UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Prediction Errors Improve Declarative Learning but Not the Salience of Individual Words

Benada, Martin; van Opstal, F.

DOI
10.1525/collabra.164379

Publication date
2026

Document Version
Final published version

Published in
Collabra: Psychology

License
CC BY

Link to publication

Citation for published version (APA):
Benada, M., & van Opstal, F. (2026). Prediction Errors Improve Declarative Learning but Not the Salience of Individual Words. *Collabra: Psychology*, *12*(1), Article 164379. https://doi.org/10.1525/collabra.164379

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:13 Nov 2026

UNIVERSITY of CALIFORNIA PRESS | Collabra: Psychology

Benada, M., & Van Opstal, F. (2026). Prediction Errors Improve Declarative Learning but Not the Salience of Individual Words. *Collabra: Psychology*, *12*(1). https://doi.org/10.1525/collabra.164379

Cognitive Psychology

Prediction Errors Improve Declarative Learning but Not the Salience of Individual Words

Martin Benada[1], Filip Van Opstal[1][a]

[1] Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

Keywords: Reward Prediction Error, Declarative Learning, Memory, Recognition, Bias

https://doi.org/10.1525/collabra.164379

Collabra: Psychology

Vol. 12, Issue 1, 2026

Recent studies using the variable choice paradigm for a word-association learning task suggest that reward prediction errors (RPEs) enhance one-shot declarative learning. However, these studies may confound true associative learning with an "RPE bias," where RPE items become more salient or fluent and are therefore chosen during recognition. This study tested whether such a bias exists and whether RPEs genuinely strengthen word-pair learning. Across three sessions, participants (n = 48) completed a modified version of the variable choice paradigm in which English-Swahili word pairs were learned under either RPE or no-RPE conditions. Results provided moderate evidence against an RPE bias: RPE words were not more likely to be selected when incorrect. In contrast, correctness strongly predicted choice, and correct translations associated with RPEs were moderately more likely to be chosen than correct translations without RPEs, indicating a true learning benefit. Confidence analyses mirrored this pattern, showing strong evidence against an effect of RPE bias but extreme evidence for an effect of correctness. Furthermore, confidence was much higher for correct translations with RPE compared to correct translations without RPE. Together, these findings demonstrate that RPEs improve associative memory without producing choice biases. Our results validate previous conclusions about RPE-driven declarative learning benefits.

Introduction

One of the key mechanisms of learning are reward prediction errors (RPEs; Rescorla & Wagner, 1972). RPEs occur when an outcome is different from what a person expects. For example, when answering a question, if a person expects their answer to be wrong but is instead told their answer is right, that person will be positively surprised, leading to an RPE. According to the neoHebbian learning framework (Lisman et al., 2011), RPEs drive episodic learning by triggering dopamine bursts in the ventral tegmental area. These dopamine bursts are projected to the hippocampus, amplifying long term potentiation and thus improving memory. Recently, the relation between RPEs and learning was extended to declarative learning (De Loof et al., 2018; Ergo et al., 2020). De Loof and colleagues (2018) tested participants' ability to learn English-Swahili word-pairs. They found that participants were more likely to remember a word pair if it was associated with a high RPE.

De Loof et al. (2018) created a variable choice paradigm testing how RPEs affect foreign language learning. During the acquisition phase of the experiment, participants learned English-Swahili word pairs. Each English word was shown with either 1, 2, or 4 possible Swahili translations, and participants were asked to guess what they thought the correct translation was. More options meant participants had a lower chance of success. Therefore, a correct answer when there were four possible translations led to a high RPE, since their success was unexpected. During the recognition phase, participants were shown the previously learned English words alongside the same four Swahili words. The study found that English-Swahili word pairs with a high RPE during the acquisition phase were more likely to be remembered during the recall phase. As such, the study concluded that RPEs increase declarative learning.

However, rather than learning word-pairs in this variable choice paradigm, it is possible that people became biased towards choosing the high RPE Swahili words, without necessarily remembering the word-pair. Previous studies that used the variable choice paradigm gave the same options during the recognition phase as in the learning phase

[a] Corresponding author: Filip Van Opstal; Brain & Cognition, Department of Psychology; University of Amsterdam; Nieuwe Achtergracht 129, 1001NK, Amsterdam; The Netherlands
f.vanopstal@uva.nl

Downloaded from http://online.ucpress.edu/collabra/article-pdf/12/1/164379/942398/collabra_2026_12_1_164379.pdf by guest on 08 July 2026

(Calderon et al., 2021; De Loof et al., 2018; Ergo et al., 2019). This meant that there was always exactly one high RPE word, and it was the correct answer. This "RPE bias" would cause participants to prefer choosing high RPE words as translations, regardless of whether they were the correct answer. Such a bias would exist separate from any benefit RPEs give to association learning. High RPE words were never available as incorrect options, making it impossible to separate any true learning from "RPE bias".

Two heuristics suggest possible mechanisms for this "RPE bias". The first is the *recognition heuristic* (Goldstein & Gigerenzer, 2002). This heuristic is used when someone is answering a question and only recognizes one of the possible answers. The recognized item is assumed to be the correct response (Pachur et al., 2011). For this to apply, all non-salient Swahili words would have to be completely forgotten, leaving participants with only one recognized word (the high RPE word). If participants do not completely forget the alternate Swahili words, then another possible mechanism is the *fluency heuristic*. This heuristic says that if one stimulus is recognized faster than another, then that difference in time (i.e., the difference recognition fluency) can be used for recognition memory decisions (Hertwig et al., 2008; Jacoby & Brooks, 1984; Schooler & Hertwig, 2005). Whittlesea (1993) showed that fluency can be increased through the context in which a word is presented, suggesting RPEs could increase fluency (Whittlesea, 1993).

There are several ways in which RPEs could trigger these heuristics. One is that RPEs increase the attentional value of high RPE words. Anderson et al. (2013), for example, showed that the magnitude of value-driven attentional capture of a distractor was influenced by its recent RPE (Anderson et al., 2013). The Swahili words with high RPEs in the learning phase in de Loof et al. (2018) could thus retain their attentional value even when they are no longer being rewarded (i.e., once the words are presented during the recognition test). Individuals would be drawn to words with higher attentional value during the testing phase, which would bias their responses. It is also possible that individual high RPE Swahili words are more likely to be remembered. Several studies have shown that RPEs improve recognition memory for task-irrelevant images (Jang et al., 2019; Pupillo et al., 2023; Rouhani & Niv, 2021). Furthermore, Fazio and Marsh (2009) showed that RPEs can enhance task-irrelevant features of task-relevant stimuli (Fazio & Marsh, 2009). This is important given that the learning of individual words (our hypothesized "RPE bias") is task irrelevant but comes from a task-relevant feature (word pairs).

A final possible explanation is that words with high RPEs have increased hedonic value. Individuals would be drawn to these words and would mistakenly interpret this hedonic attraction as evidence that they are the correct translations. Rouhani and Niv (2021) showed that RPEs can enhance the hedonic value of images, influencing subsequent choice behavior. While this supports the idea that high RPE words have higher hedonic value, this study further showed that RPE can encode hedonic value separately from item recognition. A single RPE caused both hedonic value and recognition memory to be encoded (Rouhani & Niv, 2021). This suggests that a hedonic encoding mechanism in the variable choice paradigm can exist alongside a recognition encoding mechanism. It also suggests that true association learning can occur alongside an "RPE bias", since this would again be an encoding of two memory types through one RPE.

In the current study we wanted to investigate if declarative learning in the variable choice paradigm is due to true association learning (i.e., learning the association between a Swahili word and its English translation), an RPE bias (i.e., the recognition of only the English word because it was related to an RPE during learning), or a combination of both. We therefore examined whether individual words associated with an RPE become better memorized, and are thus more likely to be chosen in a word association test, even when it is not the correct answer. To test this, we performed a slightly modified version of the variable choice paradigm by de Loof et al. (2018).

An ongoing discussion about the relation between RPEs and learning concerns RPEs being signed or unsigned. Signed RPEs mean that RPEs for incorrect responses during learning are scored as negative values. De Loof et al. (2018) found that words with a negative RPE showed worse recall than words with no or a positive RPE. Although other studies also found this effect of signed RPEs on learning (e.g., Calderon et al., 2021; Davidow et al., 2016; Jang et al., 2019; Pupillo et al., 2023; Rouhani & Niv, 2021), others have shown an effect of unsigned RPEs, meaning that all RPEs led to better performance compared to having no RPE (e.g., Rosenbaum et al., 2022; Rouhani et al., 2018; Rouhani & Niv, 2021). This discussion is, however, beyond the scope of this study. As such, our study focuses exclusively on positive RPEs, which are the same under both the signed and unsigned hypothesis.

Experiment

Methods

Participants

48 participants (aged 18-30; 19 males, 38 females, 1 not disclosed) were recruited from the University of Amsterdam student population. This number was not determined a priori. Rather, we used Bayesian analysis and stopped collecting data once the evidence for or against the RPE bias surpassed a Bayes Factor of 3. This corresponds to an interpretation of "moderate evidence" (rather than anecdotal evidence) (Lee & Wagenmakers, 2014). Participants were not told the purpose of the experiment. Participants did not have prior knowledge of Swahili. All participants gave their informed consent prior to the experiment and were debriefed after the experiment finished. The study has been approved by the Faculty Ethics Review Board of the Faculty of Social and Behavioural Sciences of the University of Amsterdam (FMG-13072_2025). All procedures performed were in accordance with the relevant guidelines and regulations and with the Declaration of Helsinki. Participants received twenty euros for participating in the experiment.

Downloaded from http://online.ucpress.edu/collabra/article-pdf/12/1/164379/942398/collabra_2026_12_1_164379.pdf by guest on 08 July 2026

Apparatus and Stimuli

The experiment was run on Psychopy 2024.2.4 (Peirce et al., 2019). The experiment was coded using Python 3.10. Analyses were run using R 4.4.1. Two separate word sets were created for this experiment. The first set contained 1296 Swahili words taken from kamusi.org (Benjamin, 1995). All words were two syllables long. The second set contained 324 English words and were taken from the SUBTLEX US word frequency database (Brysbaert & New, 2009). All words were four syllables long, and had a Zipf value between 3.3 and 3.6, indicating medium frequency of occurrence (van Heuven et al., 2014). This meant that words were matched not only for length but also for how common they are. Four syllable English words were chosen to make memorization of word-pairs more difficult, so that the recognition task would remain difficult despite giving only two options per trial. Medium word frequency was chosen to help make the task more difficult while also having words be reasonably recognizable, unlike low frequency words which could be completely unknown to the participant. No plural words were included, as plural words would likely have a lower frequency (zipf value) than the singular version of the word, while still allowing participants to memorize the singular, higher frequency version of the word. The two word sets were created completely independently, meaning that words were chosen without considering whether there is a corresponding translation in the other word set. In all likelihood there are no true English translations of Swahili words contained within the pair of word sets.

Procedure

An experimental session consisted of three parts. In the first part, participants performed an acquisition task during which they had to learn the English translation of a Swahili word. This was followed by a filler task to reduce recency effects. In a final third part, learning performance was tested in a recognition task. Every participant completed three sessions to have enough data per participant to fit the statistical models (see *Repeat of Experiment* below).

Acquisition Task

This task was taken from de Loof et al. (2018) with some adaptations. At the start of each trial, the screen displayed one English word and four Swahili words. The words stayed on screen for four seconds, after which a frame appeared around the possible Swahili translations for the English word. Either one or four words were framed. The framed words indicated possible correct translations. In the one-option condition, participants were guaranteed to answer correctly. In the four-option condition participants had a 25% chance of choosing the correct translation. Each word position was assigned to one keyboard button (‘f’, ‘v’, ‘j’, or ‘n’). Participants were allowed to take as long as they wanted when deciding. Participants were told to answer based on their first impression.

Participants were then given feedback to indicate whether they had answered correctly. However, the feedback was secretly preprogrammed, so that a fixed number of trials would be rewarded. This meant that participants were not necessarily learning real translations of English words. For example, if a trial had been preprogrammed to have four options and be rewarded, then participants would receive positive feedback for their answer, regardless of which Swahili word they selected. This chosen word would become the word they had to memorize. This made sure we had enough of every option-reward combination (rewarded with one option, rewarded and unrewarded with four options). It also ensured that the expected probabilities aligned with actual probabilities, as 25% of four-option trials were rewarded. Furthermore, for each English word, the four Swahili words were randomly selected, and likely did not include the actual Swahili translation. This ensured that any linguistic regularity between the English and Swahili words could not affect participants’ learning.

The feedback given to participants included the English word, an equal sign and the (ostensibly) correct Swahili word displayed in the middle of the screen. During rewarded trials, a green frame appeared around the English word and the chosen Swahili word. During unrewarded trials, a red frame appeared around the English word and one of the unselected Swahili words. The words stayed on screen for five seconds. Participants were told to use this time to memorize the word pair being shown. After the feedback, participants received a 2.5 second presentation of the (informative, not monetary; see Chen, Claeys, et al., 2025) reward collected on that trial: When they had answered correctly, the message “Correct +1 point” was presented in green; if their response was incorrect, the message “False +0 point” was shown in red. In total, 108 trials were completed. 18 of those were rewarded 4-option trials, 54 were unrewarded 4-option trials, and 36 were rewarded 1-option trials. The reason there are only 18 rewarded 4-option trials compared to 54 unrewarded 4-options trials is that this maintains a 25% reward ratio, matching expected and actual reward probability. Only the rewarded 4-option trials are used during the testing phase. The reason there are 36 1-option trials rather than just 18 has to do with trial construction in the testing phase. Words with an RPE of 0 (i.e., 1-option words) are used as the distractor option of every trial, while also being used as the target option of some trials (see *Recognition Task*).

Every chosen Swahili word was now associated with an RPE classification. Rewarded words chosen out of 4 possible options were considered to be with RPE, and rewarded words out of 1 possible option considered to be without RPE. Unrewarded words were not relevant to the rest of the experiment and so were not given a classification. Figure 1 shows example trials (a rewarded 4-option and 1-option trial) of the acquisition task.

Filler Task

Before beginning the recognition task, participants completed a filler task to reduce recency effects. As in de Loof et al. (2018), a magnitude comparison task was used, in which participants categorized 400 numbers between 1 and 9 (not including 5). Participants pressed the ‘f’ button when the

Downloaded from http://online.ucpress.edu/collabra/article-pdf/12/1/164379/942398/collabra_2026_12_1_164379.pdf by guest on 08 July 2026

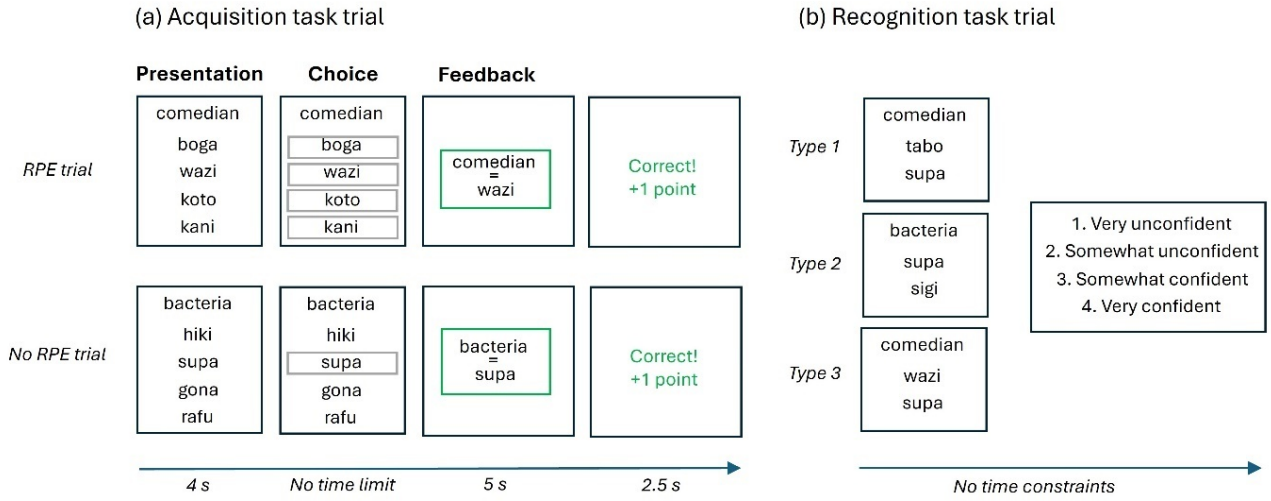

Figure 1. Design overview

Note. This figure shows examples of trials during the (a) acquisition and (b) recognition task. (a) In a high RPE trial, the four possible translations are marked (i.e., a 4-option trial); in a No RPE trial, only a single (correct) translation is marked (i.e., a 1-option trial). (b) In the recognition task, three trial Types are possible. In all trial Types, the distractor word is always a correct translation of another word from a 1-option trial. In Type 1 recognition trials, the target word is an RPE translation of another word (i.e., 'tabo' would have been a correct translation of another English word in a 4-option trial in the acquisition task). In Type 2 trials, the target word is the correct translation from a 1-option (no RPE) trial. In Type 3 trials, the target word is the correct translation from a 4-option (RPE) trial. Note that in the real experiment, none of the words (targets or distractors) were presented more than once in the recognition task.

number was smaller than 5, and the 'j' button when the number was larger than 5.

Recognition Task

Before the recognition task, participants were warned that some trials would not have a correct answer, though most trials would. For each trial, an English word appeared at the top of the screen along with two Swahili words below. As soon as the words appeared, participants could choose between the two Swahili words by using 'f' and 'j' to select the left and right word, respectively. No time constraints were given.

Each trial presented one distractor word and one target word. The distractor word was always a word that had been previously rewarded during a one-option trial (meaning it was without RPE). In total, there were three possible target words. In trial Type 1, the target word was with RPE (meaning that it was a rewarded answer among four options during the learning phase). However, it was not the correct translation as learned in the Acquisition Task. Therefore, in trial Type 1, there was no correct translation available. In trial Type 2, the target word was without RPE, and was the correct translation. In trial Type 3, the target word was with RPE. It was also the correct translation. Note that the trial type labels are purely to help describe the task. In practice the participant was not told about the concept of trial types, and the order that trials were presented was randomized. In total, 27 trials were presented (9 of each trial type). No feedback was given after recognition trials. Figure 1 displays examples of the different trial types.

After each trial, participants were shown a Likert scale asking how confident they were in their answer. The options included "very unconfident", "somewhat unconfident", "somewhat confident", "very confident". Participants answered by pressing a number from 1 to 4.

Repeat of Experiment

A single session of the experiment only gives 27 recognition trials. This is not enough to reliably estimate participant-level random effects, which we needed for data analysis. Simply increasing the number of trials per session might risk burnout, where participants would stop learning as effectively due to mental exhaustion. Therefore, we decided to have each participant complete three days of testing. Each day involved the same experiment, just with different words. The words for each day were randomly chosen during day one of each experiment. This meant that each participant would get a different set of words per day, so there was no connection between experiment day and words shown. Each word was used exactly once, so there were no repeating words and no unused words. In this way, each participant was shown every word in a random order. We therefore chose to combine data from all three days into a single overall dataset per participant. This allowed us to get 81 data points per participant. Simulations done before the experiment suggested this was enough to reliably estimate both fixed and random effects in the planned analysis. No random effects were included to distinguish what day data was collected on, as this would force our model to calculate an effect for each participant/day combination separately, which leads back to the issue of estimating random effects with only 27 trials.

Downloaded from http://online.ucpress.edu/collabra/article-pdf/12/1/164379/942398/collabra_2026_12_1_164379.pdf by guest on 08 July 2026

Data Analysis

Data Trimming

All trials in which participants indicated they were "strongly unconfident" in their answer were removed from the data. This is because strongly unconfident trials suggest that the participant did not remember the answer and simply guessed. Our study aims to look at memory, so guessed answers are not relevant to our research and only add noise to the results. Furthermore, some trials do not have a correct answer. If a participant does not (falsely) remember any answer and is forced to guess, then their choice is not a meaningful datapoint. By contrast, higher-confidence responses suggest that the participant at least believed in their choice, even if no correct option was available. This data trimming strategy differs from our preregistration. For details on the original strategy and our reasons for revising it, see osf.io/af9mx. To ensure that our results are not driven by this data trimming, we also repeated all choice analyses with a version of the data without any trials removed (see *Data Robustness* below).

Statistical Model

The primary outcome variable was a binary variable indicating whether the target word was selected on a given trial (chosen = 1; not chosen = 0). Trials were categorised into the three trial types as described before: in Type 1 trials, the target had high RPE but was incorrect; in Type 2 trials, the target was correct but had no RPE; and in Type 3 trials, the target was both correct and had a high RPE. The data was processed so that each trial corresponded to a single row of data, recording the trial type and whether the target was chosen. Code for the analysis can be found at osf.io/af9mx.

The data were analysed using Bayesian hierarchical modelling implemented in the brms R package (Bürkner, 2018). A key research question was whether the rate of target selection differed from chance (50%) within each trial type. Because the three trial types index qualitatively distinct conditions — differing in the presence of RPE, correctness, or both — each was modelled separately. Each model included only a fixed intercept and a random intercept. All models used a Bernoulli family to reflect the binary dependent variable.

The fixed intercept captured the overall log-odds of selecting the target, reflecting how strongly the defining characteristic(s) of each trial type shift choice probability away from chance level of 50%. The random intercept accounted for individual differences in this tendency. For example, in Type 2 trials, where the target is the correct answer, participants with stronger memory will select the target at much higher rates than those with weaker memory; the random intercept captures this variability. In Type 1 trials, the random intercept captures individual differences in the degree to which RPE biases choice, and in Type 3 trials, it captures variability in the combined influence of RPE and correctness across participants.

The fixed effects were assigned two-sided JZS Cauchy priors with a scale of $\sqrt{2}/2 \approx 0.707$. The JZS Cauchy is a widely used default prior in psychology that reflects minimal assumptions about effect size, enabling inference to be driven primarily by the data (Bayarri et al., 2012). Random intercept standard deviations were assigned one-sided positive JZS Cauchy priors. The restriction to positive values is because standard deviations cannot be negative.

Results

To assess how RPE biases choice probability in trials without a correct answer, we applied the Savage-Dickey density ratio method to the intercept of each trial-specific model. The Savage-Dickey ratio provides a principled Bayesian comparison of a parameter against a null hypothesis of no effect (Wagenmakers et al., 2010). It was computed using the BayestestR R package (Makowski et al., 2019). The Type 1 intercept yielded a BF_{10} of 0.21, equivalent to a Bayes Factor of 4.74 in favour of the null, providing moderate evidence that RPE does not meaningfully influence choice probability in Type 1 trials. We also looked at the parameter estimate of the intercept from the hierarchical model directly. Because the data was modelled using the Bernoulli family, the hierarchical model gave output in log odds form. To help interpretability, we converted all log odds parameters to odds ratios. The odds ratio form of the intercept represents the baseline odds of a typical participant choosing the target vs not choosing the target. For example, an odds ratio of 2.0 means that, within the trial type being tested, a typical participant was twice as likely to choose the target compared to the distractor. Because the random intercept accounts for between-participant differences, the intercept parameter can be interpreted as saying how the features unique to the target within that trial type increase odds of choice compared to not having that feature. For example, in trial Type 1, the odds ratio can be interpreted as how much a word having high RPE increases choice odds when both options are incorrect. The hierarchical model estimated the intercept to have an odds ratio of 1.11 (95% CI = [0.96, 1.28]), indicating that it is not different from 1. Participants were thus not more likely to chose the target word over the distractor in Type 1 trials.

Applying the same method to the intercept of the Type 2 model yielded a BF_{10} of 1.11e+06, constituting extreme evidence that correctness significantly influences choice probability. The hierarchical model estimated the intercept to have an odds ratio of 2.20 (95% CI = [1.77, 2.75]). Similarly, the Type 3 model intercept produced a BF_{10} of 1.12e+12, indicating extreme evidence for a combined effect of correctness and RPE on choice probability. The hierarchical model estimated the intercept to have an odds ratio of 2.86 (95% CI = [2.39, 3.49]).

However, these tests alone cannot determine whether, when the answer is correct, RPE contributes additional predictive value beyond correctness. To address this, we conducted a Bayesian paired *t*-test using the BayesFactor R package (Morey & Rouder, 2022), in which we compared the average proportion of correct choices between Type 2 and Type 3 trials. The analysis yielded a BF_{10} of 1.92, provid-

Downloaded from http://online.ucpress.edu/collabra/article-pdf/12/1/164379/942398/collabra_2026_12_1_164379.pdf by guest on 08 July 2026

Table 1. Bayes Factors from the Savage-Dickey ratio and posterior intercept estimates for all three model types under each prior

Model	Prior			
	Narrow		Wide	
	$BF_{direction}$	Intercept [95% CI]	$BF_{direction}$	Intercept [95% CI]
Type 1	BF_{01} = 3.38	1.11 [0.96, 1.27]	BF_{01} = 6.58	1.11 [0.96, 1.28]
Type 2	BF_{10} = 7.83e+05	2.18 [1.77, 2.72]	BF_{10} = 3.72e+06	2.20 [1.79, 2.77]
Type 3	BF_{10} = 9.23e+13	2.86 [2.39, 3.46]	BF_{10} = 9.52e+12	2.89 [2.41, 3.49]

ing anecdotal evidence in favour of a difference, with Type 3 trials showing a higher proportion of correct choices. The Cohen's d effect size was 0.32 (95% CI = [0.03, 0.61]).

It should be noted that this comparison cannot, in isolation, establish a causal learning benefit of RPE, as Type 3 trials remain confounded with RPE bias. However, the Type 1 model provides evidence against a systematic RPE bias, and if this null result holds, the observed difference between trial types can be tentatively interpreted as anecdotal evidence for a learning benefit given by RPE.

Prior Robustness

To assess sensitivity to prior specification, all analyses were also performed using both a wide Cauchy prior (scale = 1) and a narrow Cauchy prior (scale = 0.5). The direction and interpretation of evidence remained consistent across all prior specifications. For the Type 1 model, both priors yielded moderate evidence in favour of no difference from 50% chance and a 95% CI for the intercept (odds ratio) that contains 1. For the Type 2 and Type 3 models, both priors yielded extreme evidence in favour of a difference from 50% chance and a 95% CI for the intercept well above 1. Detailed results are reported in Table 1.

The Bayesian paired *t*-test comparing Type 2 and Type 3 choice rates was similarly robust: The wide Cauchy prior yielded a BF_{10} of 1.44 (Cohen's d = 0.33, 95% CI = [0.04, 0.61]), and the narrow Cauchy prior yielded a BF_{10} of 2.23 (Cohen's d = 0.30, 95% CI = [0.03, 0.59]), both consistent with the JZS result.

Data Robustness

As a further robustness check, all primary analyses were rerun on the untrimmed dataset, which retains trials in which participants reported a confidence rating of 1. The direction and interpretation of evidence remained consistent across all prior specifications. The Type 1 model yielded moderate evidence in favour of no difference from 50% chance (BF_{01} = 5.21). Intercept estimates in odds ratio form were 1.09 (95% CI = [0.96, 1.23]). The Type 2 model yielded extreme evidence in favour of a difference from 50% chance (BF_{10} = 3.91e+06). Intercept estimates in odds ratio form were 1.92 (95% CI = [1.60, 2.32]). The Type 3 model yielded extreme evidence in favour of a difference from 50% chance (BF_{10} = 1.38e+08). Intercept estimates in odds ratio form were 2.27 (95% CI = [1.92, 2.72]). All hierarchical model results are very similar to the results with the trimmed data.

Table 2. Observed choice behavior and Model estimates for every trial type.

Trial Type	% Chosen Target	
	Observed Mean	Model Mean [95% CI]
Type 1	52.51	52.52 [48.94, 56.10]
Type 2	67.63	68.67 [63.93, 73.33]
Type 3	73.69	74.12 [70.59, 77.66]

The Bayesian paired *t*-test on the untrimmed data, however, yielded a BF_{01} of 2.13 in favour of the null hypothesis of no difference between Type 2 and Type 3 choice rates. This contrasts with the anecdotal evidence for a difference found in the trimmed dataset. Taken together, the evidence regarding a difference in accuracy between trial types should be considered inconclusive. It is important to note that a Bayes Factor near 1 does not constitute evidence that no effect exists. Rather, it indicates that the current data are insufficient to distinguish between hypotheses and that further evidence is needed.

Choice Probability

Table 2 shows the observed pattern of choice behavior from the data. We also used the 95% CI of the intercept from each hierarchical model to estimate the mean and 95% CI of choice probability for each condition. As can be seen from Table 2, correct words are chosen significantly above chance, though the evidence for correct words with RPE to be chosen more often than correct words without RPE is inconclusive.

Confidence Analyses

To examine whether participants exhibited higher confidence on trials where they selected the target option (chosen = 1) compared to trials where they selected a non-target option (chosen = 0), we conducted a series of one-sided Bayesian paired *t*-tests, separately for each trial Type.

For each Type, confidence ratings were averaged across trials within each participant and choice condition (chosen = 1 and chosen = 0), yielding a pair of within-participant means per Type. T-tests were one sided to reflect the hypothesis that target options (which were correct and/or had

Downloaded from http://online.ucpress.edu/collabra/article-pdf/12/1/164379/942398/collabra_2026_12_1_164379.pdf by guest on 08 July 2026

higher RPE) would be chosen more frequently than non-target options. All Bayesian *t*-tests were conducted using the BayesFactor package (Morey & Rouder, 2022) in R. We once again used a JZS prior. For Type 1 trials this yielded a BF_{01} of 10.2 in favour of the null, indicating strong evidence that the RPE does not increase confidence in the absence of correctness. For Type 2 and Type 3 trials, the BF_{10} was 6.89e+07 and 1.07e+12, respectively, indicating extreme evidence that participants were more confident when choosing the correct answer. While this finding is trivial, it does support the validity of our experiment and the confidence measures.

To examine whether confidence ratings differed across stimulus Types, mean confidence was computed for each participant within each Type. Differences in confidence between Types were assessed using a series of two-sided Bayesian paired *t*-tests across all three pairwise comparisons. Two-sided tests were used here as no directional prediction was made regarding which Type would elicit higher confidence in the absence of choice information. All tests were conducted using the ttestBF function from the BayesFactor package (Morey & Rouder, 2022) in R, with a JZS Cauchy prior. A comparison between Type 1 and Type 2 gave a BF_{10} of 9.80e+04, indicating extreme evidence for a difference in average confidence. The average confidence values showed that Type 2 trials had higher confidence than Type 1 trials. A comparison between Type 1 and Type 3 also gave a BF_{10} of 3.19e+0, indicating extreme evidence for a difference in average confidence. The average confidence values showed that Type 3 trials had higher confidence than Type 1 trials. A comparison between Type 3 and Type 2 gave a BF_{10} of 14.8, indicating strong evidence for a difference in average confidence. The average confidence values showed that Type 3 trials had higher confidence than Type 2 trials. These findings strongly suggest that Type 3 trials yield the highest confidence, followed by Type 2 trials, followed by Type 1 trials.

To ensure that the difference between confidence in Type 2 vs Type 3 trials was not driven by confidence during incorrect answers, we repeated the above analysis using only data from trials where participants chose the correct answer. This gave a BF_{10} of 72.6, indicating very strong evidence in favour of a difference between groups.

Prior Robustness

To assess the sensitivity of our findings to the choice of prior, we repeated all *t*-tests under two additional Cauchy prior scales alongside the default medium prior ($r = \sqrt{2}/2 \approx 0.707$): A wide prior (r = 1.0), corresponding to Jeffreys's original standard Cauchy, and a narrow prior (r = 0.5) (Jeffreys, 1998).

For the one-sided *t*-tests comparing confidence on chosen versus unchosen trials, the BF_{01} = 14.35 in favour of the null hypothesis for trial Type 1, and BF_{10} = 7.52e+07, and 1.29e+12 for trial Types 2, and 3 respectively. Under the narrow prior BF_{01} = 7.41 in favour of the null for trial Type 1, and BF_{10} = 5.75e+07, and 8.32e+11 against the null for trial Types 2 and 3, respectively. For the pairwise two-sided comparisons of mean confidence across Types, the Bayes factors under the wide prior were BF_{10} = 9.81e+04, 3.37e+06, and 12.3 for the Type 1 vs. Type 2, Type 1 vs. Type 3, and Type 2 vs. Type 3 comparisons respectively, and under the narrow prior were BF_{10} = 8.8e+04, 2.73e+06, and 16.5. For the two-sided *t*-test comparing Type 2 and Type 3 confidence restricted to correct-answer trials, the Bayes factor under the wide prior was BF_{10} = 62.6 and under the narrow prior was BF_{10} = 76.7. In most cases, Bayes factors remained in the same evidential category as those obtained under the medium prior, indicating that our conclusions are robust to the choice of prior scale. The exception is in the *t*-test comparing confidence for target vs non-target options for Type 1 trials, as it dropped to moderate evidence under a narrow prior, though it remained well above the anecdotal threshold.

Discussion

In this study, we tested whether one shot declarative learning in the variable choice paradigm was caused by other effects of RPE. Specifically, we investigated whether RPEs can create a bias effect, so that words associated with high RPEs are more likely to be chosen in a later recognition test, even when they are not the correct answer. We used a modified version of the semantic memory task from the variable choice paradigm developed by de Loof et al. (2018). During the acquisition phase, participants were presented with an English word and either one or four possible Swahili translations. The combination of choice set size and response accuracy determined the RPE associated with each trial. After completing a filler task, participants performed a recognition test. In this phase, they were shown an English word along with two previously rewarded Swahili words and asked to select the correct translation. In some trials, the correct answer was associated with no RPE. In some trials, the correct answer was associated with a high RPE. In some trials, there was no correct answer, though one of the incorrect options was associated with a high RPE.

Our findings suggest that RPEs do not create a bias effect. Bayesian hierarchical modelling gave moderate evidence against the existence of a bias effect. Words associated with higher RPEs during the acquisition phase were not more likely to be chosen during the recognition phase when they were not the correct translation. Our findings also suggest that RPEs do improve association learning. We performed a Bayesian paired *t*-test between the correct options chosen when the correct option had a high RPE compared to when it had no RPE. This gave anecdotal evidence in favour of a difference between these trial types. Although this suggests only a relatively small RPE benefit in associative learning, results from the confidence analysis showed strong evidence that confidence was higher in Type 3 trials compared to Type 2 trials, especially when the target was correctly chosen. While this analysis cannot differentiate between confidence from learning and confidence from an RPE bias, the strong evidence for an absence of an effect of RPE on confidence in the absence of correctness (i.e., in Type 1 trials) suggests the RPE bias does not exist. We can therefore attribute any difference between correct options

Downloaded from http://online.ucpress.edu/collabra/article-pdf/12/1/164379/942398/collabra_2026_12_1_164379.pdf by guest on 08 July 2026

with RPE and correct options without RPE to an effect of RPE.

Similar to De Loof et al. (2018), we further verified if the effect of RPE could be explained by a 'time-on-task' account. According to that account, the 4-option trials would receive more elaborate processing and this would be reflected in a longer deliberation time per option when the Choice display (see Figure 1a) shows four options compared to only one. As in De Loof et al. (2018) we approximated the time devoted to each option on a trial by dividing the deliberation time (i.e., the response time on the Choice display) by the number of options. The resulting time-on-task per option ('time-on-option') was 1.31 s (SD = 1.02) when one option was presented compared to 0.443 s (SD = 1.48) per option when four options were presented. This result is in line with what was reported by De Loof et al. (2018) and would argue against more elaborate processing when four options are presented as recognition performance increased with the number of options while the time-on-word decreased when more options were available.

Our results suggest that the variable choice paradigm is a valid measure of the learning effect of RPEs. Individuals are not biased towards choosing specific words but instead use RPEs exclusively to learn word associations. These results therefore validate the findings reported in studies that used the variable choice paradigm (Calderon et al., 2021; De Loof et al., 2018; Ergo et al., 2019), or paradigms very similar to it to investigate the nature of the "testing effect" (Chen, Hauspie, et al., 2025; Chen, Verbeke, et al., 2025). Interestingly, a recent study presented a series of five experiments to test the robustness of the results obtained with the variable choice paradigm (Chen, Claeys, et al., 2025). In their Experiment 4, the distractors in the testing phase were randomly sampled from the full set of (correct) Swahili translations in the learning phase. Results from this experiment replicated earlier results, but the overall accuracy rate appeared to be smaller as compared to a similar experiment that used no Swahili translations as distractors (see their Experiment 3). Although this already provides some evidence for reliable word association learning with the variable choice paradigm, the current results provide strong evidence that RPEs do not systematically bias participants' responses in the recognition phase.

One limitation of our study is that participants completed three days of data collection. This design minimized fatigue and attention drop-off, which could impair memory for words learned later in a single session, but it reduced temporal specificity. While there is no theoretical reason to expect RPEs to affect learning differently across days, this possibility cannot be ruled out, and our study cannot separate RPE effects by day. This does means that our study cannot make claims about when the effect of RPE takes place. However, we do not feel that this is relevant to our research question. If an effect of RPE is found, but (unbeknownst to us) it only exists in the second and third day, this does not make conclusions about the effect of RPE invalid. Rather, it means that future studies are needed to determine how the effect of RPE interacts with time.

In conclusion, our study found that RPEs do not bias participants toward individual words with high RPEs, and was inconclusive on whether they affect participant choice when considering a correct answer. However, our study shows a strong confidence improvement associated with RPEs that was only present when the answer was correct, thus suggesting that RPEs do enhance learning at some level of awareness. These findings clarify a potential confound in earlier research using the variable choice paradigm and therefore give additional validation for those studies' conclusions.

Author Contributions

Contributed to conception and design: MB, FVO

Contributed to acquisition of data: MB

Contributed to analysis and interpretation of data: MB, FVO

Drafted and/or revised the article: MB, FVO

Approved the submitted version for publication: MB, FVO

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Funding

The authors did not receive support from any organization for the submitted work.

Ethics Statement

Approval was obtained from the Faculty Ethics Review Board of the Faculty of Social and Behavioural Sciences of the University of Amsterdam (FMG-13072_2025). The procedures used in this study adhere to the tenets of the Declaration of Helsinki. Informed consent was obtained from all individual participants included in the study.

Data Accessibility Statement

The datasets and code generated during the current study are publicly available. The study was preregistered on June 12th 2025. Deviations from the preregistration are fully described in a separate document. Everything can be found in the OSF repository, osf.io/af9mx.

Editors: Sebastiaan Mathôt (Senior Editor)

Submitted: May 19, 2026 PDT. Accepted: June 22, 2026 PDT.

Published: July 08, 2026 PDT.

Downloaded from http://online.ucpress.edu/collabra/article-pdf/12/1/164379/942398/collabra_2026_12_1_164379.pdf by guest on 08 July 2026

Downloaded from http://online.ucpress.edu/collabra/article-pdf/12/1/164379/942398/collabra_2026_12_1_164379.pdf by guest on 08 July 2026

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at http://creativecommons.org/licenses/by/4.0 and legal code at http://creativecommons.org/licenses/by/4.0/legalcode for more information.

References

Anderson, B. A., Laurent, P. A., & Yantis, S. (2013). Reward predictions bias attentional selection. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00262

Bayarri, M. J., Berger, J. O., Forte, A., & García-Donato, G. (2012). Criteria for bayesian model choice with application to variable selection. *The Annals of Statistics*, *40*(3). https://doi.org/10.1214/12-aos1013

Benjamin, M. (1995). *Kamusi Gold (global online living dictionary)*. https://kamusi.org/

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017

Calderon, C. B., Loof, E. D., 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), 1716–1726. https://doi.org/10.1523/JNEUROSCI.1785-20.2020

Chen, H., Claeys, W., Ergo, K., Calderon, C. B., & Verguts, T. (2025). Robust Effects of Signed Prediction Error on Declarative Memory (Sqjgz_v3). *PsyArXiv*. https://doi.org/10.31234/osf.io/sqjgz_v3

Chen, H., Hauspie, C., Ergo, K., Buc Calderon, C., & Verguts, T. (2025). Predictive learning as the basis of the testing effect. *Communications Psychology*, *3*(1), 18. https://doi.org/10.1038/s44271-025-00200-1

Chen, H., Verbeke, P., Mattioni, S., Calderon, C. B., & Verguts, T. (2025). Neural and computational evidence for a predictive learning account of the testing effect. *Proceedings of the National Academy of Sciences*, *122*(32), e2506530122. https://doi.org/10.1073/pnas.2506530122

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

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. (2020). Reward Prediction Error and Declarative Memory. *Trends in Cognitive Sciences*, *24*(5), 388–397. https://doi.org/10.1016/j.tics.2020.02.009

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

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*(1), 75–90. https://doi.org/10.1037/0033-295X.109.1.75

Hertwig, R., Herzog, S. M., Schooler, L. J., & Reimer, T. (2008). Fluency heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(5), 1191–1206. https://doi.org/10.1037/a0013025

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In *Psychology of learning and motivation* (Vol. 18, pp. 1–47). Elsevier. https://doi.org/10.1016/S0079-7421(08)60358-8

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

Jeffreys, H. (1998). *Theory of Probability*. Oxford University Press. https://doi.org/10.1093/oso/9780198503682.001.0001

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

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

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Morey, R. D., & Rouder, J. N. (2022). *BayesFactor Computation of Bayes Factors for Common Designs* [Computer software]. https://cran.r-project.org/package=BayesFactor

Pachur, T., Todd, P. M., Gigerenzer, G., Schooler, L., & Goldstein, D. G. (2011). The Recognition Heuristic: A Review of Theory and Tests. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00147

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Downloaded from http://online.ucpress.edu/collabra/article-pdf/12/1/164379/942398/collabra_2026_12_1_164379.pdf by guest on 08 July 2026

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 Science of Learning*, *8*(1), 18. https://doi.org/10.1038/s41539-023-00166-x

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In *Classical conditioning, Current research and theory* (Vol. 2, pp. 64–99). Appleton-Century-Croft. https://cir.nii.ac.jp/crid/1572543025504096640

Rosenbaum, G. M., Grassie, H. L., & Hartley, C. A. (2022). Valence biases in reinforcement learning shift across adolescence and modulate subsequent memory. *eLife*, *11*, e64620. https://doi.org/10.7554/eLife.64620

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

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

Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, *112*(3), 610–628. https://doi.org/10.1037/0033-295X.112.3.610

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cognitive Psychology*, *60*(3), 158–189. https://doi.org/10.1016/j.cogpsych.2009.12.001

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(6), 1235–1253. https://doi.org/10.1037/0278-7393.19.6.1235

Downloaded from http://online.ucpress.edu/collabra/article-pdf/12/1/164379/942398/collabra_2026_12_1_164379.pdf by guest on 08 July 2026