabe, L. (2023). Multiple routes to enhanced memory for emotionally relevant events. *Trends in Cognitive Sciences*, *27*(9), 867–882. https://doi.org/10.1016/j.tics.2023.06.006

Rouhani, N., Norman, K. A., & Niv, Y. (2018). Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(9), 1430–1443. https://doi.org/10.1037/xlm0000518

Rouhani, N., Norman, K. A., Niv, Y., & Bornstein, A. M. (2020). Reward prediction errors create event boundaries in memory. *Cognition*, *203*, Article 104269. https://doi.org/10.1016/j.cognition.2020.104269

Schlüter, H., Hackländer, R. P., & Bermeitinger, C. (2019). Emotional oddball: A review on memory effects. *Psychonomic Bulletin & Review*, *26*(5), 1472–1502. https://doi.org/10.3758/s13423-019-01658-x

Sinclair, A. H., & Barense, M. D. (2019). Prediction error and memory reactivation: How incomplete reminders drive reconsolidation. *Trends in Neurosciences*, *42*(10), 727–739. https://doi.org/10.1016/j.tins.2019.08.007

Trapp, S., O'Doherty, J. P., & Schwabe, L. (2018). Stressful events as teaching signals for the brain. *Trends in Cognitive Sciences*, *22*(6), 475–478. https://doi.org/10.1016/j.tics.2018.03.007

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). SUN database: Large-scale scene recognition from abbey to zoo. *2010 IEEE computer society conference on computer vision and pattern recognition* (pp. 3485–3492). https://doi.org/10.1109/CVPR.2010.5539970

Received March 1, 2024
Revision received October 15, 2024
Accepted October 30, 2024 ■