Neuroscience and Biobehavioral Reviews 153 (2023) 105371

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

# Signed and unsigned effects of prediction error on memory: Is it a matter of choice?

Francesco Pupillo [a,b,*], Rasmus Bruckner [c,d]

[a] Department of Psychology, Goethe-Universität Frankfurt, Germany
[b] Tilburg School of Social and Behavioral Sciences, Tilburg University, Netherlands
[c] Department of Education and Psychology, Freie Universität Berlin, Germany
[d] Max Planck Research Group NeuroCode, Max Planck Institute for Human Development, Berlin, Germany

ARTICLE INFO

*Keywords:*
Prediction error
Episodic memory
Computational models
Choice-confirmation bias
Positivity bias

ABSTRACT

Adaptive decision-making is governed by at least two types of memory processes. On the one hand, learned predictions through integrating multiple experiences, and on the other hand, one-shot episodic memories. These two processes interact, and predictions – particularly prediction errors – influence how episodic memories are encoded. However, studies using computational models disagree on the exact shape of this relationship, with some findings showing an effect of signed prediction errors and others showing an effect of unsigned prediction errors on episodic memory. We argue that the choice-confirmation bias, which reflects stronger learning from choice-confirming compared to disconfirming outcomes, could explain these seemingly diverging results. Our perspective implies that the influence of prediction errors on episodic encoding critically depends on whether people can freely choose between options (i.e., instrumental learning tasks) or not (Pavlovian learning tasks). The choice-confirmation bias on memory encoding might have evolved to prioritize memory representations that optimize reward-guided decision-making. We conclude by discussing open issues and implications for future studies.

## 1. Introduction

To flexibly adjust to the environment, humans have the ability to both incrementally learn from several past episodes (e.g., the average quality of a restaurant in the neighbourhood) and to encode temporally specific, detailed representations of single episodes (e.g., remembering a previous dinner with friends). These two elementary cognitive functions are thought to rely on partly dissociable brain circuits. Whereas incremental learning depends on dopaminergic activity originating in the midbrain and targeting the striatum (Daw, 2011; McClure et al., 2003; Schultz et al., 1997), episodic memory relies on the structures of the medial temporal lobe, especially the hippocampus (Eichenbaum et al., 2007). Functional interactions and anatomical connections between these neural systems (e.g., Foerde and Shohamy, 2011; Lisman and Grace, 2005; Shohamy and Turk-Browne, 2013) could suggest that incremental learning modulates episodic memory encoding, potentially via dopaminergic signals (Jang et al., 2019).

Consistent with this idea, the predictive processing framework suggests that new episodic memories should mainly be created when predictions formed through learning are violated (Henson and Gagnepain, 2010). Accordingly, the brain tries to constantly predict incoming information and minimize surprise (Clark, 2013; Friston, 2010). This process is hierarchically organized such that higher-level areas attempt to predict lower-level activity. Information not in line with these predictions constitutes a prediction error (PE), which is passed to the higher levels and used to update beliefs for optimizing future predictions. Creating a temporally specific, perceptually detailed episodic representation of our ongoing experiences is considered higher-level activity and should mainly occur in response to larger PEs, indicating that previous predictions could not fully explain the ongoing experience (Henson and Gagnepain, 2010). Therefore, when events that are not in line with our expectations occur, the PE causes increased learning and better episodic encoding so that predictions will be more precise in the future. Consequently, one would expect that the likelihood of encoding episodic representations should increase as a function of PEs.

The reinforcement learning (RL) framework (Sutton and Barto, 2018) provides one way to formalize PE-driven learning. RL offers mechanistic computational models that describe how expectations are

* Correspondence to: Tilburg School of Social and Behavioral Sciences, TIlburg University, 5037 DB Tilburg, Netherlands.
*E-mail address:* F.Pupillo@tilburguniversity.edu (F. Pupillo).

https://doi.org/10.1016/j.neubiorev.2023.105371
Received 27 April 2023; Received in revised form 18 August 2023; Accepted 23 August 2023
Available online 24 August 2023

updated in response to the PE (see Box 1). These models have frequently been used in the learning literature and show a marked correspondence with behavioural and neural data (Daw and Tobler, 2014). Moreover, RL algorithms have been used in the majority of studies on the relationship between PE and episodic memory (Calderon et al., 2021; Davidow et al., 2016; De Loof et al., 2018; Pupillo et al., 2023; Rouhani and Niv, 2021; Rouhani et al., 2018; Wimmer et al., 2014), and we therefore also focus on an RL approach in this work.

The studies linking incremental learning and episodic memory produced contrasting findings regarding the relationship between PE and episodic encoding. A major critical open question concerns the sign of the effect of PEs (Ergo et al., 2020). While several studies showed a positive relationship between signed PEs (i.e., ranging from negative to positive) and episodic encoding (Calderon et al., 2021; Davidow et al., 2016; De Loof et al., 2018; Jang et al., 2019; Pupillo et al., 2023), other studies reported a positive effect of unsigned PEs (i.e., absolute PEs; Rouhani and Niv, 2019, 2021; Rouhani et al., 2018). A positive relationship between signed PEs and memory suggests that episodic encoding is better for better-than-expected outcomes compared to worse-than-expected outcomes. On the contrary, a positive effect of unsigned PEs implies that unexpected, surprising outcomes improve episodic encoding per se, regardless of valence. In this opinion paper, we focus on these apparently divergent findings and suggest that the effects of signed PEs on episodic memory observed when outcomes are delivered are essentially driven by a choice-confirmation bias often arising in instrumental learning tasks (e.g., Palminteri and Lebreton, 2022).

## 2. What we know: the effect of PE on memory

The interaction between PEs and episodic memory is typically studied in experiments that feature a learning task in which PEs are experienced and events are encoded, followed by a recognition memory phase where participants are asked to recognize previously encoded items among new items. In particular, previous studies employed learning tasks in which participants learned the values of different stimuli and actions and were presented with outcomes that matched or violated their expectations, generating PEs of various degrees (Fig. 2). These studies have often derived PEs from RL algorithms and linked them to recognition-memory performance (Fig. 1). Therefore, this approach allows researchers to test whether the PE experienced while encoding an object facilitates or hinders subsequent object recognition in the memory test.

Studies taking this approach generally found a positive relation between PE and memory (Calderon et al., 2021; Davidow et al., 2016; De Loof et al., 2018; Jang et al., 2019; Pupillo et al., 2023; Rosenbaum et al., 2022; Rouhani and Niv, 2019, 2021; Rouhani et al., 2018). This PE-driven enhancement of recognition memory does not depend on specific characteristics of the task, such as whether the values of the stimuli were learned before or during the encoding task, whether the outcome was monetary or not, and whether the recognition test was intentional or incidental. Also, the retention interval does not seem to affect the PE-memory relationship, as the effects were observed for both immediate and one-day delayed recognition tests. All in all, these findings suggest that the effects of PEs on memory are quite robust.

One potential case in which PEs might be associated with worse recognition memory is when the timing between the presentation of objects and PEs is delayed. The study by Wimmer et al. (2014) found that PEs were detrimental to the encoding of objects (see Fig. 2h-i). Importantly, the major difference to the aforementioned studies was the timing of the presentation of the feedback. In fact, while in most of the studies, the to-be-remembered objects were presented together with the corresponding PEs, Wimmer and colleagues presented monetary feedback reflecting PEs one second after the objects. This delay might have led to a working memory conflict between prioritizing episodic memory or incremental learning representations. In line with this idea, when participants encoded the objects better (higher episodic memory performance), reward had a decreased influence on the next choice (lower incremental learning performance).

## 3. What we do not know: signed vs. unsigned PE

Although there is mounting evidence that PEs positively affect episodic memory, the findings diverge when it comes to the sign of the PE. On the one hand, a series of studies found an effect of signed PEs on recognition memory (Calderon et al., 2021; Davidow et al., 2016; De Loof et al., 2018; Jang et al., 2019; Pupillo et al., 2023; Rouhani and Niv, 2021). That is, these studies suggest that objects related to better-than-expected outcomes (positive PE) are remembered better than objects related to worse-than-expected outcomes (negative PE; Fig. 2b,g). On the other hand, other studies have found an effect of unsigned PEs on memory, indicating that objects presented together with surprising outcomes are remembered better, regardless of their valence (Fig. 2d,e; Rosenbaum et al., 2022; Rouhani and Niv, 2019, 2021; Rouhani et al., 2018).

**Box 1**
- Reinforcement Learning Models.

RL models can be used to formalize how expectations gradually change in response to feedback. For example, when visiting a new restaurant for dinner and you adjust your expectation about the quality of the food across multiple visits. In the canonical RL model, an expectation is quantified by the expected value $V_t^s$ for a stimulus $s$ on a given trial $t$, which is updated as follows after a reward has been received:

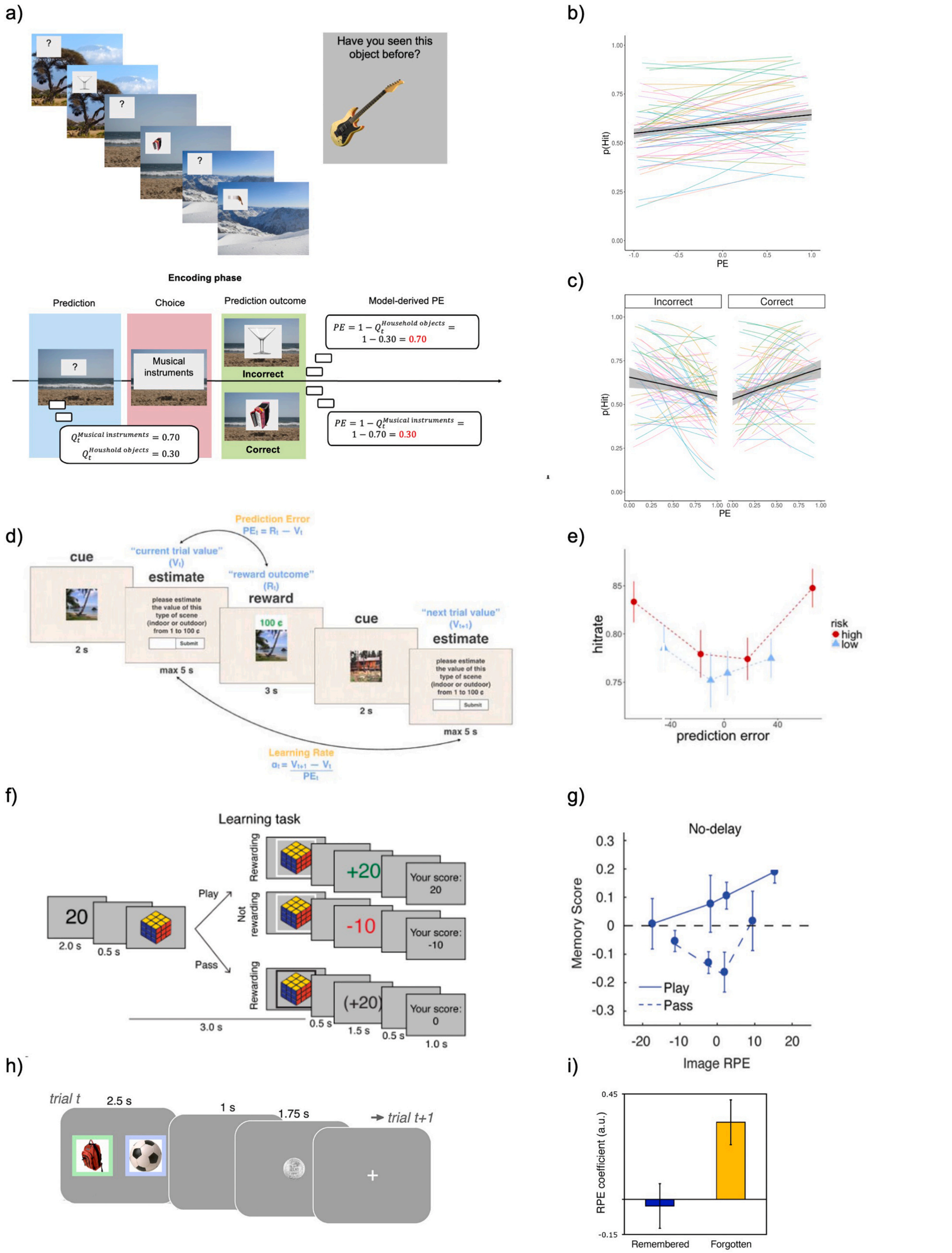
$$V_{t+1}^s = V_t^s + \alpha \delta_t^s, \quad (1)$$

where the learning rate $\alpha$ is a parameter ranging between zero and one and indicates the extent to which the PE $\delta_t^s$ is taken into account to update the expected value $V_{t+1}^s$. A larger $\alpha$ assigns a higher weight to the most recent PE when updating the expected value. The PE $\delta_t^s$ is computed as:

$$\delta_t^s = r_t - V_t^s, \quad (2)$$

where $r_t$ represents the reward on trial $t$ (e.g., quality of a dish). The PE $\delta$ has a positive sign if the reward $r_t$ is larger than $V_t^s$ (e.g., the food is more delicious than expected), and vice versa if the reward is lower than expected. By contrast, taking the unsigned PE, i.e., its absolute value, has traditionally been interpreted as to how surprising an experienced outcome is, regardless of whether it is positive or negative (e.g., Pearce and Hall, 1980). The learning rate $\alpha$ is a free parameter, obtained by fitting the RL model to participants' data (e.g., value estimations or BOLD activity, Wilson and Collins, 2019).

To examine the relation between PE and memory, the PE elicited during the encoding of an object can then be linked to the likelihood of recognizing that object in a subsequent recognition test (Fig. 1).




(caption on next page)

**Fig. 2.** PE and episodic encoding: Different paradigms and results. a) A typical instrumental learning task (top-left). Participants are asked to predict the object category associated with a particular scene (e.g., savanna - household objects, beach - musical instruments). A trial-unique image belonging to one specific object category is then presented (e.g., glass, accordion). In a subsequent memory test (top-right), participants are then asked to recognize the previously encoded objects among distractors. Bottom: The calculation of PE in the encoding task. b) Effect of positive PEs on recognition memory. c) Interaction between prediction outcome and PE. d) Pavlovian learning task. Participants are presented with a reward-predicting cue and instructed to estimate the reward that they expect to receive for a specific cue. PE and learning rate are then calculated based on an RL model. f) Play/pass paradigm. Participants are presented with a number showing them the reward that they would receive if they win. A trial-unique object related to one object category (e.g., animate/inanimate) is then presented and serves as a cue indicating the probability of the reward. Participants can decide whether to play or pass. After the decision, the object is presented together with the amount of reward or loss. h) Paradigm from (Wimmer et al., 2014). Participants are presented with two trial-unique objects, each one having a different colour frame. They have to decide between the two colours. After the decision is made, participants are presented with the amount of monetary reward received. i) Findings from (Wimmer et al., 2014) showing that higher reward PE was related to a higher number of forgotten items.
(a) Reproduced from (Pupillo et al., 2023). (b) Reproduced from (Pupillo et al., 2023). (c) Reproduced from (Rouhani et al., 2018) e) Unsigned PE effects, reproduced from (Rouhani et al., 2018). (d) Reproduced from (Jang et al., 2019) g) Results related to the play/pass paradigm, reproduced from (Jang et al., 2019), showing an effect of signed reward PEs for "play" trials, and an unsigned effect of signed reward PEs for "pass" trials.

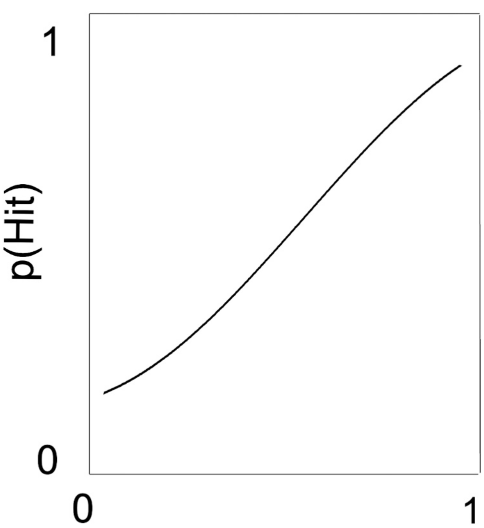



**Fig. 1.** PE and Memory. Logistic regression model used to link the computationally-derived PE at encoding to the likelihood of recognizing an old item (p(Hit)).

Findings by Rouhani and Niv (2021) suggest that the effect of the PE could depend on the task stage in which the PE is elicited. The authors manipulated the PE at two different stages: (1) when a cue was presented that signalled the average reward that could be expected on a trial and (2) when the reward itself was delivered (see Fig. 2d). This design yielded a signed PE effect on memory for the objects presented together with the cue and an unsigned PE effect on memory for the objects presented when reward was delivered. These findings are in line with the study by Jang et al. (2019), also showing positive, signed PE effects for images presented together with a reward-predictive cue (Fig. 2f,g). These studies provide consistent evidence in favour of positive, signed PE effects on memory encoding regarding reward-predictive cues. However, regarding PE effects at outcome, the findings diverge (Ergo et al., 2020). In fact, while some studies have found unsigned PEs effects at outcome (Rouhani and Niv, 2021; Rouhani et al., 2018), others have reported outcome-related signed PEs effects (Calderon et al., 2021; De Loof et al., 2018; Pupillo et al., 2023).

We propose that this discrepancy between signed and unsigned PEs on memory with respect to outcomes might be driven by a choice-confirmation bias that underlies participants' RL behaviour in instrumental but not Pavlovian tasks. In the next paragraphs, we will describe recent results supporting the presence of a choice-confirmation bias in RL and suggest how these findings could reconcile the seemingly contradictory effects of signed and unsigned PEs elicited during outcome delivery on episodic memory.

### *3.1. The choice-confirmation bias*

The observed effect of signed PEs on memory suggests that episodic encoding is stronger when outcomes are better than expected compared to worse-than-expected outcomes. A similar effect of positive PEs has been shown to affect learning (see Box 2). This "positivity" bias yields stronger belief updating in response to positive than negative outcomes. As a consequence, individuals tend to overestimate the likelihood of positive events and underestimate the likelihood of negative events (Sharot and Garrett, 2016). Examples include high-level beliefs, such as the likelihood of getting cancer or becoming divorced (which would be underestimated), but also updating more low-level expectations like choice preferences in an RL task (Lefebvre et al., 2017).

The potential computational origin of the positivity bias is an asymmetry in learning rates for positive and negative PEs, where the "positive" learning rate (after experiencing a positive PE) tends to be larger than the "negative" learning rate (Lefebvre et al., 2017). Interestingly, in the study by Lefebvre et al. (2017), individual differences in asymmetric learning were predicted by the activation of the striatum in response to the PE during the presentation of the outcome. Participants who learned more from positive outcomes showed a stronger activation than participants who learned equally from positive and negative outcomes. These findings suggest that preferentially learning from positive outcomes is rooted in the brain's reward circuits.

Crucially, the positivity bias might ultimately be driven by outcomes that confirm choices, which are in the vast majority of choices positive (see Box 2). This so-called choice-confirmation bias was more closely examined in tasks in which participants could learn from both chosen and unchosen options. Palminteri et al. (2017) argued that if learning was only biased by the valence of the PE, the learning rate for positive outcomes of both chosen and unchosen options should be larger compared to "negative" learning rates. By contrast, if the bias depended on whether the outcomes confirmed or disconfirmed choices, "positive" learning rates in response to chosen options and "negative" learning rates of unchosen options should be larger (see Fig. 3a). Indeed, several recent findings show that participants tend to preferentially learn from PEs that confirm their choices, supporting the presence of a choice-confirmation bias (Palminteri and Lebreton, 2022; Palminteri et al., 2017; Schüller et al., 2020). Therefore, the choice-confirmation bias could be a generalization of the positivity bias, suggesting that the observed positivity effect is a byproduct of a preference for "being right".

Accordingly, the positivity bias would not be present in conditions where individuals cannot choose. A recent study supported this idea, showing that when participants do not have the opportunity to choose, the positivity bias disappears, and participants similarly learn from both positive and negative outcomes (Chambon et al., 2020, Fig. 3b). These findings suggest the intriguing possibility that a bias for positive vs. negative information occurs only in learning conditions where the outcomes inform participants' choices.

**Box 2**
- Positivity and Choice-Confirmation Biases.

To model choice behaviour, RL models can be turned into "agents" that make decisions on the basis of expectations by implementing a specific action selection rule (see Wilson and Collins, 2019). In this case, the expected value $Q_t^a$ learned by the agent corresponds to the reward expected if option $a$ is chosen. $Q_t^a$ is then updated similarly to the $V$ in equation:

$$\begin{aligned} Q_{t+1}^{a} &= Q_t^{a} + \alpha\left(r_t - Q_t^{a}\right) \\ \delta_t^{a} &= r_t - Q_t^{a}. \end{aligned} \tag{3}$$

To test whether participants' choices are differently influenced by positive outcomes compared to negative ones, researchers typically estimate two different learning rates depending on the sign of the PE:

$$Q_{t+1}^{a} = Q_t^{a} + \begin{cases} \alpha^{+}\delta_t^{a}, \delta_t^{a} > 0 \\ \alpha^{-}\delta_t^{a}, \delta_t^{a} < 0 \end{cases} \tag{4}$$

where $\alpha^{+}$ represents a freely estimated learning rate for updating the expected values for better-than-expected outcomes, and $\alpha^{-}$ for updating the expected value in response to worse-than-expected outcomes. Higher values of $\alpha^{+}$ compared to $\alpha^{-}$ suggest a positivity bias.

Moreover, the agents can not only learn from the outcomes related to the chosen option but also from the outcomes of the unchosen options. For example, when having dinner at a restaurant, we might update our expectations about the quality of the food not only based on our dish but also the dishes of our friends. To model this sort of counterfactual learning, the RL model needs to differentiate between chosen and unchosen options. Therefore, the chosen values are updated as follows:

$$\begin{aligned} Q_{t+1}^{c} &= Q_t^{c} + \begin{cases} \alpha^{c+}\delta_t^{c}, \delta_t^{c} > 0 \\ \alpha^{c-}\delta_t^{c}, \delta_t^{c} < 0 \end{cases} \\ Q_{t+1}^{u} &= Q_t^{u} + \begin{cases} \alpha^{u+}\delta_t^{u}, \delta_t^{u} > 0 \\ \alpha^{u-}\delta_t^{u}, \delta_t^{u} < 0, \end{cases} \end{aligned} \tag{5}$$

where $c$ represents the chosen option and $u$ the counterfactual, unchosen option.

Higher values of $\alpha^{c+}$ and $\alpha^{u-}$ compared to $\alpha^{c-}$ and $\alpha^{u+}$ suggest the presence of a confirmation bias (see Fig. 3a).

A reduced version of the aforementioned four-learning rate model that accounts for both factual and counterfactual learning has also been proposed (Palminteri, 2022). In this model, only two learning rates account for the presence of confirmatory and disconfirmatory learning:

$$\begin{aligned} \alpha_{CON} &= \alpha^{c+} = \alpha^{u-} \\ \alpha_{DIS} &= \alpha^{c-} = \alpha^{u+} \end{aligned} \tag{6}$$

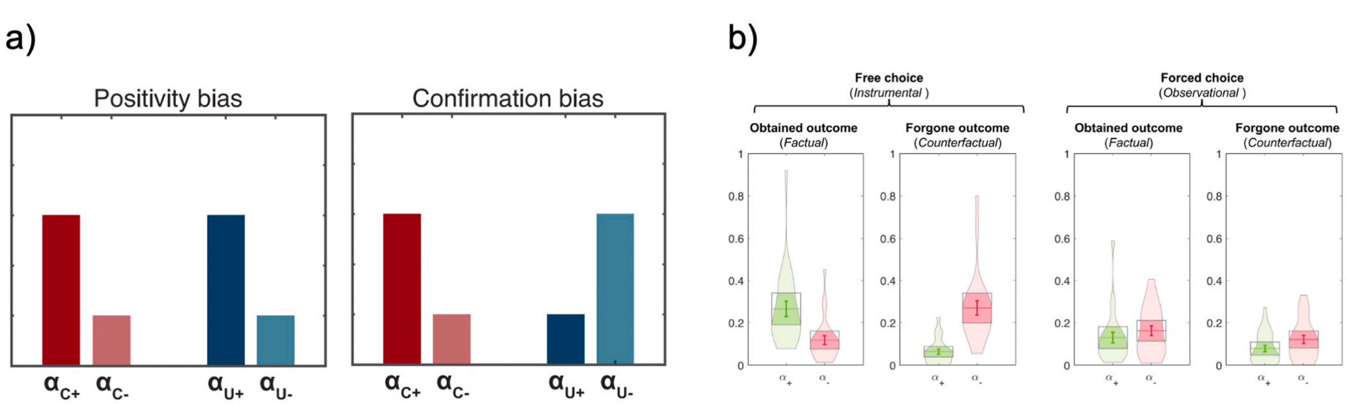



**Fig. 3.** Positivity and Confirmation Bias. a) Different predicted learning rate asymmetries for positivity and confirmation bias. b) Differences in estimated learning rates for free-choice and forced-choice tasks.
(a) Reproduced from (Palminteri et al., 2017). (b) Reproduced from (Chambon et al., 2020).

The choice-confirmation bias seems to be quite robust and might be an adaptive heuristic in some learning conditions. In fact, it generalizes even to tasks that do not show counterfactual outcomes to participants (Lebreton et al., 2019). Moreover, it has been suggested to represent the optimal strategy in specific learning contexts in which individuals need to maximize rewards and minimize losses. Simulation studies showed that RL agents with choice-confirmation bias outperformed unbiased agents in typical RL tasks (Lefebvre et al., 2022). This effect might be because the choice-confirmation bias helps learners better deal with negative PEs due to random outcome variability. In particular, when these negative PEs occur only occasionally and when they should be discarded as uninformative outliers. That way, the choice-confirmation bias yields an overestimation of expected values of favourable options and an underestimation of the value of sub-optimal options, thereby emphasizing value differences that are more robust against random outcome variability (Palminteri and Lebreton, 2022).

In addition, the choice-confirmation bias affects metacognitive confidence judgements. In the study by Salem-Garcia et al. (2023), overestimated expected values due to the choice-confirmation bias led to higher confidence ratings. This effect could be explained by the increased value difference between the choice options (due to the overestimated value of the favourable option), which decreases the subjectively perceived choice difficulty, thereby increasing confidence.

## 4. PE-memory effects: the role of choice

The findings on the choice-confirmation bias highlight an essential learning mechanism that shares similarities with some of the described effects of PEs on memory. The preferential updating for positive information, especially when it confirms one's predictions, means that the updating of expectations is stronger when an outcome is better-than-expected compared to when it is worse than expected. This effect is reminiscent of the effects of signed PEs on episodic memory, showing better memory for objects presented together with better-than-expected outcomes, suggesting that these two effects might share a common underlying mechanism.

Our own recent work showed that choice confirmation was sufficient to modulate the effect of PEs on episodic memory (Pupillo et al., 2023). Using a task without explicit reward, we manipulated participants' expectations leading to PEs of varying degrees. We then fitted an RL model to participants' choice data to compute a PE that depended on the expectancy of the appearance of an object category (see Fig. 2 a-c). Results revealed that stronger (positive) PEs elicited when participants had weak outcome expectations and chose the correct option (i.e., choice was confirmed) led to increased episodic encoding. In contrast, stronger (negative) PEs in response to outcomes disconfirming strong expectations led to decreased encoding. These findings are consistent with the choice-confirmation bias on learning and suggest that we do not only favor choice-confirming PEs when incrementally learning from outcomes but also when selecting what episodes to encode preferentially.

We argue that one crucial factor governing whether signed or unsigned PEs in response to outcomes drive episodic memory encoding is the distinction between instrumental and Pavlovian tasks. At first sight, findings on the role of the choice-confirmation bias in episodic memory encoding discussed thus far may appear at odds with reported effects of unsigned PEs on episodic encoding (Rouhani and Niv, 2019, 2021; Rouhani et al., 2018) according to which the magnitude and not the sign of the PE affects episodic encoding. However, the studies showing unsigned PE effects used Pavlovian tasks in which participants had to predict the reward that they would receive conditional on different stimuli. Each stimulus was associated with a specific reward probability that participants could learn throughout the task, but this learning did not translate into choices (see Fig. 2d). In contrast, the studies finding effects of signed PEs on episodic encoding had in common that they employed instrumental tasks in which the PE was related to participants' choices (see Fig. 2a).

Supporting the idea that the freedom of choosing between options is necessary for the emergence of the choice-confirmation bias, Chambon et al. (2020) showed the presence of the choice-confirmation bias in conditions in which participants could freely choose between options but not in forced-choice conditions (see Fig. 3b). Therefore, also in the above-mentioned studies on the interaction between PEs and episodic memory, it is likely that the effect of unsigned PEs reflected the reduced choice-confirmation bias due to the lack of free choice. Crucially, in instrumental tasks, the effects of positive PEs seem to affect both the learning rate (Lefebvre et al., 2017; Palminteri and Lebreton, 2022) and episodic encoding (Calderon et al., 2021; De Loof et al., 2018; Pupillo et al., 2023), thereby suggesting a shared underlying mechanism. By contrast, in Pavlovian tasks, positive and negative PEs might not lead to different learning rates and seem to have similar effects on episodic encoding.

Findings from Rosenbaum et al. (2022) support the idea that asymmetries in learning rates mirror asymmetries in the effect of PEs on episodic encoding. Using an instrumental task, the authors showed that participants who learned more from positive feedback also remembered images associated with positive PEs better than images with negative PEs. In a separate study that was based on the Pavlovian task from Rouhani et al. (2018), the authors fitted a dual learning rate model (see Box 2, Eq. 4) and showed that participants tended to have similar learning rates for positive and negative outcomes. Participants who presented unbiased learning (no systematic difference between the two learning rates) also had improved memory for images related to both positive and negative PEs at the outcome, and thus an unsigned PE effect.

## 5. Discussion

A considerable body of literature shows that PEs affect episodic memory processes. However, previous studies disagree on the nature of the relationship between PEs and episodic encoding, with some results showing a positive relationship between signed PEs and memory and others showing a positive relationship between unsigned PEs and memory. We propose that this discrepancy regarding the sign of the effect of PEs elicited when outcomes are delivered can be explained by the choice-confirmation bias. The choice-confirmation bias might lead to episodic memory encoding asymmetries between outcomes from chosen and unchosen options, particularly stronger memory encoding after choice-confirming, positive PEs. These asymmetries are not present, or at least considerably weaker, in tasks without explicit choices, such as in Pavlovian learning.

The choice-confirmation bias yields stronger learning from outcomes that confirm a choice but less learning from stochastic, negative PEs. In instrumental tasks, this heuristic strategy could help maximize rewards compared to an unbiased RL strategy, particularly because it might make expected value representations more robust in the face of uncertainty (Lefebvre et al., 2022). Choice-confirmation biases on episodic memory might serve a similar purpose and lead to memory representations that optimize reward-guided decision-making. That is, when a vast amount of potential episodes could be encoded in memory, the choice-confirmation bias might ensure that episodes associated with successful decision-making are preferentially encoded. Therefore, the decision-maker is more likely to recall and utilize such memory representations for future choices.

One potential neural underpinning of the choice-confirmation bias is dopaminergic signalling. It has been suggested that free choices amplify reward-PE signals compared to no-choice trials. Therefore, free choices might be associated with stronger dopaminergic bursts in the striatum in the service of learning (Cockburn et al., 2014). Accordingly, the study from Calderon et al. (2021) showed that the effects of signed PEs on episodic memory are also supported by a PE-related activation of the striatum. This result indicates that some areas that play a role in updating expected values during value-based learning in instrumental tasks (Lefebvre et al., 2017) are also involved in PE-related effects on episodic encoding. Similarly, it is well known that the hippocampus receives dopaminergic input from the striatum that modulates its plasticity (Lemon and Manahan-Vaughan, 2006). Therefore, the connectivity between the striatum and the hippocampus might be responsible for the delivery of PE-related dopaminergic signals to the hippocampus that may result in the prioritization of information that is associated with positive PEs. Moreover, a potential mechanism that may be responsible for reduced episodic encoding in response to negative PEs is hippocampal inhibition through reduced tonic firing (Rosen et al., 2015).

The neural origins of the effects of unsigned PEs on episodic encoding have been linked to arousal-related enhancements of attention to stimuli, leading to improved encoding (Rouhani and Niv, 2021). A promising technique for testing the suggested arousal-unsigned-PE relationship in humans is pupillometry. Pupil dilation is a proxy for central arousal state (McGinley et al., 2015) and has been linked to the activation of the locus coeruleus-norepinephrine system (and other neurotransmitters; Aston-Jones and Cohen, 2005; Cazettes et al., 2021; Joshi et al., 2016). Future studies could test this hypothesis by investigating whether pupil-linked arousal is related to unsigned PE effects on memory encoding in Pavlovian tasks.

While the effect of PEs related to the delivery of outcomes may be influenced by the task type (instrumental vs. Pavlovian), the reported effects of the PE elicited in response to reward-predictive cues seem to be independent of it. In fact, a signed PE effect related to cues has been observed in both instrumental (Jang et al., 2019) and Pavlovian (Rouhani and Niv, 2021) tasks. Because this signed PE occurs before an outcome is presented, it is linked to anticipated and not experienced rewards. In both Jang et al. (2019) and Rouhani and Niv (2021), participants saw different object categories associated with different probabilities of receiving an outcome, where one category was associated with a higher probability of receiving positive outcomes. Results showed that cues associated with positive PEs were linked to better encoding compared to cues associated with negative PEs. Therefore, signed PE effects at cue could be related to the anticipation of the reward, which has also been linked to the strength of episodic encoding (Stanek et al., 2019).

Nevertheless, a direct empirical examination of the hypothesis that the choice-confirmation bias observed in reward-based learning also affects episodic encoding is required. Future investigations should directly compare the effects of PEs on episodic memory across instrumental and Pavlovian tasks. Moreover, they should link the effects of the PE on memory to the behavioural and computational signatures of the choice-confirmation bias. In order to demonstrate that a choice-confirmation bias affects episodic encoding, the following three criteria could be examined (Palminteri and Lebreton, 2022): (1) using model comparison, a model with asymmetric learning rates should have a better fit than a single-learning-rate model; (2) the learning rate for chosen options should be larger for positive compared to negative PEs, while the learning rate for unchosen options should be larger for negative compared to positive PEs (see Fig. 3b); (3) descriptive, model-free behavioural results should show a qualitatively similar pattern to the model-based results related to asymmetric learning rates, like the development of a preference for an option in situations in which participants' choices are confirmed. Once the choice-confirmation bias has been demonstrated and quantified on choice data, it can be tested whether it selectively affects episodic memory in instrumental but not Pavlovian tasks.

Moreover, it has been suggested that the choice-confirmation bias depends on beliefs about the controllability of the environment (Chambon et al., 2020; Dorfman et al., 2019). Studies have shown that having the opportunity to choose enhances episodic encoding (Murty et al., 2015; Yebra et al., 2019), suggesting that the perceived control over the environment influences the prioritization of information in memory. Therefore, perceived controllability over the environment, which should be higher for free choices, could also modulate the effect of the PE on learning and episodic encoding. When perceived controllability is high, negative outcomes may not be perceived as a consequence of one's own choices but of the unpredictability of the environment. Due to a larger choice-(dis)confirmation bias, these negative outcomes could then be under-weighed during episodic encoding. Conversely, when perceived controllability is low, both positive and negative outcomes might be similarly weighed, as both types of outcomes are perceived as independent of one's own choices. Controllability is likely to be generally higher on instrumental tasks, in which participants have some degree of control over the outcomes, compared to Pavlovian tasks where the outcomes are merely observed.

Finally, a better understanding of PE-memory interactions might be relevant for progress in computational psychiatry. Computational approaches to studying learning and decision-making in depression suggest that depressive symptoms are related to a reduced reward sensitivity and signed PE signals in the striatum (Chen et al., 2015). Moreover, depression and particularly anhedonic symptoms have been linked to a reduced positivity bias in learning (Gradin et al., 2011; Kumar et al., 2018), suggesting that interactions between episodic memory and PEs might also be affected by depressive symptoms. A recent study using a Pavlovian learning paradigm examining PE effects on memory and depressive symptoms found reinforcement learning impairments in individuals with more severe depressive symptoms (Rouhani and Niv, 2019). Concerning the interplay of PEs, episodic memory, and depression, the study showed that depression modulated the interaction effect of unsigned and signed PEs on memory. On average, participants' memory was positively affected by unsigned PEs, consistent with related work based on Pavlovian tasks that found better episodic encoding after unsigned PEs (Rouhani and Niv, 2021; Rouhani et al., 2018). However, the study also identified a bias in the unsigned PE modulation of memory. In individuals with self-reported depressive symptoms, unsigned PEs improved memory more strongly when they originated from negative PEs than from positive PEs. In contrast, individuals without depressive symptoms showed the opposite effect, where positive PEs modulated the effect of unsigned PEs on memory more strongly than negative PEs.

To our knowledge, it is currently unclear how the proposed choice-confirmation bias on episodic memory is related to depressive symptoms. It has been shown that individuals with depression have lower perceived controllability and a reduced agency bias compared to healthy controls (Alloy and Abramson, 1979), suggesting that they are less likely to consider outcomes contingent on their actions. Therefore, future studies should further investigate the relationship between perceived controllability, choice-confirmation biases on learning and memory, and depressive symptoms.

## 6. Conclusion

In this review, we have argued that the choice-confirmation bias provides an explanation for the seemingly inconsistent effects of signed and unsigned PEs on episodic encoding. Effects of signed PEs on episodic encoding have been found in instrumental tasks in which participants' choices were either confirmed and associated with stronger episodic encoding or disconfirmed, which was linked to weaker encoding. In contrast, effects of unsigned PEs have been found in Pavlovian tasks in which feedback was delivered regardless of choices. A choice-confirmation bias in episodic memory might have evolved to prioritize memory representations that optimize reward-guided decision-making.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial or non-financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgements**

We thank Javier Ortiz-Tudela, Gözem Turan, Prashanti Ganesh, and two anonymous reviewers for their helpful comments on previous versions of the manuscript. This article was funded by the European Union (ERC-2018-StG-PIVOTAL-758898 for Yee Lee Shing).

## References


Alloy, L.B., Abramson, L.Y., 1979. Judgment of contingency in depressed and nondepressed students: sadder but wiser. J. Exp. Psychol.: Gen. 108 (4), 441.

Aston-Jones, G., Cohen, J.D., 2005. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu. Rev. Neurosci. 28, 403–450.

Calderon, C.B., De Loof, E., Ergo, K., Snoeck, A., Boehler, C.N., Verguts, T., 2021. Signed reward prediction errors in the ventral striatum drive episodic memory. J. Neurosci. 41 (8), 1716–1726.

Cazettes, F., Reato, D., Morais, J.P., Renart, A., Mainen, Z.F., 2021. Phasic activation of dorsal raphe serotonergic neurons increases pupil size. Curr. Biol. 31 (1), 192–197.

Chambon, V., Théro, H., Vidal, M., Vandendriessche, H., Haggard, P., Palminteri, S., 2020. Information about action outcomes differentially affects learning from self-determined versus imposed choices. Nat. Human Behav. 4 (10), 1067–1079.

Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., Kusumi, I., 2015. Reinforcement learning in depression: a review of computational research. Neurosci. Biobehav. Rev. 55, 247–267.

Clark, A., 2013. Whatever next? predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. 36 (3), 181–204.

Cockburn, J., Collins, A.G., Frank, M.J., 2014. A reinforcement learning mechanism responsible for the valuation of free choice. Neuron 83 (3), 551–557.

Davidow, J.Y., Foerde, K., Galván, A., Shohamy, D., 2016. An upside to reward sensitivity: the hippocampus supports enhanced reinforcement learning in adolescence. Neuron 92 (1), 93–99.

Daw, Nathaniel D., 2011. Trial-by-trial data analysis using computational models. Decis. Mak. Affect, Learn. Atten. Perform. XXIII 1–26.

Daw, Nathaniel D., Tobler, P.N., 2014. Value learning through reinforcement: the basics of dopamine and reinforcement learning. Neuroeconomics 283–298.

De Loof, E., Ergo, K., Naert, L., Janssens, C., Talsma, D., Van Opstal, F., Verguts, T., 2018. Signed reward prediction errors drive declarative learning. PLoS One 13 (1), e0189212.

Dorfman, H.M., Bhui, R., Hughes, B.L., Gershman, S.J., 2019. Causal inference about good and bad outcomes. Psychol. Sci. 30 (4), 516–525.

Eichenbaum, H., Yonelinas, A.P., Ranganath, C., 2007. The medial temporal lobe and recognition memory. Annu. Rev. Neurosci. 30, 123–152.

Ergo, K., De Loof, E., Verguts, T., 2020. Reward prediction error and declarative memory. Trends Cogn. Sci. 24 (5), 388–397.

Foerde, K., Shohamy, D., 2011. Feedback timing modulates brain systems for learning in humans. J. Neurosci. 31 (37), 13157–13167.

Friston, K., 2010. The free-energy principle: a unified brain theory. Nat. Rev. Neurosci. 11 (2), 127–138.

Gradin, V.B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Steele, J.D., 2011. Expected value and prediction error abnormalities in depression and schizophrenia. Brain 134 (6), 1751–1764.

Henson, R.N., Gagnepain, P., 2010. Predictive, interactive multiple memory systems. Hippocampus 20 (11), 1315–1326.

Jang, A.I., Nassar, M.R., Dillon, D.G., Frank, M.J., 2019. Positive reward prediction errors during decision making strengthen memory encoding. Nat. Hum. Behav. 3 (July), 719–732.

Joshi, S., Li, Y., Kalwani, R.M., Gold, J.I., 2016. Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. Neuron 89 (1), 221–234.

Kumar, P., Goer, F., Murray, L., Dillon, D.G., Beltzer, M.L., Cohen, A.L., Pizzagalli, D.A., 2018. Impaired reward prediction error encoding and striatal-midbrain connectivity in depression. Neuropsychopharmacology 43 (7), 1581–1588.

Lebreton, M., Bacily, K., Palminteri, S., Engelmann, J.B., 2019. Contextual influence on confidence judgments in human reinforcement learning. PLoS Comput. Biol. 15 (4), e1006973.

Lefebvre, G., Lebreton, M., Meyniel, F., Bourgeois-Gironde, S., Palminteri, S., 2017. Behavioural and neural characterization of optimistic reinforcement learning. Nat. Human Behav. 1 (4), 0067.

Lefebvre, G., Summerfield, C., Bogacz, R., 2022. A normative account of confirmation bias durinreinforcement learning. Neural Comput. 34 (2), 307–337.

Lemon, N., Manahan-Vaughan, D., 2006. Dopamine D1/D5 receptors gate the acquisition of novel information through hippocampal long-term potentiation and long-term depression. J. Neurosci. 26 (29), 7723–7729.

Lisman, J.E., Grace, A.A., 2005. The hippocampal-VTA loop: controlling the entry of information into long-term memory. Neuron 46 (5), 703–713.

McClure, S.M., Berns, G.S., Montague, P.R., 2003. Temporal prediction errors in a passive learning task activate human striatum. Neuron 38 (2), 339–346.

McGinley, M.J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C.R., McCormick, D.A., 2015. Waking state: rapid variations modulate neural and behavioral responses. Neuron 87 (6), 1143–1161.

Murty, V.P., DuBrow, S., Davachi, L., 2015. The simple act of choosing influences declarative memory. J. Neurosci. 35 (16), 6255–6264.

Palminteri, S., 2022. Choice-confirmation bias and gradual perseveration in human reinforcement learning. Behav. Neurosci. 137 (1), 78–88.

Palminteri, S., Lebreton, M., 2022. The computational roots of positivity and confirmation biases in reinforcement learning. Trends Cogn. Sci. 26 (7), 607–621.

Palminteri, S., Lefebvre, G., Kilford, E.J., Blakemore, S.-J., 2017. Confirmation bias in human reinforcement learning: evidence from counterfactual feedback processing. PLoS Comput. Biol. 13 (8), e1005684.

Pearce, J.M., Hall, G., 1980. A model for pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol. Rev. 87 (6), 532.

Pupillo, F., Ortiz-Tudela, J., Bruckner, R., Shing, Y.L., 2023. The effect of prediction error on episodic memory encoding is modulated by the outcome of the predictions. Npj Sci. Learn. 8 (1), 18.

Rosen, Z.B., Cheung, S., Siegelbaum, S.A., 2015. Midbrain dopamine neurons bidirectionally regulate CA3-CA1 synaptic drive. Nat. Neurosci. 18 (12), 1763–1771.

Rosenbaum, G.M., Grassie, H.L., Hartley, C.A., 2022. Valence biases in reinforcement learning shift across adolescence and modulate subsequent memory. eLife 11.

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.

Rouhani, N., Niv, Y., 2021. Signed and unsigned reward prediction errors dynamically enhance learning and memory. eLife 10 (Lc), 1–28.

Rouhani, N., Norman, K.A., Niv, Y., 2018. Dissociable effects of surprising rewards on learning and memory. J. Exp. Psychol. Learn. Mem. Cogn.

Salem-Garcia, N., Palminteri, S., Lebreton, M., 2023. Linking confidence biases to reinforcement-learning processes. Psychol. Rev.

Schüller, T., Fischer, A.G., Gruendler, T.O., Baldermann, J.C., Huys, D., Ullsperger, M., Kuhn, J., 2020. Decreased transfer of value to action in tourette syndrome. Cortex 126, 39–48.

Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science 275 (5306), 1593–1599.

Sharot, T., Garrett, N., 2016. Forming beliefs: why valence matters. Trends Cogn. Sci. 20 (1), 25–33.

Shohamy, D., Turk-Browne, N.B., 2013. Mechanisms for widespread hippocampal involvement in cognition. J. Exp. Psychol. Gen. 142 (4), 1159.

Stanek, J.K., Dickerson, K.C., Chiew, K.S., Clement, N.J., Adcock, R.A., 2019. Expected reward value and reward uncertainty have temporally dissociable effects on memory formation. J. Cogn. Neurosci. 31 (10), 1443–1454.

Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning: An Introduction, Second Edition,Vol. 258, pages 675–676.

Wilson, R.C., Collins, A.G.E., 2019. Ten simple rules for the computational modeling of behavioral data. eLife 8, 1–33.

Wimmer, G.E., Braun, E.K., Daw, N.D., Shohamy, D., 2014. Episodic memory encoding interferes with reward learning and decreases striatal prediction errors. J. Neurosci. 34 (45), 14901–14912.

Yebra, M., Galarza-Vallejo, A., Soto-Leon, V., Gonzalez-Rosa, J.J., Berker, A.O., de, Bestmann, S., Strange, B.A., 2019. Action boosts episodic memory encoding in humans via engagement of a noradrenergic system. Nat. Commun. 10 (1), 3534.