BRIEF REPORT

Reward prediction errors drive declarative learning irrespective of agency

Kate Ergo[1] · Luna De Vilder[1] · Esther De Loof[1] · Tom Verguts[1]

Abstract

Recent years have witnessed a steady increase in the number of studies investigating the role of reward prediction errors (RPEs) in declarative learning. Specifically, in several experimental paradigms, RPEs drive declarative learning, with larger and more positive RPEs enhancing declarative learning. However, it is unknown whether this RPE must derive from the participant's own response, or whether instead, any RPE is sufficient to obtain the learning effect. To test this, we generated RPEs in the same experimental paradigm where we combined an agency and a nonagency condition. We observed no interaction between RPE and agency, suggesting that any RPE (irrespective of its source) can drive declarative learning. This result holds implications for declarative learning theory.

Traditionally, long-term memory is divided into two major memory systems: declarative and nondeclarative memory (Doll et al., 2015; Poldrack & Gabrieli, 1997; Squire, 2004; Squire & Dede, 2015). Nondeclarative memory encompasses several subtypes of memory, one of which is procedural memory, or the memory for skills and actions. In contrast with declarative learning, procedural learning is characterized by a slow process requiring repeated practice (e.g., learning how to ride a bicycle). Declarative memory instead refers to factual knowledge, experiences, and concepts that are accessible through conscious recall (Squire, 2004, 2009). Here, memoranda are learned fast, often after a single encounter (Eichenbaum, 2004; Shohamy & Adcock, 2010). The process of acquiring such memories is referred to as declarative learning.

One important framework to understand learning and memory is reinforcement learning (RL). Here, learning is guided by the experience of reward prediction errors (RPEs; i.e., mismatches between reward outcome and reward expectation; Sutton & Barto, 2018). Indeed, nondeclarative learning and memory formation have long since been associated with RPEs. These RPEs are thought to originate from the dopaminergic midbrain and project to cortical and subcortical (e.g., striatal) pathways to support learning (Schultz et al., 1997). Recent work, however, also demonstrated an important role for RL and RPE in declarative learning and memory (Calderon et al., 2021; Davidow et al., 2016; Gershman & Daw, 2017; Jang et al., 2019; Mattar & Daw, 2018; Rouhani et al., 2018; Rouhani et al., 2020).

RPEs in declarative learning have been studied with two main approaches (Ergo et al., 2020b). In the multiple-repetition approach, the same stimuli are presented multiple times, and participants have to estimate their probability of success based on their certainty and the feedback they receive. For instance, in Butterfield and Metcalfe (2001) participants answered general information questions from a wide variety of topics, rated their confidence after each answer, after which they received (corrective) feedback. Perhaps counterintuitively, stimuli accompanied with large absolute RPEs (i.e., high-confidence errors) were better remembered; a phenomenon referred to as the hypercorrection effect (HCE). The HCE has been found after both immediate and delayed (re)testing (Butterfield & Mangels, 2003; Eich et al., 2013; Metcalfe & Eich, 2019; Metcalfe & Finn, 2012). Given that making (unexpected) mistakes might be accompanied by a sense of surprise, one possible explanation for the HCE has been attributed to attentional capture due to surprise (Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009). In the

✉ Kate Ergo
kate.ergo@ugent.be

[1] Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium

RL framework, surprise can be conceptualized as a prediction error, which is known to guide attention.

A second way of generating and measuring the effect of RPEs is the reward-prediction approach. Here, participants learn declarative information while reward is sampled from a statistical distribution that must be estimated by the participant. For example, in Jang et al. (2019), participants were given the opportunity to gamble on each trial. On each trial, a potential reward was shown, followed by an image indicating the reward probability of a correct gamble, with different images inducing different reward probabilities. Based on these two pieces of information (potential reward and reward probability), participants could compute their reward expectation (reward prediction) on that trial. After making their choice, they were given reward feedback (reward outcome); thus, the participant could compute the RPE (i.e., RPE = reward outcome − reward prediction), and different images thus led to different RPE values. Performance on the subsequent recognition test was enhanced with increasing RPEs. Similarly, using the recent variable-choice experimental paradigm, we have repeatedly shown that RPEs boost declarative learning on both behavioral and neural levels (Calderon et al., 2021; De Loof et al., 2018; Ergo et al., 2019; Ergo, De Loof, Debra, et al., 2020a). In one version of this paradigm, participants learned Dutch–Swahili word pairs. Different Dutch words were associated with a different number of Swahili options to choose from (in a multiple-choice format), and thus with different RPEs. Throughout multiple experiments, we repeatedly found a signed RPE (SRPE) effect. Specifically, the larger and more positive the RPE, the better subsequent memory performance was. RPEs may thus be fundamentally important in the formation of new declarative memories.

In all results discussed thus far, the RPE was the result of participants' own actions. It is unclear whether RPEs elicited without the participant's active contribution (i.e., without agency) would also facilitate declarative learning. Agency is defined as the perceived control over learning and the opportunity to make choices (Murty et al., 2015). This feeling of control is considered valuable (Fujiwara et al., 2013; Leotti et al., 2010), increases engagement with the material, and facilitates declarative learning (Markant et al., 2016). In addition, evidence suggests that the opportunity to choose is inherently rewarding itself and modulates mesolimbic dopaminergic (Leotti & Delgado, 2011) and striatal (Leotti & Delgado, 2014; Wang & Delgado, 2019) pathways. In Murty et al. (2015), participants were presented with trial-unique objects hidden behind two occluders. Enhanced memory performance for the memoranda was observed when participants had the opportunity to choose which occluder to remove, compared with computer choices. Memory performance was correlated with stronger functional connectivity between the striatum and hippocampus during encoding. This connectivity was more pronounced for participant

choices compared with computer choices. Using a similar paradigm, DuBrow et al. (2019) replicated the finding that the opportunity to choose increased memory for the chosen item. In addition, preference increased for chosen items compared with yoked items. In another study investigating the effect of active versus passive choices on memory, Rotem-Turchinski et al. (2019) showed participants video clips where they had the opportunity to either choose themselves (active condition) or let the computer (passive condition) choose how the video clips ended. Participants were then tested on details conveyed in the video clips and the choice that was made (by themselves or by the computer) after either a 2-day or a 1-week delay. The opportunity to choose the outcome of the video clips positively influenced recognition memory, even when tested after a significant delay. Moreover, even being able to choose the order and the timing of item presentation, has been associated with increased memorization in both adult (Markant et al., 2016) and child (Ruggeri et al., 2019) learners. Overall, evidence suggests that being in control of one's own learning experience by having the opportunity to choose is critical for successful declarative learning (but see Katzman & Hartley, 2020, and the Discussion section).

Despite the evidence that both RPE and agency improve declarative memory and are associated with dopaminergic activity in the midbrain, their mutual relation remains unclear. Knowledge of this interaction would, however, clarify the nature of RPE in declarative learning. Specifically, in RL theory, a fundamental distinction concerns learning about states (i.e., the environment) versus learning about one's own actions (Sutton & Barto, 2018). Traditionally, RPEs have been associated with learning about one's own actions (as is usually the case in procedural learning). For example, in operant (or instrumental) conditioning paradigms, RPEs are utilized to learn stimulus–action associations (Skinner, 1990; Thorndike, 1932). However, RL models suggest that RPEs may also be used to learn about states (including in declarative memory; Rouhani et al., 2018; Rouhani et al., 2020), in which case they are referred to as state prediction errors (Mattar & Daw, 2018). Historically, state prediction errors were studied in classical (or Pavlovian) conditioning paradigms (Pavlov, 1902). From this perspective, an RPE effect in declarative learning may occur either because the agent learns about an action (e.g., I choose this Swahili word), or it may occur because RPEs also drive learning about states (e.g., this Dutch word translates into that Swahili word). Neurally, these two learning signals are known to coexist within the human brain (Gläscher et al., 2010).

Unfortunately, earlier RPE declarative learning paradigms could not disentangle these two theoretical possibilities. However, it is possible to do so with an experimental design where agency and RPE are crossed. Specifically, if an RPE in declarative memory only helps because one is learning about one's own actions, the RPE effect should only occur in an

agency condition—that is, in a condition where one chooses (acts) oneself. It should not have an effect in a nonagency condition. In this case, an interaction effect is expected between RPE and agency. In contrast, if an RPE effect in declarative memory also helps for learning about states (such as a Dutch–Swahili word pair), then the RPE should also have a beneficial effect in a nonagency condition (e.g., when the computer chooses a word, rather than the participant). In the latter case, no interaction effect is expected between agency and RPE.

To address this issue, we used a variable-choice paradigm where participants learned 84 Dutch–Swahili word pairs. Each word pair was associated with a unique RPE value. Half of the trials were assigned to an agency condition, while the remaining trials were assigned to a nonagency condition. We expected to replicate our previous finding of RPEs driving declarative learning. More specifically, we anticipated that large, positive RPEs would lead to increased memory performance. In addition, we sought to evaluate whether RPE interacts with agency or not.

Methods

Participants

All participants were recruited through Ghent University's online recruitment platform. We tested a total of 37 participants. One participant was removed from further analysis due to below-chance-level performance (<25%) on the recognition test (33 females, range: 18–40 years, M = 19.5 years, SD = 4.5 years). Participants were given partial course credit. Before partaking in the study, participants signed an informed consent form. No participant had prior knowledge of Swahili. The participant with the best performance in the recognition test was additionally rewarded with a gift voucher worth €20.

Material

A total of 420 words (84 Swahili words and 336 Dutch words; see Tables 1 and 2 in Appendix A) were used. The experiment was programmed in PsychoPy2 (Peirce, 2007).

Procedure

Familiarization task

To familiarize participants with the stimuli used in the experiment and to control for the novelty of the Swahili words, a familiarization task was included at the start of the experiment. All words (N = 420) were randomly presented for 2 seconds. Participants were instructed to press the space bar only when a Dutch word was presented.

Acquisition task

Participants learned Swahili translations of 84 Dutch words. On each trial, one Dutch word was presented on top of the screen together with four Swahili translations below, of which only one was the correct translation (see Fig. 1a). After four seconds, a cue was presented on top of the screen indicating whether the participant (agency condition; "You choose") or the computer (nonagency condition; "Computer chooses") had to make a choice. As an additional cue, the color of the Dutch word changed to cyan or magenta (counterbalanced across participants). In addition, frames surrounded either one, two, or four Swahili translations. These frames indicated which Swahili translations were eligible as the translation for the Dutch word. In the one-option condition, only one Swahili translation was framed and the probability of being correct was 100%. In the two-option condition, two Swahili translations were framed and there was a 50% probability of obtaining reward. This probability of being correct was reduced to 25% when presented with the four-option condition in which all four Swahili translations were framed. Each Swahili translation was assigned one response key ('f', 'v', 'n', or 'j') and participants had to respond with the middle and index finger of the left and right hand, respectively. For participant choices, there was no time limit to respond, while the response time for computer choices was drawn from a uniform distribution between 1 and 4 seconds. After a choice was made (by the participant [agency condition] or the computer [nonagency condition]), feedback was presented in which the to-be-learned Dutch–Swahili word pair was shown for 5 seconds. The Dutch word, an equation sign, and its Swahili translation appeared on the screen. Participants were instructed to use this time to encode the word pair as they knew their memory for all 84 word pairs would be tested in a subsequent recognition test. If the chosen Swahili translation was rewarded, a green frame was presented around the Dutch–Swahili word pair. Alternatively, if the chosen Swahili translation was unrewarded, a red frame appeared around the (to-be-learned) Dutch–Swahili word pair. Each trial ended with a reward update for 2 seconds. Participants received 1 point on rewarded trials (in agency and nonagency conditions), while no points were added on unrewarded trials (in agency and nonagency conditions). The participant with the highest number of points was rewarded a gift voucher worth €20. In Fig. 1a, the two-option condition with rewarded choice is illustrated.

Design For each participant, agency was manipulated by randomly assigning half of the trials (N = 42) to the agency condition and the other half of the trials (N = 42) to the nonagency condition (see Fig. 1b).

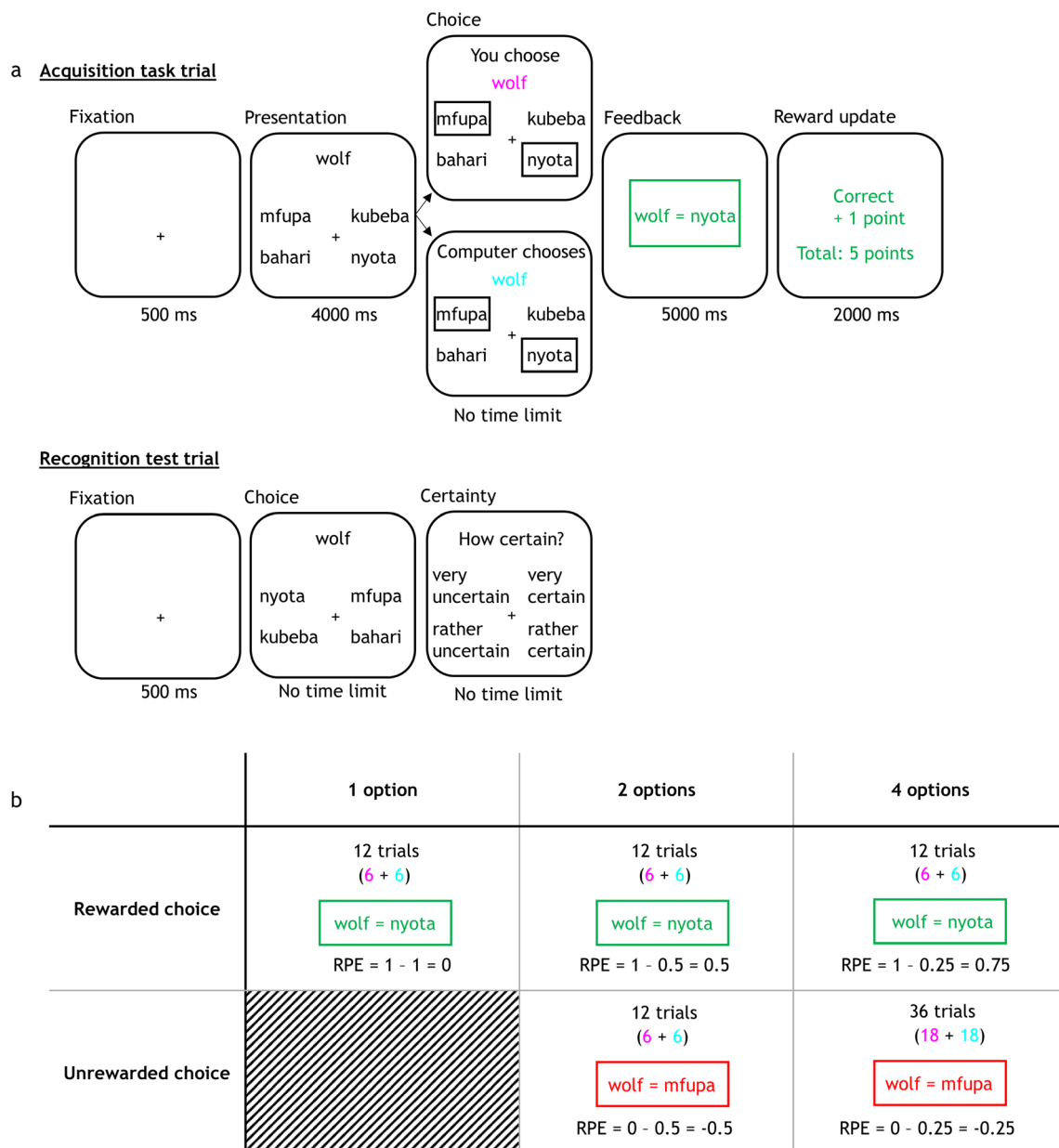RPE magnitudes were parametrically modulated by a priori determining the number of options (one, two, or

Fig. 1 Overview of the (**a**) Experimental Trials and (**b**) Experimental Design. **a** In the acquisition task, the participant (agency condition) or computer (nonagency condition) chose between one, two, or four Swahili translations. An acquisition trial is illustrated from the "2 options, rewarded choice" cell in the experimental design, and the participant/ computer chose 'nyota' as the translation for 'wolf'. Hence the word pair to encode was 'wolf = nyota'. If the participant/computer would alternatively have chosen 'mfupa', the feedback would again (by design) have been positive, but in this case, the word pair to encode would have been 'wolf = mfupa'. After the 84 acquisition trials, the recognition test was performed immediately after the filler task. **b** The 2 (agency/nonagency) × 2 (obtained reward) × 3 (number of options) experimental design, including the number of trials and associated RPE. RPEs were calculated by subtracting the probability of reward from the obtained reward

four options) as well as the reward (reward/no reward) on each trial. By doing so, an RPE for each cell of the design could be computed (see Fig. 1b). Note that by predetermining reward feedback at each trial, participants did not necessarily learn the real Swahili translations of the Dutch words. For example, if a trial was rewarded, participants received positive feedback irrespective of their choice (but this choice then became their to-be-learned translation of the Dutch word). Participants were debriefed about this manipulation afterward. For rewarded trials, reward outcome was equal to one, whereas reward outcome was equal to zero for unrewarded trials. Reward probability was determined by the number of eligible options (one, two, or four options). RPEs were obtained by subtracting reward probability (which in this case equals reward prediction) from reward outcome.

Filler task

To avoid recency effects, participants performed a magnitude comparison task immediately after the acquisition task. A total of 400 digits ranging from 1 to 9, with the exclusion of 5, were sequentially presented on the screen. Participants pressed 'f' for digits smaller than 5 and 'j' for digits larger than 5.

Recognition test

Participants were again presented with the 84 Dutch–Swahili pairs from the acquisition task. On each trial, the Dutch word appeared on top of the screen together with the same four Swahili translations from the acquisition task (see Fig. 1a). To avoid that participants would choose solely based on spatial location, the position of the Swahili translations was randomized. This time, no frames surrounded the Swahili translations. No time constraint was imposed. Participants made their choice by pressing one of the four designated response buttons ('f', 'v', 'n', or 'j'). After they made their choice, participants were asked how certain they were of their answer: 'very uncertain', 'rather uncertain', 'rather certain', or 'very certain' (measured on a scale from 1 = 'very uncertain' to 4 = 'very certain').

Data analysis

All behavioral data were analyzed using the linear mixed-effects framework in R software (R Core Team, 2014). For continuous dependent variables (e.g., certainty ratings in the recognition test), linear mixed-effects models were used, while for categorical dependent variables (e.g., recognition accuracy), generalized linear mixed-effects models were applied. A random intercept for participant was included in each model. All (fixed effects) predictors were mean centered. RPEs were treated as a continuous predictor, allowing the inclusion of all 84 trials per participant to estimate its regression coefficient. Nonagency trials on which a button press was made by the participant were removed from further analysis, resulting in a loss of 3% of the total number of trials. While building our statistical models, we used a bottom-up modeling approach to control the risk of Type I errors and to verify the validity of adding random slopes. This approach allowed us to leave out insignificant random slopes from the start. In general, this modeling process involved four steps: (1) We evaluated the fixed effects and their random slopes by fitting four models. The first model was the baseline model with only the random intercept for participant. The second model additionally had a fixed effect for one of the variables of interest (e.g., reward). The third model had again a random intercept for participant, no fixed effect, but a random slope for the same variable of interest. Finally, the fourth model had both a fixed and a random slope for this variable (and again, a random intercept for participant). Next, we compared all of these models to test whether adding the random slope and/or fixed effect made the model significantly better. The significant random and fixed effects of the first predictor remained in the model. We then added in a second predictor (e.g., number of options), again checking whether the random and fixed effects of the second predictor were significant. This continued until all predictors had been tested. (2) Here, we only fitted fixed effects. We started with the most significant predictor (as determined in Step 1) and kept on adding weaker (i.e., less statistically significant) predictors until we ran out of (significant) predictors. The goal of this series of models was to check whether we obtained similar p values as in Step 1 without the random slopes. If this was the case, we could leave out the random slopes from the model. (3) We then moved on to the interactions. To fit the interaction effects, we started with all predictors in a huge interaction model, and then gradually removed the predictors that were not involved in any statistically significant interactions. Note that we only tested for interactions in the fixed effects, not in the random slopes. (4) In the end, we combined all the effects that we found in one model: The significant random slopes that could not be avoided because they would otherwise cause a Type 1 error in the fixed effects obtained from Step 1, the significant fixed main effects from Step 2, and the significant interactions from Step 3. The four steps described above were performed for all the dependent variables of interest. We reported the χ^2 statistics from the ANOVA Type III tests.

In addition to frequentist statistics, we also reported Bayesian repeated measures analyses of variance (ANOVAs) that were performed in JASP (Version 0.13; JASP Team, 2020). In Bayesian ANOVAs, recognition accuracy and certainty ratings were analyzed as a function of SRPE and agency. Bayes factors (BFs) quantify the evidence in favor of the null hypothesis (BF_{01}; e.g., agency does not influence memory performance) or the alternative hypothesis ($BF_{10} = 1/BF_{01}$; e.g., agency influences memory performance). BF_{01} was reported when the Bayesian analysis provides relatively more evidence for the null hypothesis; BF_{10} was instead reported when the analysis provides relatively more evidence for the alternative hypothesis. As a test on the robustness of our findings, we also investigated the effect of the prior specification (specifically, the Cauchy distribution scale parameter) for all analyses (see Rouder et al., 2012). All priors were centered around zero. Jeffreys' benchmarks (Jeffreys, 1961) were used to determine the strength of evidence, with BFs corresponding to anecdotal (0–3), substantial

(3–10), strong (10–30), very strong (30–100), or decisive (>100) evidence.

Results

Recognition accuracy

Our bottom-up modeling approach revealed that no random slopes were necessary for the models described below. We first examined the effect of reaction times (RTs) in the acquisition task (i.e., how much time participants spent making their choice) on recognition accuracy. We found that participants spent more time in the nonagency condition ($M = 2.84$ seconds, $SD = .19$ seconds, range: 2.49-3.24 seconds) compared with the agency condition ($M = 1.73$ seconds, $SD = .76$ seconds, range: .57-4.24 seconds). To investigate whether RT could thus be a confound, we verified whether RT in the nonagency condition affects recognition accuracy, which turned out not to be the case $\chi^2(1, N = 36) = .17, p = .68$. The effect of RT in the agency condition is harder to interpret, given that it may correlate with number of options. Nevertheless, for completeness, we checked the effect of RT in the agency condition as well, and observed no effect, $\chi^2(1, N = 36) = .05, p = .82$. Hence, we did not add RT as a confound regressor to the model.

The data revealed a significant main effect of reward, $\chi^2(1, N = 36) = 16.16, p < .001$. Recognition accuracy was lower for unrewarded choices ($M = 55.2\%, SD = 14.0\%$, range: 31%–90%) compared with rewarded choices ($M = 61.1\%, SD = 13.4\%$, range: 33%–86%). Furthermore, recognition accuracy increased with number of options, $\chi^2(1, N = 36) = 12.21, p < .001$ (one-option: $M = 55.1\%, SD = 19.2\%$, range: 17%–100%; two-option: $M = 57.2\%, SD = 15.5\%$, range: 25%–88%; four-option: $M = 58.8\%, SD = 13.7\%$, range: 35%–88%). Finally, the interaction between reward and number of options was not significant, $\chi^2(1, N = 36) = .03, p = .86$.

Next, we analyzed whether recognition accuracy linearly increased with SRPEs. The data revealed a significant positive effect of SRPE, $\chi^2(1, N = 36) = 24.36, p < .001$, with larger and more positive SRPEs leading to increased recognition accuracy (see Fig. 2a–b; see Table 3 in Appendix B for mean accuracies for each SRPE and agency condition). To make sure that the observed SRPE effect was not a mere reward effect, we reran the analysis for number of options, separately for rewarded versus nonrewarded trials. There was a main effect of number of options for both rewarded, $\chi^2(1, N = 36) = 8.02, p = .005$, and for unrewarded, $\chi^2(1, N = 36) = 4.67, p = .031$, trials. Additionally, we also reran the analysis including both rewarded and unrewarded trials with a reward and SRPE regressor. This revealed similar results. There was still a main effect of

SRPE, $\chi^2(1, N = 36) = 10.69, p = .001$; interestingly, there was no effect of reward, $\chi^2(1, N = 36) = 2.87, p = .090$. Together, these results suggest that our effect was indeed driven by SRPE, not by mere reward.

There was a main effect of agency on recognition accuracy, $\chi^2(1, N = 36) = 13.01, p < .001$ (agency: $M = 62.7\%, SD = 15.5\%$, range: 36%–95%, nonagency: $M = 53\%, SD = 13.6\%$, range: 26%–81%). The interaction between SRPE and agency was, however, not significant, $\chi^2(1, N = 36) = 1.67, p = .20$. SRPEs did not influence accuracy differently on agency versus nonagency trials. Additionally, we verified whether SRPE increased declarative learning within each condition. To do so, we reran the analyses separately for the agency and nonagency conditions. The data revealed that SRPE drives declarative learning in both the agency, $\chi^2(1, N = 36) = 19.42, p < .001$, and nonagency condition, $\chi^2(1, N = 36) = 6.74, p = .009$.

Bayesian repeated-measures ANOVAs provided substantial evidence for an agency effect ($BF_{10} = 7.94$, compared with the null model). The data were about 8 times more likely under the alternative hypothesis than under the null hypothesis. The evidence for the SRPE effect was decisive ($BF_{10} > 100$, compared with the null model). In addition, there was strong evidence against the interaction of SRPE and agency ($BF_{01} = 19.33$, compared with the two-main-effects model). Importantly, Bayes factor robustness checks showed that the results yielded strong evidence for an SRPE effect and against an interaction effect of SRPE and agency over a wide range of prior scales (see Figs. 3a–c, and 4a in Appendix C).

Certainty ratings

Our bottom-up modeling approach revealed that no random slopes were necessary in the models. For the certainty ratings there was a significant main effect of recognition accuracy, $\chi^2(1, N = 36) = 823, p < .001$, indicating that participants were more certain of correctly recognized word pairs (Fig. 2c–d). There was also a main effect of agency, $\chi^2(1, N = 36) = 8.16, p = .004$. Participants were more certain on agency versus nonagency trials. The interaction between recognition accuracy and SRPE was also significant, $\chi^2(1, N = 36) = 8.51, p = .004$. Follow-up analysis revealed that SRPE only influenced certainty for correctly recognized word pairs, $\chi^2(1, N = 36) = 1.63, p = .001$, but not for incorrectly recognized word pairs, $\chi^2(1, N = 36) = 823, p = .20$.

A Bayesian repeated-measures ANOVA revealed strong evidence for an agency effect ($BF_{10} = 27.32$, compared with the null model). For the SRPE effect, the evidence was decisive ($BF_{10} > 100$, compared with the null model). We also found strong evidence against the interaction of SRPE and agency ($BF_{01} = 17.35$, compared with the two-main-effects model). Crucially, Bayes factor robustness checks revealed
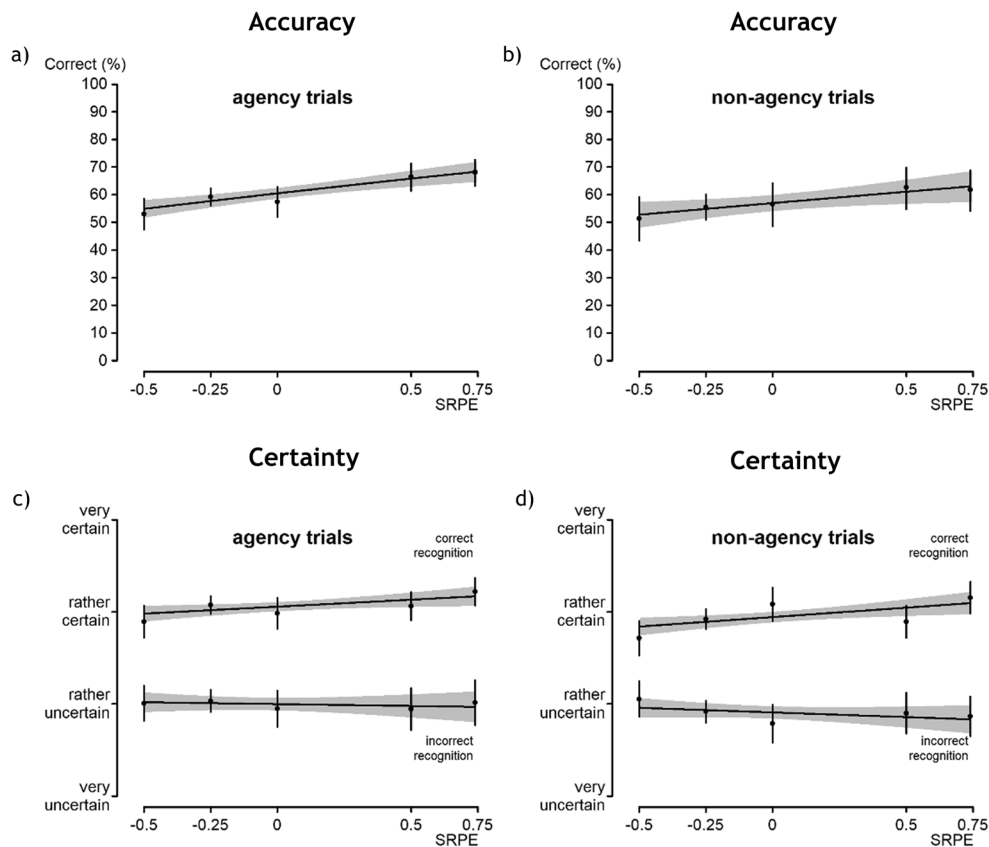
Fig. 2 Behavioral Results. **a–b** Recognition accuracy as a function of SRPE for the agency and nonagency trials, respectively. Recognition accuracy increased linearly with larger and more positive SRPEs in both trial types. The interaction between agency and SRPE was not significant, indicating that SRPEs drive declarative learning irrespective of its origin (i.e., coming from the participant's own action or the participant's environment). **c–d** Certainty rating for the agency and nonagency trials, respectively. SRPE significantly predicted certainty for correctly recognized word pairs, but not for incorrectly recognized word pairs on both trial types

strong evidence for the null hypothesis of our effect of interest (i.e., the interaction effect SRPE × Agency) across a wide range of priors (see Figs. 3d–f, and 4b in Appendix C).

Discussion

In the current study, we investigated whether agency influenced the RPE effect in declarative learning. To do so, we used a variable-choice paradigm in which participants learned 84 Dutch–Swahili word pairs. In half of the trials, participants made a choice themselves (i.e., agency condition), whereas in the other half of the trials, the computer chose for them (i.e., nonagency condition). We replicated our previous finding of SRPE-driven declarative learning, with increased word pair recognition for large, positive RPEs. In line with earlier studies (DuBrow et al., 2019; Murty et al., 2015; Rotem-Turchinski et al., 2019), we found a main effect of agency, with increased recognition on trials where participants chose themselves, compared with computer choices. As an important extension to the previous line of studies discussed, we

combined RPEs and agency within the same experiment. The interaction between agency and RPE was not significant (and Bayesian statistics provided strong evidence against an interaction), indicating that the RPE effect on declarative learning is not modulated by agency. This finding can be interpreted within an RL framework, according to which cognitive agents utilize RPEs to collect knowledge about states in their environment and their own actions. Traditional empirical work in procedural learning has focused on actions only; the absence of a significant interaction between agency and RPE in the current study indicates that participants use RPEs to learn about both states and their own actions while performing a declarative learning task (see also Rouhani et al., 2018; Rouhani et al., 2020).

Surprisingly, in contrast to the current and earlier studies where a mere effect of agency was found (DuBrow et al., 2019; Murty et al., 2015; Rotem-Turchinski et al., 2019), Katzman and Hartley (2020) argued that agency itself is not sufficient to enhance memory. In their experiment, participants performed a memory task where the utility of agency (i.e., the degree to which participants' choices were rewarded
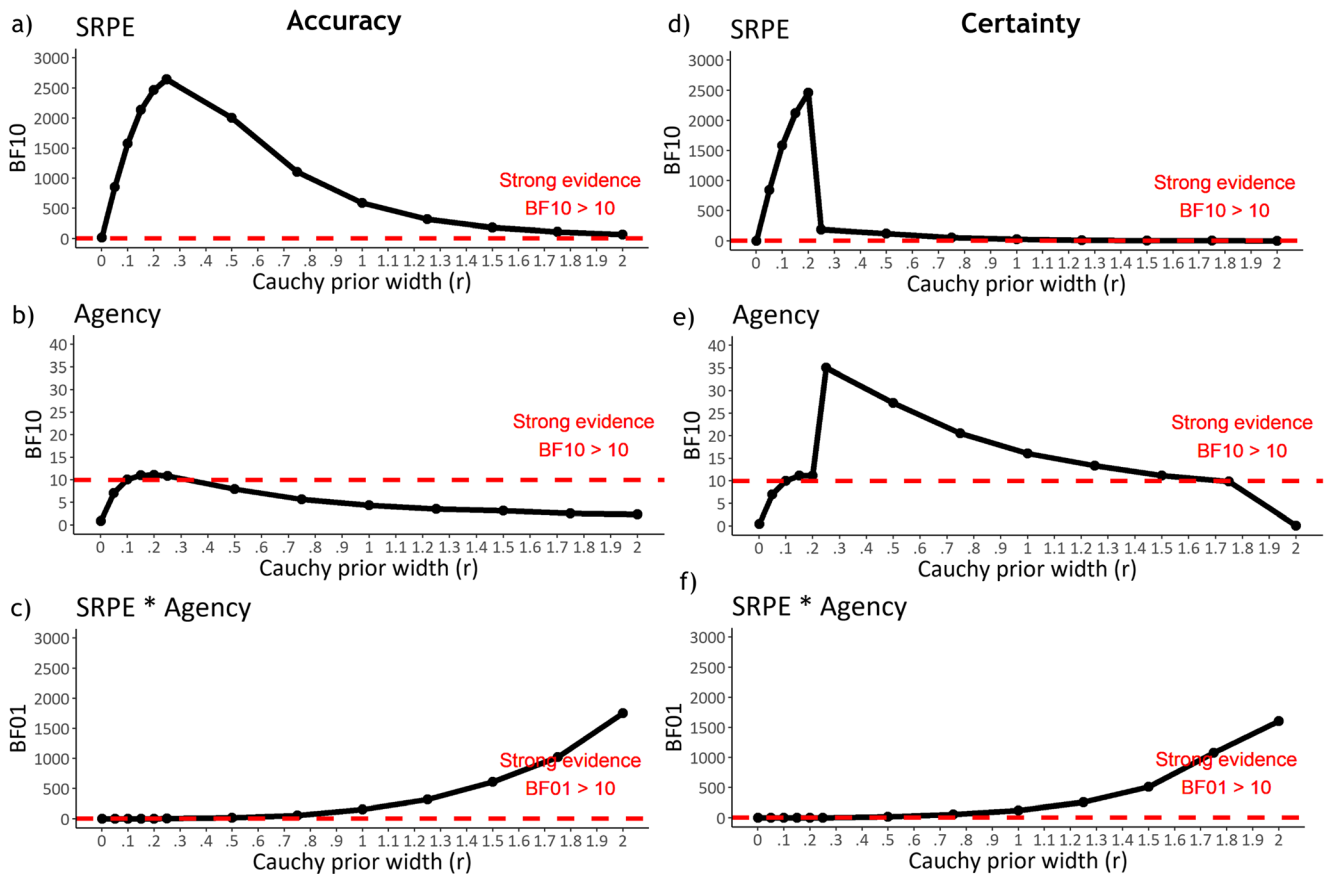
Fig. 3 Bayes factor robustness checks. Bayes factor (*BF*) robustness checks for all predictors of interest. In general, the results remained roughly the same over a wide range of Cauchy priors with scale parameter *r* ranging from 0 to 2. We reported BF_{10} for the main effects of SRPE and agency (higher *BF* is more evidence for the alternative hypothesis), while for the interaction effect SRPE × Agency, BF_{01} is reported (higher *BF* is more evidence for the null hypothesis). **a** BF_{10}

robustness check for the effect of SRPE on recognition accuracy. **b** BF_{10} robustness check for the effect of agency on recognition accuracy. **c** BF_{01} robustness check for the interaction effect of SRPE × Agency on recognition accuracy. **d** BF_{10} robustness check for the effect of SRPE on certainty. **e** BF_{10} robustness check for the effect of agency on certainty. **f** BF_{01} robustness check for the interaction effect of SRPE × Agency on certainty

or not) was manipulated. On each trial, participants were first given a context cue (i.e., which galaxy they were in), followed by an agency cue (i.e., the computer makes a choice [nonagency] versus the participant makes a choice [agency] about what planet to travel to). Different planets were associated with different reward probabilities. After a choice was made, they were shown a trial-unique image followed by feedback (i.e., reward; implemented by whether the inhabitants of the planet considered the trial-unique image treasure or trash). Agency was manipulated in such a way that three learning environments could be distinguished: A nonlearnable, nonagency environment, and two learnable environments where there was no and high utility of agency, respectively. Participants performed old/new judgments on the images after a 1-day delay. The data revealed that the high-utility but not the no-utility agency condition increased memorization relative to the nonagency condition. Interestingly, this result contradicts our (as well as earlier studies') finding of a main effect of agency in declarative learning, irrespective of its utility.

Indeed, in our experiment, agency had no utility (i.e., reward rates were exactly equal for agency and nonagency conditions in the current experiment). However, in which circumstances agency improves memory, remains to be investigated more systematically.

Incidentally, our study ruled out a potential confound in the variable-choice design. Specifically, in this design, higher RPE values necessarily derive from word pairs with a larger number of eligible options. In principle, it is possible that with more eligible options, there is a higher probability of choosing an intuitively attractive word-word association. So, one could argue that the advantage for a higher number of options does not derive from its relation with RPE, but simply because more choice options are associated with a higher probability of an attractive word-word association. The current data allowed addressing this hypothesis because this argument could only work if participants chose themselves; not in the nonagency condition, in which there were simply no choice options. Therefore, if this alternative hypothesis were true, the

data should have revealed a significant interaction between agency and RPE. More specifically, under the confound hypothesis, the RPE effect should have disappeared on nonagency trials, as participants did not have the opportunity to choose themselves between the eligible options. Instead, the RPE effect should have appeared on agency trials, as participants made active choices themselves on these trial types. The data revealed no significant interaction between agency and RPE, demonstrating that the RPE effect in our experiment was not merely driven by the number of eligible options participants could actively choose from.

Whereas RPEs were mainly studied within procedural learning, which naturally focuses on learning from actions, recent studies have shown a role for RPEs in declarative learning as well. Here, we speculate about the potential mechanism underlying the effect of RPEs in declarative learning. Neurally, RPEs are computed in the dopaminergic midbrain (i.e., ventral tegmental area [VTA] and substantia nigra [SN]) and are projected to various brain regions, including the hippocampus (Shohamy & Adcock, 2010) and ventral striatum (Watabe-Uchida et al., 2017). Midbrain VTA activation (triggered by RPEs) plays a significant role in RL (Montague et al., 1996) and has been associated with increased declarative learning (Calderon et al., 2021; Gruber et al., 2016; Wittmann et al., 2005). According to the neoHebbian framework (Lisman et al., 2011), dopaminergic RPEs promote declarative memory by increasing synaptic learning efficiency directly during the acquisition process. Alternatively, RPEs may modulate learning during off-line hippocampal replay (Skaggs & McNaughton, 1996; Wilson & McNaughton, 1994). During hippocampal replay, neural activity patterns (representing environmental states) in hippocampal pyramidal neurons (e.g., place cells) that occurred during activity, are sequentially reactivated. Hippocampal replay can take place during sleep and/or (off-line) wakefulness (Pfeiffer, 2020) and has been evidenced in nondeclarative learning (Momennejad et al., 2018). However, to efficiently learn from replay, the brain has to decide *which* memories to replay. Hippocampal replay is sometimes considered to be modulated by unsigned RPEs (URPEs) where the absolute value of an RPE is computed (e.g., Khamassi & Girard, 2020; Momennejad et al., 2018; Roscow et al., 2019). However, some computational models argue for the importance of SRPE in hippocampal replay instead. For example, Mattar and Daw (2018) proposed that prioritization of which memories to replay is facilitated by SRPEs. More precisely, stimuli associated with large SRPEs are placed higher on the priority list. As a consequence, these highly prioritized stimuli are replayed more often and thus better remembered. Moreover, evidence suggests that hippocampal replay is sensitive to VTA signaling (Gruber

et al., 2016; Ólafsdóttir et al., 2018; Tompary et al., 2015). One possibility to investigate the importance of RPE-based replay at the behavioral level is by manipulating the subject's activity in the retention interval. Specifically, using our variable-choice paradigm, one could compare a condition where participants are subjected to a filler task in one condition, but no filler task in the other condition (similar to Dewar et al., 2014), who obtained a wakeful-rest versus filler task advantage for unintentionally studied words). This would allow explicitly testing whether cognitive processes occurring during off-line (but wakeful) periods (such as replay), boosts the RPE effect in declarative learning.

The current study has some limitations. First, although we found a main effect of agency on memory, it remains possible that the agency effect was driven by the fact that participants only had to press a button on agency trials and not on nonagency trials. In Yebra et al. (2019), action trials where participants made button presses, consistently led to better memory performance. A possible follow-up study would be to let participants also make a button press on nonagency trials after the computer has made its selection. Note, however, that this confound does not influence our primary effect of interest—namely, the interaction between agency and RPE. Indeed, if there is no interaction in the current design, there will presumably also not be an interaction in an experiment where the agency and nonagency conditions are even more tightly matched. Second, even though participants were aware that they would be tested on all 84 word pairs, they might have paid less attention to the word pairs presented in the nonagency condition. One way of objectively measuring this would be by using an eye tracker, to verify whether participants pay equal attention to word pairs presented in agency versus nonagency trials. Finally, another limitation of the current design is the limited range of RPEs that is probed. Specifically, RPEs ranged from minimally −.5 to maximally .75 (hence higher in absolute value). This asymmetry might have biased our results into finding an SRPE rather than a URPE pattern; indeed, a URPE effect has also been documented in the declarative learning literature (Rouhani et al., 2018).

In conclusion, the current results add to the growing body of evidence that RPE (independent of its source: stemming from the participant's own actions or coming from the states in a participant's environment) enhances declarative learning. In addition, we showed that the RPE effect cannot be solely explained by the number of eligible options. Introducing agency in the context of RPE and declarative learning, provides novel insights into declarative learning theory, with potential implications in applied psychology.

Appendix

Table 1   Stimulus material: 84 Dutch words

agent	appel	augurk	auteur	badjas
ballon	bezem	bloem	borst	brief
brood	cactus	cadeau	dakpan	dief
doos	draad	eend	eiland	emmer
ezel	fiets	gebouw	geit	gelaat
hamer	heuvel	insect	kaars	kaas
kajuit	kikker	kist	klok	knie
kogel	kooi	lamp	lawine	lepel
long	majoor	mand	naald	natuur
neus	oase	oceaan	olijf	oven
paard	parfum	peer	plank	poort
raam	riem	rijst	schaar	schip
sigaar	sjaal	slang	spade	spin
steen	ster	stoel	strand	tempel
tomaat	touw	traan	trein	varken
vinger	voet	vork	walvis	wiel
winkel	wolf	worm	zuil	

Table 2   Stimulus material: 336 Swahili words

adhabu	adui	afya	aibu	akili
almasi	amani	anga	ardhi	asili
askari	bafuni	bahari	baiski	barua
basi	bata	bega	bendi	buli
bullet	bunifu	chaki	chanzo	cheo
chombo	choori	chui	chuma	chumvi
chungu	chupi	chura	dafu	daima
dakika	dama	damu	daraja	darasa
dari	dizeli	dubu	duka	elfu
enzi	farasi	fedha	funzi	fupi
furaha	fuvu	gari	garisi	geza
godoro	goti	gugu	gundi	guruwe
habari	haki	hamsi	hasira	hatari
hatua	hawa	hazini	hekalu	hewa
hofu	homa	hukumu	ijayo	imani
ingi	ishara	ishiri	jangwa	jani
jansa	jeraha	jibini	jibu	jikoni
jino	jipu	jirani	jiwe	jokofu
joto	juisi	jumatu	juuya	kaburi
kahawa	kalamu	kamba	kamwe	katika
kaya	kazi	kelele	kemia	kesho
kiatu	kibodi	kichwa	kidole	kifua
kihozi	kijiko	kijivu	kikapu	kike
kilima	kimysa	kinywa	kioo	kisiwa
kisu	kitabu	kiti	kito	kitovu
kofia	koleo	kondoo	kovuli	kuacha
kubale	kubeba	kubwa	kuhesa	kuku
kumba	kununa	kunywa	kupima	kuruka
kuzama	kweli	leso	loma	lugha

Table 2   (continued)

adhabu	adui	afya	aibu	akili
mafuta	mageho	maisha	maji	mali
mamba	mapafu	mashua	maua	mavazi
maziwa	mbali	mbao	mbegu	mbolea
mboo	mbuzi	mbwa	mchawi	mchezo
mchuzi	mdomo	mdudu	mekno	meli
mende	meza	mfalme	mfano	mfuko
mfupa	mguu	miaka	minyoo	miwani
mizizi	mkanda	mkasi	mkate	mkoba
mkono	mkuu	mlango	moshi	moyo
mpira	mpishi	mraba	msitu	mtawa
mtirka	mtoto	mungu	mvua	mvuke
mwanga	mwezi	mwili	mwizi	mzungu
nafaka	nafasi	nakala	nanga	nchi
ndege	ndevu	ndizi	ndogo	ndoora
ndugu	neno	ngamia	ngao	ngazi
ngombe	ngome	ngono	ngozi	nguo
nopya	nyange	nyasi	nyeusi	nyoka
nyota	nyumba	nyundo	nyundu	nzuri
onyo	orodha	osha	paka	panya
pazia	petye	picha	piga	pigo
pole	pombe	pote	punda	pwani
pweza	radi	rafiki	rangi	rombus
ruka	sabuni	sahani	sakafu	samaki
sauti	sayari	sehemu	seri	shairi
shamba	shida	shingo	shule	siagi
simu	skrini	soko	somo	surali
suti	swali	tafuta	taifa	tamaa
tanuri	tatizo	tatu	tausi	tawi
tembo	trekta	twai	uadui	ufagio
ugomvi	uhuru	ukame	ukanda	ukimya
ukuta	ukweli	umeme	umri	unga
uongo	upepo	upya	usiku	uteo
uyoga	uzio	viatu	vifaa	vyeo
wadudu	wakala	wakati	washia	watoto
waya	wazo	welder	wema	wimbo
wingi	wingu	wizi	yake	yangu
yatima	ziwa			

Appendix B

Table 3   Mean Accuracy for each SRPE and Agency Condition

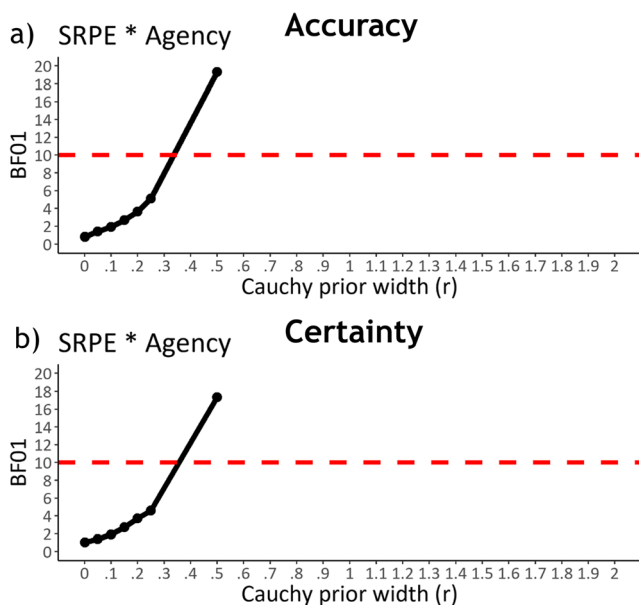|	Agency		Nonagency	
SRPE	*Mean*	*SD*	*Mean*	*SD*
-0.5	54.11%	25.05%	48.97%	24.86%
-0.25	61.78%	20.12%	51.06%	15.88%
0	57.44%	21.22%	52.75%	25.00%
0.5	68.97%	20.32%	56.97%	20.82%
0.75	72.75%	20.36%	58.28%	19.78%

Appendix C

Fig. 4 Bayes factor robustness checks. Bayes factor robustness checks with an adjusted *y*-axis scale ranging from 0 to 20, showing more clearly with which Cauchy prior width (r) the threshold for strong evidence (i.e., $BF_{01} > 10$) is crossed. **a** BF_{01} robustness check for the interaction effect of SRPE × Agency on recognition accuracy. **b** BF_{01} robustness check for the interaction effect of SRPE × Agency on certainty

Declarations

Conflict of interest None.

References

Butterfield, B., & Mangels, J. A. (2003). Neural correlates of error detection and correction in a semantic retrieval task. *Cognitive Brain Research*, *17*(3), 793–817. https://doi.org/10.1016/S0926-6410(03)00203-9

Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1491–1494. https://doi.org/10.1037/0278-7393.27.6.1491

Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning*, *1*(1), 69–84. https://doi.org/10.1007/s11409-006-6894-z

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. *Journal of Neuroscience, 41*(8) 1716–172. https://doi.org/10.1523/JNEUROSCI.1785-20.2020

Davidow, J. Y., Foerde, K., Galvan, A., & Shohamy, D. (2016). An upside to reward sensitivity: The hippocampus supports enhanced reinforcement learning in adolescence. *Neuron*, *92*(1), 93–99. https://doi.org/10.1016/j.neuron.2016.08.031

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), Article e0189212. https://doi.org/10.1371/journal.pone.0189212

Dewar, M., Alber, J., Cowan, N., & Sala, S. Della (2014). Boosting long-term memory via wakeful rest: Intentional rehearsal is not necessary, consolidation is sufficient. *PLOS ONE*, *9*(10), Article 0109542. https://doi.org/10.1371/journal.pone.0109542

Doll, B. B., Shohamy, D., & Daw, N. D. (2015). Multiple memory systems as substrates for multiple decision systems. *Neurobiology of Learning and Memory*, *117*, 4–13. https://doi.org/10.1016/j.nlm.2014.04.014

DuBrow, S., Eberts, E. A., & Murty, V. P. (2019). A common mechanism underlying choice's influence on preference and memory. *Psychonomic Bulletin and Review*, *26*(6), 1958–1966. https://doi.org/10.3758/s13423-019-01650-5

Eich, T. S., Stern, Y., & Metcalfe, J. (2013). The hypercorrection effect in younger and older adults. *Aging, Neuropsychology, and Cognition*, *20*(5), 511–521. https://doi.org/10.1080/13825585.2012.754399

Eichenbaum, H. (2004). Hippocampus: Cognitive processes and neural representations that underlie declarative memory. *Neuron*, *44*(1), 109–120. https://doi.org/10.1016/j.neuron.2004.08.028

Ergo, K., De Loof, E., Debra, G., Pastötter, B., & Verguts, T. (2020a). Failure to modulate reward prediction errors in declarative learning with theta (6 Hz) frequency transcranial alternating current stimulation. *PLOS ONE*, *15*(12), e0237829. https://doi.org/10.1371/journal.pone.0237829

Ergo, K., De Loof, E., Janssens, C., & Verguts, T. (2019). Oscillatory signatures of reward prediction errors in declarative learning. *NeuroImage*, *186*, 137–145. https://doi.org/10.1016/j.neuroimage.2018.10.083

Ergo, K., De Loof, E., & Verguts, T. (2020b). Reward prediction error and declarative memory. *Trends in Cognitive Sciences*, *24*(5), 388–397. https://doi.org/10.31234/OSF.IO/XNGWQ

Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review*, *16*(1), 88–92. https://doi.org/10.3758/PBR.16.1.88

Fujiwara, J., Usui, N., Park, S. Q., Williams, T., Iijima, T., Taira, M., Tsutsui, K. I., & Tobler, P. N. (2013). Value of freedom to choose encoded by the human brain. *Journal of Neurophysiology*, *110*(8), 1915–1929. https://doi.org/10.1152/jn.01057.2012

Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and episodic memory in humans and animals: An integrative framework. *Annual Review of Psychology*, *68*, 101–128. https://doi.org/10.1146/annurev-psych-122414-033625

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Gruber, M. J., Ritchey, M., Wang, S. F., Doss, M. K., & Ranganath, C. (2016). Post-learning hippocampal dynamics promote preferential retention of rewarding events. *Neuron*, *89*(5), 1110–1120. https://doi.org/10.1016/j.neuron.2016.01.017

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

JASP Team. (2020). *JASP (Version 0.13)[Computer software]*. https://jasp-stats.org/

Jeffreys, H. (1961). *The theory of probability* (3rd ed.). Oxford University Press.

Katzman, P. L., & Hartley, C. A. (2020). The value of choice facilitates subsequent memory across development. *Cognition*, *199*, 104239. https://doi.org/10.1016/j.cognition.2020.104239

Khamassi, M., & Girard, B. (2020). Modeling awake hippocampal reactivations with model-based bidirectional search. *Biological Cybernetics*, 114(2), 231–248. https://doi.org/10.1007/s00422-020-00817-x

Leotti, L. A., & Delgado, M. R. (2011). The inherent reward of choice. *Psychological Science*, 22(10), 1310–1318. https://doi.org/10.1177/0956797611417005

Leotti, L. A., & Delgado, M. R. (2014). The value of exercising control over monetary gains and losses. *Psychological Science*, 25(2), 596–604. https://doi.org/10.1177/0956797613514589

Leotti, L. A., Iyengar, S. S., & Ochsner, K. N. (2010). Born to choose: The origins and value of the need for control. *Trends in Cognitive Sciences,* 14(10), 457–463. https://doi.org/10.1016/j.tics.2010.08.001

Lisman, J., Grace, A. A., & Duzel, E. (2011). A neoHebbian framework for episodic memory: Role of dopamine-dependent late LTP. *Trends in Neurosciences*, 34(10), 536–547. https://doi.org/10.1016/j.tins.2011.07.006

Markant, D. B., Ruggeri, A., Gureckis, T. M., & Xu, F. (2016). Enhanced memory as a common effect of active learning. *Mind, Brain, and Education*, 10(3), 1–11.

Mattar, M. G., & Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature Neuroscience*, 21(11), 1609–1617. https://doi.org/10.1038/s41593-018-0232-z

Metcalfe, J., & Eich, T. S. (2019). Memory and truth: correcting errors with true feedback versus overwriting correct answers with errors. *Cognitive Research: Principles and Implications*, 4(1), 1–18. https://doi.org/10.1186/s41235-019-0153-8

Metcalfe, J., & Finn, B. (2012). Hypercorrection of high confidence errors in children. *Learning and Instruction*, 22(4), 253–261. https://doi.org/10.1016/j.learninstruc.2011.10.004

Momennejad, I., Otto, A. R., Daw, N. D., & Norman, K. A. (2018). Offline replay supports planning in human reinforcement learning. *ELife*, 7. https://doi.org/10.7554/eLife.32548

Montague, P., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *The Journal of Neuroscience*, 76(5), 1936–1947.

Murty, V. P., Dubrow, S., & Davachi, L. (2015). The simple act of choosing influences declarative memory. *The Journal of Neuroscience*, 35(16), 6255–6264. https://doi.org/10.1523/JNEUROSCI.4181-14.2015

Ólafsdóttir, H. F., Bush, D., & Barry, C. (2018). The role of hippocampal replay in memory and planning. *Current Biology,* 28(1), R37–R50. https://doi.org/10.1016/j.cub.2017.10.073

Pavlov, I. P. (1902). *The work of the digestive glands.* Charles Griffin and Company.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1/2), 8–13.

Pfeiffer, B. E. (2020). The content of hippocampal "replay" *Hippocampus*, 30(1), 6–18. https://doi.org/10.1002/hipo.22824

Poldrack, R. A., & Gabrieli, J. D. (1997). Functional anatomy of long-term memory. *Journal of Clinical Neurophysiology*, 14(4), 294–310.

R Core Team. (2014). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing. https://www.r-project.org/

Roscow, E. L., Jones, M. W., & Lepora, N. F. (2019). *Behavioural and computational evidence for memory consolidation biased by reward-prediction errors.* bioRxiv (p. 716290). https://doi.org/10.1101/716290

Rotem-Turchinski, N., Ramaty, A., & Mendelsohn, A. (2019). The opportunity to choose enhances long-term episodic memory. *Memory*, 27(4), 431–440. https://doi.org/10.1080/09658211.2018.1515317

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56, 356–374.

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/10.1101/111070

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

Ruggeri, A., Markant, D. B., Gureckis, T. M., Bretzke, M., & Xu, F. (2019). Memory enhancements from active control of learning emerge across development. *Cognition*, 186, 82–94. https://doi.org/10.1016/j.cognition.2019.01.010

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences*, 14(10), 464–472. https://doi.org/10.1016/j.tics.2010.08.002

Skaggs, W. E., & McNaughton, B. L. (1996). Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science*, 271(5257), 1870–1873. https://doi.org/10.1126/science.271.5257.1870

Skinner, B. (1990). *The behavior of organisms: An experimental analysis.* BF Skinner Foundation.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, 82(3), 171–177. https://doi.org/10.1016/j.nlm.2004.06.005

Squire, L. R. (2009). Memory and brain systems: 1969–2009. *Journal of Neuroscience*, 29(41), 12711–12716.

Squire, L. R., & Dede, A. J. O. (2015). Conscious and unconscious memory systems. *Cold Spring Harbor Perspectives in Biology*, 7(3), Article a021667. https://doi.org/10.1101/cshperspect.a021667

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction.* MIT Press.

Thorndike, E. (1932). *The Fundamentals of Learning.* New York: Teachers College Press.

Tompary, A., Duncan, K., & Davachi, L. (2015). Consolidation of associative and item memory is related to post-encoding functional connectivity between the ventral tegmental area and different medial temporal lobe subregions during an unrelated task. *Journal of Neuroscience*, 35(19), 7326–7331. https://doi.org/10.1523/JNEUROSCI.4816-14.2015

Wang, K. S., & Delgado, M. R. (2019). Corticostriatal circuits encode the subjective value of perceived control. *Cerebral Cortex*, 29(12), 5049–5060.

Watabe-Uchida, M., Eshel, N., & Uchida, N. (2017). Neural circuitry of reward prediction error. *Annual Review of Neuroscience*, 40(1), 373–394. https://doi.org/10.1146/annurev-neuro-072116-031109

Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, 265(5172), 676–679. https://doi.org/10.1126/science.8036517

Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H.-J., & Düzel, E. (2005). Reward-related fMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*, 45(3), 459–467. https://doi.org/10.1016/j.neuron.2005.01.010

Yebra, M., Galarza-Vallejo, A., Soto-Leon, V., Gonzalez-Rosa, J. J., de Berker, A. O., Bestmann, S., Oliviero, A., Kroes, M. C. W., & Strange, B. A. (2019). Action boosts episodic memory encoding in humans via engagement of a noradrenergic system. *Nature Communications*, 10(1), 1–12. https://doi.org/10.1038/s41467-019-11358-8