



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Sewell, DK;Warren, HA;Rosenblatt, D;Lyons, M;Bode, S

Title:

Feedback discounting in probabilistic categorization: Converging evidence from EEG and cognitive modelling

Date:

2018

Citation:

Sewell, D. K., Warren, H. A., Rosenblatt, D., Lyons, M. & Bode, S. (2018). Feedback discounting in probabilistic categorization: Converging evidence from EEG and cognitive modelling. *Computational Brain & Behavior*, 1 (2), pp.165-183. <https://doi.org/10.1007/s42113-018-0012-6>.

Persistent Link:

<https://hdl.handle.net/11343/300040>



Feedback Discounting in Probabilistic Categorization: Converging Evidence from EEG and Cognitive Modeling

David K. Sewell^{1,2}  · Hayley A. Warren² · Daniel Rosenblatt² · Daniel Bennett² · Maxwell Lyons¹ · Stefan Bode²

© Springer Nature Switzerland AG 2018

Abstract

In simple probabilistic learning environments, the informational value of corrective feedback gradually declines over time. This is because prediction errors persist despite learners acquiring the contingencies between stimuli and outcomes. An adaptive solution to the problem of unavoidable prediction error is to discount feedback from the learning environment. We provide novel neural evidence of feedback discounting using a combination of behavioral modeling and electroencephalography (EEG). Participants completed a probabilistic categorization task while EEG activity was recorded. We used a model-based analysis of choice behavior to identify individuals that did and did not discount feedback. We then contrasted changes in the feedback-related negativity (FRN) for these two groups. For individuals who did not discount feedback, we observed learning-related reductions in the FRN that reflected incremental changes in choice behavior. By contrast, for individuals who discounted feedback, we found that the FRN was effectively eliminated due to the rapid onset of feedback discounting. The use of a feedback discounting strategy was linked to superior performance on the task, highlighting the adaptive nature of discounting when trial-to-trial outcomes are variable, but the long-term contingencies relating cues and outcomes are stable.

Keywords Feedback discounting · EEG · Feedback-related negativity · Category learning · Computational modeling

Many of the decisions we face in daily life are based on uncertain information, when cues that are available at the time of decision-making are only imperfect predictors of possible outcomes. For example, dark clouds on the horizon may predict rain 70% of the time, or an itch may predict the presence of a skin rash 10% of the time. In these kinds of *probabilistic learning environments*, people are adept at tracking the probabilistic relationship between cues and their respective outcomes (Castellan 1973; Edwards 1961; Lagnado et al. 2006; see Little and Lewandowsky 2012, for a brief review). It is widely held that learning in probabilistic environments—just like learning in deterministic environments, where cues are perfect predictors of outcomes—is driven by prediction error generated by corrective feedback (Kruschke and Johansen 1999). However, in probabilistic learning environments, corrective feedback has diminishing utility, as error-free

performance is not possible, even under an optimal decision strategy. One way to adapt to the continuing presence of prediction error is to progressively ignore, or discount, feedback as one gains more experience with the learning environment and the contingencies relating stimuli to outcomes. To date, however, the evidence for feedback discounting has been surprisingly sparse with only a handful of studies convincingly demonstrating the need to assume a discounting mechanism (e.g., Craig et al. 2011).

In this article, we seek to provide novel support for feedback discounting by relating a model-based analysis of the learning process to changes in an electrophysiological correlate of feedback processing, the *feedback-related negativity* (FRN; Miltner et al. 1997). The FRN is an event-related potential (ERP) that arises approximately 200–400 ms following onset of corrective feedback, and is characterized by a negative deflection in the activity of the electroencephalogram (EEG) over fronto-central scalp regions. The FRN has been linked theoretically to error-detection processes involved in learning (Holroyd and Coles 2002; for recent reviews, see Luft 2014; Walsh and Anderson 2012). If feedback discounting is indeed a defining property of associative learning, changes in the FRN should reflect the degree to which individuals discount corrective feedback. Specifically,

✉ David K. Sewell
d.sewell@uq.edu.au

¹ School of Psychology, The University of Queensland, St. Lucia, QLD 4072, Australia

² Melbourne School of Psychological Sciences, The University of Melbourne, Parkville 3010, VIC, Australia

learners who progressively discount feedback should show greater reductions in the FRN over time, compared to individuals who do not discount feedback. One of the complicating factors in investigating feedback discounting in this way, however, is that changes in the FRN can only be attributed to feedback discounting once the naturally occurring reductions in prediction error—due to learning—are already accounted for. We argue that this necessitates application of cognitive models to establish the presence of feedback discounting before examining changes in the FRN over time.

We investigated feedback discounting by having participants complete a probabilistic cue-outcome learning task, where their goal was to predict the category outcomes associated with a set of simple perceptual stimuli. The FRN was measured on each trial via continuous EEG recording. Feedback discounting was assessed behaviorally at the individual participant level via cognitive modeling of choice data. Our model-based analysis revealed individual differences in whether people discounted corrective feedback during the task. We investigated changes in the evolution of the FRN for people who showed evidence of feedback discounting compared to those who did not. By operationalizing feedback discounting behaviorally via modeling, we were able to relate changes in the FRN over time to individual differences in the cognitive mechanisms underlying feedback processing. Our principal finding is that people who did not discount feedback produced a stronger FRN response compared to those who discounted feedback. Indeed, we found that feedback discounting effectively abolished the FRN. These results are striking in light of the differences in the behavioral choice profiles between the two groups of individuals: Participants who discounted feedback produced response profiles that more closely approached optimal responding compared to those who did not discount feedback. Put another way, participants who consistently relied on feedback over the course of the experiment tended to perform *worse* than those who progressively discounted, or ignored, corrective feedback.

We proceed by summarizing the behavioral evidence for feedback discounting during probabilistic category learning. We then describe the evidence linking the FRN to error processing as well as findings that have identified attenuation of the FRN over the course of learning. We note that while changes in the FRN are consistent with the notion of feedback discounting, to date, no study has accounted explicitly for discounting when relating the FRN to learning performance.

Feedback Discounting in Probabilistic Categorization: Behavioral Evidence

In probabilistic learning environments, where contingencies relating cues to category outcomes are uncertain, there is a question about how extensively learners should process

corrective feedback, and whether the degree of processing should change over the course of learning. Kruschke and Johansen (1999) proposed that the unavoidable nature of prediction error in probabilistic environments should encourage reduced reliance on corrective feedback once performance has stabilized and cue-outcome contingencies have been learned. During the early stages of learning, people should rely heavily on feedback so that they can rapidly reduce prediction error. However, the probabilistic nature of the cue-outcome relationships imposes a strict upper limit on accuracy, and so there is little reason to process feedback during the later stages of learning, as the effort spent in doing so will no longer improve accuracy. It follows that people should become less reliant on feedback as they become more experienced with the task. In the extreme, corrective feedback may be ignored entirely, resulting in no changes in associations relating cues to outcomes, regardless of whether the learner's response was correct or incorrect.

Feedback discounting is an appealing way to minimize cognitive effort while achieving accurate performance in probabilistic environments; however, the existing support for such a discounting mechanism remains sparse. Edgell (1983) used a probabilistic categorization task incorporating cues that were individually, but imperfectly, predictive of outcomes. Additional predictive information, in the form of *configurations* of cues, was introduced after a variable delay (e.g., from the outset of the task, or after 40 to 200 learning trials). Edgell found that learners underutilized the configural cue information when it was introduced after a delay compared to when predictive configural information was available from the outset, reflecting a reduced capacity to learn about newly available predictive information. Kruschke and Johansen (1999) subsequently showed, via computational modeling, that feedback discounting was sufficient to account for the pattern of results reported by Edgell. A limitation of the feedback discounting interpretation of Edgell's results is that they are equally compatible with associative blocking (Kamin 1969), which can occur independently of other processes such as feedback discounting (Rescorla and Wagner 1972; Sewell and Lewandowsky 2012). Under blocking, a previously learned association between a cue, *A*, and an outcome, *X*, prevents learning of an association between a second cue, *B*, and the same outcome, when *A* and *B* are paired during learning. In the case of Edgell's study, early learning of the relationships between the individually predictive cues and the outcomes could have blocked subsequent learning of the configural information, as the configural information was defined by combinations of individually predictive cues. Using a more elaborate experimental design, Blair and Homa (2005) showed that underutilization of predictive information introduced part-way through learning was at least partly caused by blocking, though feedback discounting also played a role in driving the effect.

Other studies provide stronger support for feedback discounting. For example, Edgell and Morrissey (1987; Experiments 4–5) showed systematic underutilization of predictive configural information even when no cues were initially predictive of category outcomes, eliminating blocking as a viable explanation. More recently, Craig et al. (2011) investigated probabilistic categorization using a set of unidimensional stimuli—another way of ruling out blocking. Each stimulus had a unique feedback probability, and Craig et al.’s principal manipulation was the number of trials before the feedback probabilities were reversed (e.g., an item associated with category A 80% of the time suddenly became associated with category B 80% of the time). Craig et al. fit a variety of learning models to their data and found that, regardless of the representational assumptions of the models, feedback discounting improved fits to data.

Taken together, the studies reviewed above suggest that feedback discounting plays an important role in learning in probabilistic environments. However, stronger support for feedback discounting could be obtained by linking model-based behavioral evidence of discounting with relevant neural data (Forstmann et al. 2011; Love and Gureckis 2007). One of the best candidates for investigating feedback discounting processes in learning is the FRN (Miltner et al. 1997; for reviews, see Luft 2014; Nieuwenhuis et al. 2004a; Walsh and Anderson 2012), which serves as a neural assay of feedback processing. We now provide a brief overview of studies linking EEG data to feedback processing in simple learning environments.

Electrophysiological Correlates of Learning

Over the last several decades, there has been a burgeoning literature investigating the neural correlates of learning. A key early finding was the differential neural response to correct versus error decisions in choice response time tasks with high accuracy (Falkenstein et al. 1991; Gehring et al. 1993). Approximately 80–100 ms after an error response is made, there is a negative deflection in EEG activity measured over fronto-central scalp regions. Because this deflection is reliably stronger following error responses than correct responses, it was termed the *error-related negativity* (ERN) by Gehring et al. The ERN principally manifests when stimulus–outcome contingencies have been clearly established by instruction, as in simple RT tasks, or by earlier learning, as in probabilistic choice tasks. Because the ERN is typically observed when people do not require overt feedback about whether they have committed an error, it has been interpreted as reflecting the operation of an error-monitoring system that responds when outcomes deviate from expectations (e.g., Holroyd and Coles 2002; Stahl 2010). Subsequently, Miltner et al. (1997) identified an electrophysiological correlate of

receiving error *feedback*, which had a similar form and spatial distribution to the ERN, but peaked approximately 200–400 ms after onset of feedback rather than the response. This deflection was termed the *feedback-related negativity* (FRN). Due to the similarities between the ERN and FRN, it has been argued that the FRN reflects the operation of the same error-monitoring system that has been implicated by the ERN (e.g., Holroyd and Coles 2002; Miltner et al. 1997; Stahl 2010).

The Feedback-Related Negativity

The FRN has received particular interest, as it is elicited by feedback signaling negative task outcomes, such as monetary losses (e.g., Gehring and Willoughby 2002), prediction errors (e.g., Luque et al. 2012), and obtaining poorer reward outcomes relative to other possible choice options (e.g., Nieuwenhuis et al. 2004b). The broad range of circumstances under which the FRN is observed strongly suggests that it reflects domain-general processing of outcome feedback. Although there remains controversy about whether the FRN reflects a binary good–bad evaluation of an outcome, as opposed to a continuous measure of prediction error (e.g., Hajcak et al. 2006; Yeung and Sanfey 2004; but see Sambrook and Goslin 2015 for a recent meta-analysis), there is general consensus that the FRN captures important properties of outcome feedback that can be used to promote learning and improve task performance (Cohen et al. 2011; Luft 2014; Walsh and Anderson 2012). Indeed, one of the properties of the FRN that makes it especially attractive in relation to studying feedback discounting is that its amplitude¹ has been shown to be modulated by learning itself (e.g., Bellebaum and Daum 2008; Cohen et al. 2007; Eppinger et al. 2008; Krigolson et al. 2009; Pietschmann et al. 2008; Sailer et al. 2010). We now discuss some of the key findings examining learning-related changes in the FRN.

Learning and the FRN Changes in the FRN over the course of a task are thought to reflect people’s reduced reliance on external feedback in monitoring their performance. In the early part of a task, when contingencies relating different stimuli to outcomes are not yet known, performance must be guided exclusively by corrective feedback. Once the contingencies have been learned, however, people should become less reliant on feedback, as they are able to monitor their own performance based on prior experience (i.e., learned associations). Numerous studies that have examined the FRN before and after learning have shown systematically reduced FRN amplitudes in the latter parts of learning tasks (Bellebaum and Daum 2008; Eppinger et al. 2008; Krigolson et al. 2009;

¹ Throughout this article, we follow convention and use the term “amplitude” as a shorthand to describe the signed voltage difference relative to some reference level.

Pietschmann et al. 2008). That these changes in the FRN are more pronounced on trials where outcomes match people's expectations—rather than violate them—strongly suggests that reductions in the FRN reflect learning (Bellebaum and Daum 2008; Hajcak et al. 2007). Importantly, surprising events also elicit an enhanced FRN, reflecting processing of outcomes that violate expectations established by prior learning. For example, Cohen et al. (2007) used a probabilistic reward prediction task, and found that, late in the task, the FRN for reward outcomes was enhanced for stimuli that typically failed to predict reward. It is also notable that learning-related reductions in FRN amplitude appear to be restricted to stimuli that are predictive of outcomes, and that no modulation is observed for non-predictive stimuli (Eppinger et al. 2008; Krigolson et al. 2009). A particularly striking demonstration of how learned expectations relate to the FRN was provided by Luque et al. (2012) who utilized a blocking paradigm, showing that, while FRN amplitude reduced over the course of learning, the FRN was consistently larger for predictive cues compared to blocked cues. Luque et al.'s results reinforce the idea that established predicted relationships can “block” learning about other equally predictive cues that are introduced later in the task. That the FRN was diminished for a blocked cue supports the idea that the FRN reflects the extent to which people process feedback information. Reductions in FRN amplitude over time imply the development of established expectations about cue-outcome contingencies. The link between “quality” of learning (i.e., the extent to which the learner processes feedback information) is further supported by a close relationship between rapid reductions in FRN amplitude on the one hand, and superior learning outcomes on the other (Arbel and Wu 2016).

Converging neural data also suggest that reductions in FRN amplitude reflect establishing learned expectations about task contingencies. For example, several studies have shown an inverse relationship between the amplitudes of the ERN and the FRN over the course of a task. Whereas the FRN tends to be larger than the ERN in the early learning phases of a task, the opposite is true in the later parts of the task. This pattern is indicative of shifting reliance from external sources of feedback to guide performance to monitoring performance via internally generated predictions (e.g., Bellebaum and Daum 2008; Eppinger et al. 2008; Holroyd and Coles 2002; Krigolson et al. 2009; Pietschmann et al. 2008).

More fine-grained evidence that changes in the FRN reflect learned associations has been reported via model-based analyses. For example, Chase et al. (2011) characterized behavioral choice data from a probabilistic reversal learning task using a simple reinforcement learning model. They showed that the trial-by-trial variation in the amplitude of the FRN was correlated with the model-derived prediction errors. Walsh and Anderson (2011) showed a similar correspondence between model-based prediction errors and FRN

amplitude. Learning-related changes in the FRN have also been simultaneously related to model-based estimates of prediction error as well as subjective estimates of reward probability (Ichikawa et al. 2010).

Feedback Discounting and the FRN The studies reviewed above show that there is substantial evidence that the amplitude of the FRN reduces over the course of learning. While these changes are consistent with the idea that people come to ignore feedback that is encountered after contingencies between stimuli and outcomes have been established—and are therefore suggestive of the presence of feedback discounting—they are also consistent with the fact that learning entails a reduction in prediction error. To the extent that the FRN is an index of prediction error, the changes in the FRN reviewed above are expected on the basis of existing theory (e.g., Holroyd and Coles 2002) without the need to posit an additional discounting mechanism. A critical question, then, is whether there is evidence that feedback discounting might serve to reduce the FRN over and above what can be attributed to reductions in prediction error due to learning. As is the case in the behavioral literature, direct evidence for feedback discounting in the electrophysiology literature is sparse.

Preliminary support for feedback discounting relating to changes in the FRN was reported by Bland and Schaefer (2011). They investigated performance in a probabilistic rule reversal task, where the stability of the learning environment—operationalized by the mappings of stimuli to response alternatives—varied across different blocks of trials. By manipulating the frequency with which rule reversals occurred, Bland and Schaefer were able to control how informative trial-by-trial feedback was to learners. When the environment is stable, unexpected prediction errors can be explained away by attributing them to the probabilistic nature of the feedback, and need not be interpreted as signaling a rule reversal. By contrast, when the learning environment is volatile, and rule shifts occur frequently, unexpected prediction errors can be highly informative, as they potentially indicate a rule reversal. Bland and Schaefer found that the FRN was diminished when the feedback environment was stable, compared to volatile. Although suggestive of feedback discounting, this result could also reflect more frequent prediction errors due to the presence of rule reversals in the more volatile learning environment.

Compelling support for feedback discounting comes from a very recent study by Schiffer et al. (2017), who showed that *expectations* about how stable or volatile the learning environment is can modulate the FRN. These authors focused on performance on pre-reversal trials in a probabilistic rule reversal task. Critically, people's expectations about the environment were manipulated by advising learners that a rule reversal was either likely or unlikely

to occur at some point during an upcoming block of trials. By manipulating people's expectations about the probability of a rule reversal, rather than the objective probability, the FRN during the early stages of learning (i.e., before a rule reversal, if present, occurred) could be investigated without being confounded with prediction error. The key result, with regard to feedback discounting, was that the FRN was indeed reduced when people expected the learning environment to be stable, rather than volatile, even though the initial learning environment was identical. This result strongly suggests that the degree to which people process feedback can vary, depending on expectations about the learning environment, indicative of a discounting mechanism.

Current Study

In this study, we adopt an individual difference approach to investigate feedback discounting. Here, we ask whether there are differences in the way the FRN changes over the course of learning for people who do and do not discount feedback. We present people with a probabilistic cue-outcome learning task, where the feedback contingencies are held constant over the course of the entire experiment. To identify people who utilize feedback discounting, we fit different versions of a popular category learning model (RASHNL; Kruschke and Johansen 1999) to each participants' behavioral data; one that incorporates a feedback discounting mechanism, and one that does not. By comparing the fits of these models to the behavioral data, we are able to identify feedback discounting independent of the FRN. We then contrast the changes in the FRN data for people who were identified as discounting and not discounting feedback. We hypothesize that individuals who discount feedback will show a reduced FRN compared to those who do not discount feedback. Such a result would be consistent with the idea that feedback discounting contributes to changes in the FRN over and above what would be expected on the basis of reduced prediction error alone.

Method

Participants

Forty participants were recruited via advertisements posted around the University of Melbourne as well as through social networking sites, and were remunerated A\$20 to cover travel costs. All participants reported normal or corrected-to-normal visual acuity, and no history of psychotropic medication or previous neurological disorders. The study was approved by the Human Research Ethics Committee at the University of Melbourne and was conducted in accordance with the Declaration of Helsinki.

Apparatus and Stimuli

The four stimuli used in the current study, adapted from Craig et al. (2011), were circles of fixed radius that varied along a single quasi-continuous dimension: the orientation of a radial line. Depending on the stimulus, the radial line could take on an orientation of -54° , -18° , 18° , or 54° , relative to vertical. On each trial, participants were shown a single stimulus and were required to categorize it as belonging to category A or category B. Following Craig et al., feedback probabilities assigned to the stimuli systematically increased or decreased as the orientation of the line moved from left to right, and participants were randomly assigned to either the ascending ($N=21$) or descending ($N=19$) probability conditions. For the ascending condition, feedback probabilities for the -54° , -18° , 18° , and 54° stimuli were set to $P(A)=0.2, 0.4, 0.6$, and 0.8 . For the descending condition, the respective feedback probabilities were $P(A)=0.8, 0.6, 0.4$, and 0.2 . Feedback regarding the correct classification was provided at the end of each trial.

The experiment was controlled on a Windows-based computer running a MATLAB program developed using the Psychophysics toolbox (Brainard 1997; Pelli 1997), which presented all stimuli and recorded all responses. Stimuli were displayed on a 19-in. Sony Trinitron G420 CRT monitor at a resolution of 1600×1200 pixels and frame rate of 120 Hz. Responses were recorded using the *F* and *J* keys on a standard computer keyboard. Participants were seated approximately 60 cm from the display monitor with viewing position stabilized using a chinrest.

EEG Recording

EEG was recorded continuously at 512 Hz from 64 active Ag/AgCl electrodes (interfacing a BioSemi Active Two system using ActiView acquisition software) placed according to the international 10–20 system in a BioSemi fabric head cap. External electrodes were placed above bilateral mastoid bones. The vertical and horizontal electrooculogram (EOG) was recorded from electrodes placed on the infraorbital ridge and outer canthi of the left eye, respectively. All channels were referenced to the implicit reference of the BioSemi system during recording.

Procedure

Participants were tested individually in a quiet booth. At the outset of the experiment, participants were provided with instructions that advised them of the probabilistic nature of the task. Specifically, they were informed that while perfect performance was not possible, they could use the feedback they received on each trial to improve their performance. Participants were instructed to try and respond correctly on

as many trials as possible, and to respond quickly, but without sacrificing accuracy.

Figure 1 summarizes the time course of an experimental trial. Each trial began with a fixation cross presented to the center of the screen for 1000 ms. To reduce eye movement artifacts in the EEG recording, participants were instructed to maintain central fixation for the duration of each trial. The stimulus was then presented centrally for 500 ms before response options (labeled “A” and “B”) appeared underneath. The 500 ms interval between stimulus onset and onset of response alternatives was used to avoid temporal smearing of the ERPs associated with attentional processes related to stimulus onset with response-preparatory ERPs elicited by presentation of the response alternatives (e.g., Takeda et al. 2008). The stimulus and response options remained visible until a category response was made by pressing either the F or J key. Mappings relating categories to response keys were randomly determined on each trial, and were reflected in the left-right positioning of the “A” and “B” labels (i.e., when “A” appeared to the left of the “B” label, the F key indicated a category A response, when “A” appeared to the right of “B”, the F key indicated a category B response). Corrective feedback was presented 500 ms after the participant’s response, to avoid temporal smearing of response- and feedback-locked ERP components. Feedback was presented underneath the stimulus for 1000 ms. The feedback on each trial identified the correct category label (i.e., “A” or “B”) as well as the accuracy of the participant’s response (i.e., “+” for a correct

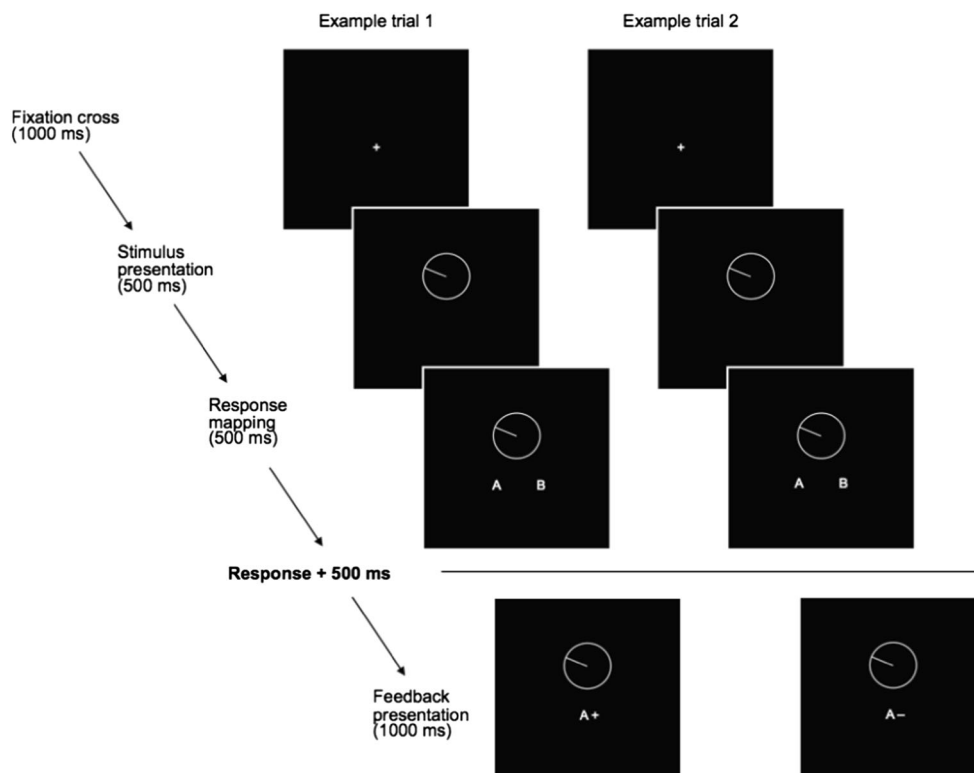
response or “-” for an incorrect response). For example, “A+” indicated that the stimulus belonged to category A and the participant’s response was correct. Feedback was coded in this way to minimize eye movements, as pilot testing revealed that participants struggled to maintain fixation when presented with longer text-based feedback. Trials were separated by a blank screen presented for 100 ms.

Participants completed 720 trials of the probabilistic categorization task, divided into 18 40-trial blocks. Each block comprised 10 presentations of each of the four stimuli. Trial order was uniquely randomized for each participant. Self-paced breaks were provided every 20 trials. An experimental session lasted for approximately 1 h and 45 min (1 h to complete the task, and 45 min for EEG setup and administering task instructions).

EEG Pre-processing

Offline, all scalp electrodes were re-referenced to the average of the left and right mastoids. Data were band-pass filtered at 0.1 to 70 Hz, and a standard 45 to 55 Hz notch filter was applied to minimize line noise. Filtering was conducted using standard routines included in the EEGLAB toolbox for MATLAB (Delorme and Makeig 2004). To obtain feedback-locked ERPs, an epoch was used that ranged from 100 ms before until 1500 ms after feedback onset. Average EEG activity from the 100 ms interval preceding feedback onset served as a baseline. All trials were visually inspected, and

Fig. 1 Two example trial sequences. Each trial began with central presentation of a fixation cross for 1000 ms. The stimulus was then presented centrally. Response mappings relating category outcomes to response keys were presented 500 ms after stimulus onset. The stimulus and response mappings remained visible until a response was made. 500 ms after a response, the response mapping was replaced by corrective feedback indicating the correct category outcome for that trial (i.e., “A” or “B”) as well as response accuracy (i.e., “+” for a correct response or “-” for an error response). Feedback remained on screen alongside the stimulus for 1000 ms



those containing artifacts unrelated to eye blinks (e.g., facial muscle movements or excessive skin conductance) were manually removed. The data were subjected to an independent component analysis (ICA) as implemented in the EEGLAB toolbox, so that eye blink- and eye movement-related artifacts could be identified and then removed from the data using EEGLAB’s standard routines. Finally, an additional artifact screening was performed which removed contaminant trials that had maximum/minimum amplitudes exceeding $\pm 200 \mu\text{V}$ (e.g., Bode et al. 2014).

Electrode selection for ERP measurement was based on sites commonly used in the study of the FRN (i.e., Fz, FCz, Cz, and CPz). We confirmed that our choice of electrode sites was reasonable by inspecting the local maximum of the grand-averaged difference waveforms for each electrode, participant, and the whole group, along with FRN scalp topography. These checks confirmed that the FRN was present at these electrodes at both the whole group and individual level. Analogous methods were used to identify the time window for measuring the FRN in that our choice of time window was informed by both our own data and by the existing literature. Visual inspection of changes in the topography of the FRN through time revealed a clear FRN that was observable between 220 and 380 ms following feedback onset. Because this interval was narrowly contained within the 200–400 ms window reported by Miltner et al. (1997), we did not make any further adjustments to this time window. For each participant, EEG epochs were time-locked to feedback onset and were analyzed using the ERPLAB plug-in (Lopez-Calderon and Luck 2014) for the EEGLAB toolbox. For ERP analyses of the FRN, data from four midline electrode sites (Fz, FCz, Cz, and CPz) were pooled by collapsing across the four stimuli and averaging waveforms across electrodes for each block (e.g., Bellebaum and Daum 2008; Potts et al. 2011). For each participant, we computed the mean ERP amplitude from 220 to 380 ms after feedback onset separately for correct and error feedback trials. Mean amplitude was used because it is less susceptible to high frequency noise (cf. peak amplitude measured within an interval), and is consequently more robust when comparing measurements of waveforms based on different numbers of trials (e.g., Luck 2005).

Overview of RASHNL

RASHNL is an exemplar-based model of probabilistic category learning developed by Kruschke and Johansen (1999). Here, we summarize the model, as we have implemented it for the purposes of the current study, following relevant precedent set by Craig et al. (2011). Our use of unidimensional stimuli obviates the need for the dimensional attention-shifting mechanisms of the original model, and so our description omits these details; interested readers are referred to Kruschke and Johansen (1999) for complete details.

RASHNL assumes that presentation of a stimulus activates exemplar-based representations of previously encountered stimuli according to the following similarity function,

$$a_i = \exp(-c|\psi_i - s|). \tag{1}$$

Equation 1 states that the activation of exemplar i , a_i , is an exponentially decreasing function of the similarity between the orientation value of exemplar i , ψ_i , and the presented stimulus, s . Exemplar activation declines with decreasing similarity at a rate set by the exemplar specificity constant, c . For the purposes of modeling, we normalized the range of radial line orientations to fall between 0 and 1, such that s could take on the values of 0, 0.33, 0.67, and 1.

The summed exemplar activation propagates to category outcome nodes, causing activation in proportion to the associative weights linking exemplar i with category outcome j , w_{ij} . The activation of category outcome node j is given by,

$$\omega_j = \sum_i w_{ij} a_i. \tag{2}$$

Category node activations are converted to choice probabilities via an exponentially scaled version of Luce’s (1963) choice rule. The predicted probability of a category A response, $P(A)$, is given by,

$$P(A) = \frac{\exp(\phi\omega_A)}{\exp(\phi\omega_A) + \exp(\phi\omega_B)}, \tag{3}$$

where ϕ is a response scaling parameter that controls how deterministically the model generates category responses. As ϕ increases, even small differences in category outcome activation lead to deterministic responding favoring the category node with higher activation.

RASHNL adjusts its behavior over the course of learning in response to corrective feedback using a standard error-driven learning rule (e.g., Rescorla and Wagner 1972). Changes in model behavior are governed by changes in the associative weights linking exemplars to category outcome nodes according to,

$$\Delta w_{ij} = \lambda(t_j - \omega_j) a_i, \tag{4}$$

where t_j is a *teacher value*, determined by the correct category outcome on a given trial. When a stimulus is assigned to category j , $t_j = 1$, otherwise $t_j = 0$. Equation 4 states that associative weights are adjusted in proportion to prediction error (i.e., the discrepancy between the observed category outcome, expressed by the teacher value, and the activation of the category nodes), which is scaled by a learning rate parameter, λ .

Importantly, RASHNL assumes that the learning rate in Eq. 4 can be progressively reduced over the course of a task, resulting in feedback discounting. To achieve this, on each

trial of the task, the learning rate, λ , is multiplied by a discount rate, $r(n)$. On trial n of a task, the discount rate is given by

$$r(n) = \frac{1}{1 + \rho n}, \tag{5}$$

where ρ is a feedback discounting parameter. Higher values of ρ reflect earlier onset of feedback discounting, resulting in task performance being strongly influenced by feedback received very early on in the task. As ρ decreases, feedback discounting appears progressively later in learning and has a less pronounced effect on responding, allowing feedback encountered even late in the task to affect learning and influence performance.

Models were fit to data by estimating values for the learning rate, λ , exemplar specificity, c , response scaling, ϕ , and, where relevant, the feedback discounting parameter, ρ , by minimizing the (negative) binomial likelihood,

$$-\ln L = - \sum_{i=1}^{18} \sum_{j=1}^4 \sum_{k=1}^2 f_{ijk} \ln(p_{ijk}), \tag{6}$$

where i indexes the 18 blocks of trials, j indexes the four stimuli, and k indexes the two category outcomes; f_k denotes the observed frequency of category k responses, and p_k denotes the proportion of category k responses predicted by the model.

Results

Data Screening

Data from 40 participants were collected. However, nine produced EEG data sets that were unusable, either because of excessive eye movement artifacts, head movements during the recording, and/or slow voltage drifts during the experiment due to high skin conductance. Data from these individuals were not included in the final analysis and were identified by the experimenter as potentially problematic while the participant completed the task. To ensure that the behavioral data from the remaining 31 individuals reflected meaningful performance on the categorization task, we summarized each participant’s performance by computing a *probability matching score* (PM score; Friedman and Massaro 1998). The probability matching score provides a measure of how closely people’s responding tracked feedback probabilities for each item during the task, and is defined as

$$PM_i = [P(A|i) - R(A|i)]s_i, \tag{7}$$

where i indexes the four stimuli used in the task. For each stimulus, an individual’s PM score is determined by the difference between the probability of A feedback, given

presentation of stimulus i , $P(A|i)$, minus the proportion of A responses made by the participant, given presentation of stimulus i , $R(A|i)$. This difference is multiplied by a stimulus-dependent sign value, s_i . The sign value is set to +1 when $P(A|i) > 0.5$, and to -1 otherwise. To get a global PM score for each participant, we averaged PM scores across all four stimuli.

The PM score is a useful way of summarizing performance in probabilistic learning environments, as it distinguishes random responding from other meaningful, but potentially sub-optimal, response strategies. In probabilistic learning environments with stable cue-outcome contingencies, the optimal response strategy is to *maximize* the proportion of responses that align with the most frequently occurring category outcome for each stimulus. Maximizing entails assigning stimuli to category A whenever their feedback probabilities exceed 0.5, and assigning them to category B whenever their feedback probabilities are less than 0.5. Typically, however, people do not maximize in probabilistic environments (e.g., Shanks et al. 2002), and instead use a so-called *probability matching* strategy. Probability matching entails assigning each stimulus to category A in roughly equal proportion to the rates at which the individual stimuli are paired with category A feedback. For our task, maximizing would result in a global PM score of -0.3, probability matching would result in a global PM score of 0, and random responding would result in a global PM score of 0.2. Two participants failed to respond at better than chance levels, indicated by PM scores greater than 0.2. Data from these two participants were removed from further analysis, leaving a final sample of 29 participants.

Behavioral Data

We first examined group level performance to confirm that (1) people were sensitive to the different feedback probabilities and responded differently to the four stimuli, and (2) the pattern of performance was not affected by the way changes in line orientation mapped onto changes in feedback probability. To this end, a 2 (condition: ascending vs. descending) \times 4 (feedback probability) \times 18 (block) between-within ANOVA was conducted on the proportion of a category A responses, with condition as a between participants variable. The analysis returned significant effects of feedback probability, indicating that people successfully differentiated the four stimuli, $F(3, 81) = 53.60, MS_e = 0.65, p < 0.001, \eta_p^2 = 0.67$. There was also a significant interaction between feedback probability and block, reflecting changes in response probabilities over the course of the task that were consistent with learning, $F(51, 1377) = 1.64, MS_e = 0.25, p = 0.003, \eta_p^2 = 0.06$. Condition did not have a significant effect, nor did it enter into any higher-order interactions with any other variable(s), $F_s < 0.79, p_s > 0.71$. For all subsequent analyses, the data were collapsed across conditions. The leftmost panel of Fig. 2 shows

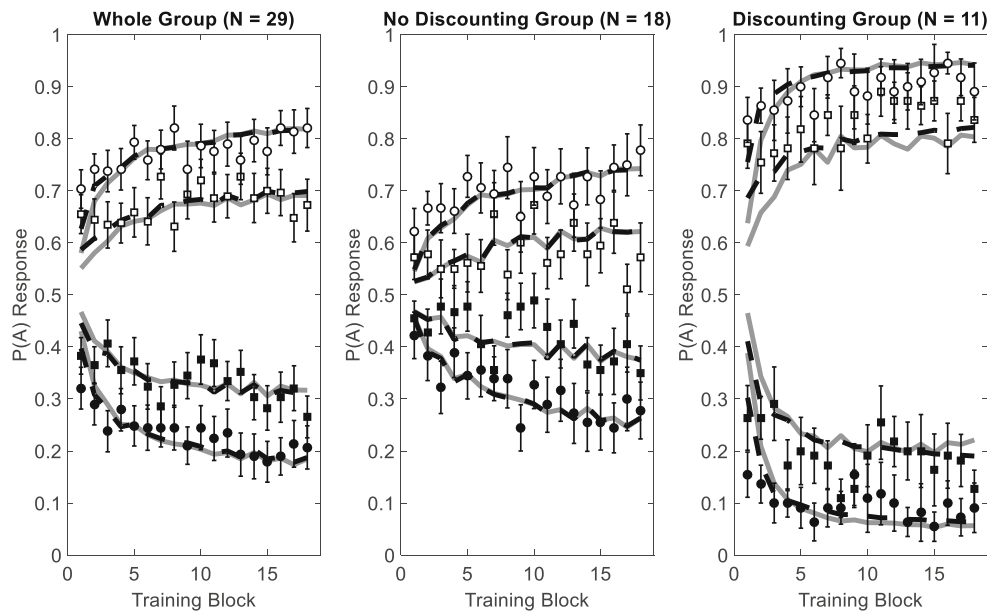


Fig. 2 Group-averaged feedback probabilities for each stimulus over the course of the entire learning task. Plotting symbols show choice probabilities for the 0.2, 0.4, 0.6, and 0.8 stimuli (filled circles, filled squares, open squares, open circles, respectively). Error bars show the standard error of the mean. Data for the whole group are shown in the left panel. The middle and right panels show the sample data split according to whether behavioral modeling with RASHNL provided support for feedback discounting (right panel) or not (middle panel).

The effect of feedback discounting on the choice probabilities for each group is shown by comparing the gray lines (model predictions with no feedback discounting) to the dashed black lines (model predictions with feedback discounting). For the middle panel, the model predictions are virtually unchanged regardless of whether feedback discounting is included. For the right panel, the model predictions, especially in the early stages of the task, depend strongly on whether feedback discounting was included

behavioral data, averaged across all participants. The probability of a category A response is plotted in each block, for each of the four stimuli. Circles denote stimuli with more diagnostic feedback probabilities (i.e., 0.8 and 0.2 feedback probabilities), while squares denote stimuli with less diagnostic feedback probabilities (i.e., 0.6 and 0.4). Filled symbols denote stimuli with feedback probabilities less than 0.5, unfilled symbols denote stimuli with feedback probabilities greater than 0.5. The pattern of responding is typical of probabilistic categorization tasks (e.g., Craig et al. 2011). There is a tendency for participants to slightly overshoot the assigned feedback probabilities, though, in the aggregate, the overall pattern of performance is broadly consistent with probability matching (e.g., Shanks et al. 2002).

Modeling Individual Choice Behavior with RASHNL

We evaluated feedback discounting by fitting a model of probabilistic category learning (RASHNL; Kruschke and Johansen 1999) to each individual’s behavioral data. For each participant, we contrasted the fits of two models, one that included feedback discounting (i.e., with ρ free to vary), and one that did not (i.e., with ρ fixed to 0). Because these models were nested, differing only in terms of a single free parameter (viz. the feedback discounting rate), likelihood ratio tests were used to determine if the feedback discounting model provided a significantly better fit to the behavioral data compared to the

model without feedback discounting. This provides us with a purely behavioral assessment of feedback discounting, which we subsequently use to set up key contrasts when analyzing the EEG data.

Participants for whom the feedback discounting model provided a significantly better fit to the behavioral data ($N = 11$) were identified as the *discounting group*, while the remaining participants, for whom discounting provided no significant improvement to model fit ($N = 18$), were identified as the *no discounting group*. Figure 2 shows performance profiles for the no discounting group (middle panel) as well as the discounting group (right panel) along with model predictions without feedback discounting (gray lines) as well as with feedback discounting included (dashed black lines). Best fitting parameter estimates averaged across individuals in the discounting and no discounting groups, respectively, are shown in Table 1. The averaged fits and parameter estimates show that for the no discounting group, allowing ρ to vary freely produced virtually no improvement in fit compared to when ρ was fixed to zero. By contrast, for the discounting group, the fit of the model was improved by assuming a non-zero value for ρ .

It is clear from the middle and right panels of Fig. 2 that the discounting and no discounting groups responded quite differently over the course of the task. Specifically, the responses of the no discounting group appear to reflect approximate probability matching, whereas the responses of the

Table 1 Best fitting parameter estimates, averaged across individuals, for the no discounting group, and discounting group. Parameter estimates for models with (free ρ) and without feedback discounting (fixed ρ) are

shown. The negative log likelihood for each model, $-\ln L$, is also reported. Information about the preferred model for the no discounting and discounting groups, based on likelihood ratio tests, is presented in italic

Participant group and model	c	λ	ϕ	ρ	$-\ln L$
<i>No discounting group—fixed ρ</i>	9.02	0.08	9.82	0*	440.59
No discounting group—free ρ	12.61	0.09	16.72	0.005	440.33
Discounting group—fixed ρ	4.94	0.03	6.79	0*	271.78
<i>Discounting group—free ρ</i>	10.79	1.12	18.97	0.80	264.04

*Denotes a fixed parameter

discounting group clearly overshoot the feedback probabilities and are more akin to maximizing. These qualitative differences in responding were confirmed via a 2 (group: discounting vs. no discounting) \times 4 (feedback probability) \times 18 (block) between-within ANOVA, which returned a group \times feedback probability interaction, $F(3, 81) = 22.36$, $MS_e = 0.36$, $p < 0.001$, $\eta_p^2 = 0.45$. The differences in the response scaling parameter, ϕ , reflect the differences in responding across groups, with larger mean values for the discounting group driving the stronger tendency toward more deterministic responding (i.e., a maximizing-like pattern of responding). For convenience, we refer to these behavioral differences in responding (i.e., approximate probability matching and maximizing) as “response strategies,” though we only use the term in a descriptive fashion. We do not make any assumptions about whether people deliberately chose to respond in these ways or even if they were aware of the nature of their responding. Due to the presence of these two distinct response strategies in the behavioral data, we separate data from the discounting and no discounting groups in subsequent analyses.

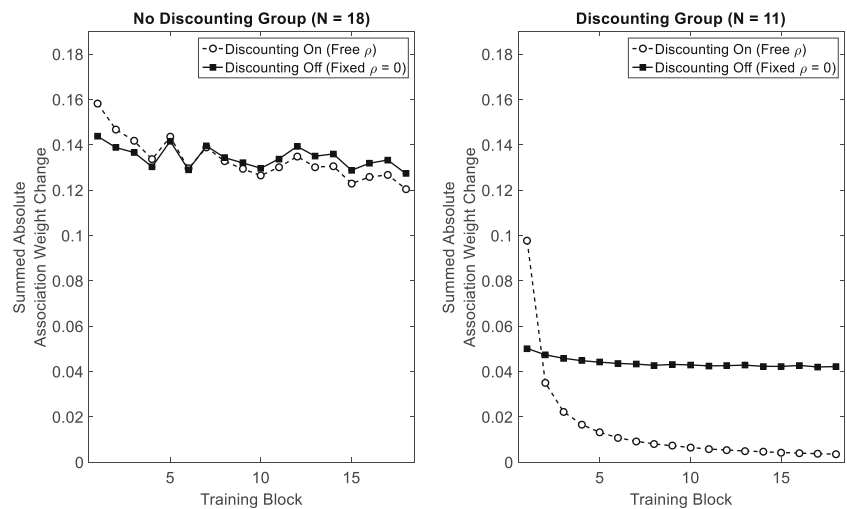
It is also clear from Fig. 2 that feedback discounting had different effects on model predictions for the two groups. As expected, for the no discounting group, allowing ρ to be freely estimated from the data produced virtually no differences in model predictions compared to when ρ was fixed to 0: The gray and dashed black lines sit on top of one another across all trial blocks. By contrast, for the discounting group, model predictions—especially in the earlier trial blocks—improved dramatically when ρ was freely estimated from the data. The effect of feedback discounting can be further understood by comparing learning rates for the discounting and no discounting groups. Although learning rates are higher for the discounting group—suggesting increased sensitivity to feedback and more volatile changes in associative weights across successive trials—this is tempered by the presence of feedback discounting. The net effect is that the discounting group makes relatively larger adjustments in cue-outcome associations early on in the task, and this rapid initial learning becomes “locked” into place due to feedback discounting. A higher learning rate that favors feedback encountered early in

the task, combined with the more extreme decision strategy (i.e., higher values of ϕ), allows choice behavior in the discounting group to rapidly tend toward a maximizing-like response strategy. By contrast, the no discounting group steadily processes corrective feedback for the duration of the task due to the learning rate, λ , holding a constant value throughout the task. This results in relatively small, but consistent changes in performance over the course of the task, where trial-to-trial responding is continuously perturbed by feedback. Probability matching emerges naturally because relative associative strengths change incrementally throughout the task. Over the course of a large number of trials, the relative associative strengths come to reflect the relative feedback probabilities.

Having identified these individual differences in the behavioral data, two key critical questions remain. The first revolves around how feedback discounting affected learning in the two groups. By analyzing differences in the way the model updated its association weights across the discounting and no discounting groups, we can gain insight into how corrective feedback was processed at different stages of the task. The second question is how the presence (or absence) of feedback discounting affected changes in the FRN over the course of learning.

Analysis of Learning via Changes in Association Weights We investigated the effect of feedback discounting on learning for the discounting and no discounting groups by tracking changes in association weights over the course of the task for both groups. Because association weights both increase and decrease depending on the feedback presented on a given trial (Eq. 4), we operationalized learning in terms of the summed absolute changes in association weights on each trial. To quantify the cumulative effects of learning, we summed the changes in association weights across trials within each block. Figure 3 shows how feedback discounting affected learning for the no discounting and discounting groups (left and right panels, respectively). For each group, changes in association weights summed within each trial block are shown when RASHNL was fit to the behavioral data with feedback discounting on (i.e., ρ free to vary; open symbols) and with

Fig. 3 Summed absolute changes in association weights averaged across participants in the no discounting group (left panel) and discounting group (right panel). For each group, the effects of learning are contrasted with feedback discounting on (open white symbols) and off (filled black symbols). Whereas feedback discounting has little effect on learning in the no discounting group, changes in association weights are rapidly suppressed in the discounting group



feedback discounting off (i.e., ρ fixed to 0; filled symbols). For the no discounting group, learning (i.e., changes in association weights) was fairly steady over the course of the task, and largely unaffected by whether feedback discounting was incorporated in the model, echoing the fits of the behavioral data shown in Fig. 2. By contrast, for the discounting group, learning was heavily influenced by whether feedback discounting was incorporated in the model. When feedback discounting was not included in the model, the changes in association weights were relatively stable for the duration of the task—much like the pattern seen for the no discounting group. However, when feedback discounting was incorporated into the model, learning was effectively restricted to the earliest trial blocks, with only minimal changes in association weights later on. We confirmed the differences in patterns of learning in the model via a 2 (group: discounting vs. no discounting) \times 2 (feedback discounting: on vs. off) \times 18 (block) between-within ANOVA. The critical three-way interaction was significant, $F(17, 459) = 3.609$, $MS_e < 0.001$, $p < 0.001$, $\eta_p^2 = 0.12$, highlighting the differential effect of feedback discounting on learning performance in the models across the two groups. The most important comparison, however, was of learning in the no discounting group when ρ was fixed to 0 (i.e., no feedback discounting) and in the discounting group when ρ was freely estimated (i.e., feedback discounting was on), as these were the models that best characterized the behavioral data for each group. Changes in association weights for these models were analyzed by a 2 (group: discounting vs. no discounting group) \times 18 (block) between-within ANOVA, which returned a main effect of group, $F(1, 27) = 4.69$, $MS_e = .377$, $p < 0.001$, $\eta_p^2 = 0.15$, a main effect of block, $F(17, 459) = 8.54$, $MS_e = 0.001$, $p < 0.001$, $\eta_p^2 = 0.24$, as well as an interaction between the factors, $F(17, 459) = 5.30$, $MS_e = 0.001$, $p < 0.001$, $\eta_p^2 = 0.16$. These effects summarize (1) the difference in overall learning, (2) the tendency for changes in association

weights to diminish over the course of the task, and (3) the different rates at which learning occurred for the two groups. Taken together, the modeling results clearly show that learning was modulated by feedback discounting. Whereas the discounting group rapidly reduced their reliance on corrective feedback, the no discounting group consistently updated association weights in response to feedback for the duration of the task.

Summary of Behavioral Modeling The results of the behavioral modeling show clear evidence of individual differences in both response strategies and feedback discounting. Interestingly, the two appear to be linked with the no discounting group favoring a probability matching strategy and the discounting group favoring a response strategy more akin to maximizing. Because responding in the discounting group more closely approximated a rational maximizing strategy, a counterintuitive finding from our study is that participants who performed better on the task were the ones who were more likely to *ignore* corrective feedback. Put another way, those who performed poorly on the task were those who continued to process corrective feedback for the duration of the experiment. This result is striking, and we discuss possible explanations for why these individual differences emerged in the “[Discussion](#).”

EEG Analysis

To investigate the effects of feedback discounting on the FRN, we contrast the EEG data for participants in the discounting and no discounting groups. As noted above, all EEG analyses used data that were pooled across four electrodes (Fz, FCz, Cz, and CPz). We define the FRN as the mean amplitude of ERPs recorded on correct feedback trials subtracted from the mean amplitude of ERPs recorded on error feedback trials.

Difference Wave Analysis First, we tested for the presence of the FRN in the grand averaged data (i.e., combining data across all participants in all blocks). A typical FRN, as measured by the difference wave (i.e., the difference in waveforms obtained on correct vs. error feedback trials), is shown in the solid black line in Fig. 4. The component waves for correct (dashed black line) and error trials (dashed gray lines) are also shown. A one-sample *t* test on the grand average difference wave confirmed that its mean amplitude was significantly different from 0, $t(28) = -2.31, p = 0.028, r^2 = 0.40$, reflecting an overall FRN in our data.

Our key focus is whether the FRN is different for people who discount feedback compared to those who do not discount feedback. Because EEG data are noisier than the corresponding behavioral data, we grouped the EEG data into six trial epochs, where each trial epoch summarized amplitude data averaged across consecutive non-overlapping three-block sets of trials (i.e., the first trial epoch comprised data from blocks one to three, the second epoch comprised data from blocks four to six, and so on). Analysis of the EEG data across these six trial epochs (as opposed to 18 blocks of trials) permits more straightforward identification of the major patterns in the data. We note that we conducted the same analyses on the EEG data across all 18 blocks, which resulted in the same conclusions and theoretical interpretations.

Figure 5 shows changes in the mean FRN amplitude over the course of the task for the discounting and no discounting groups (open and filled symbols, respectively). For the no discounting group, FRN amplitude is relatively consistent across trial epochs, maintaining a negative deflection throughout the task. By contrast, for the discounting group, FRN amplitude straddles zero, showing little systematic change. We investigated these patterns via a group (2: discounting vs. no discounting) \times trial epoch (1–6) between-within ANOVA on the difference wave amplitudes over the course

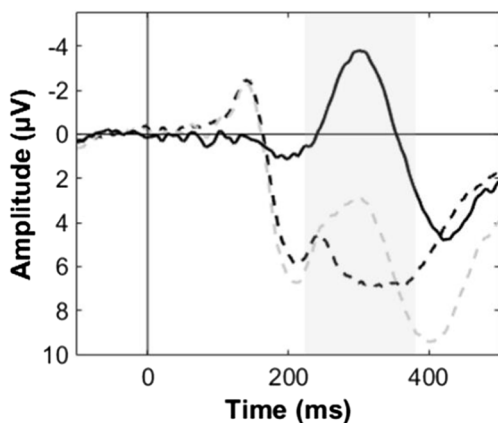


Fig. 4 Overall weighted grand-averaged feedback-locked waveforms sorted according to correct (black dashed line) and error (gray dashed line) trial outcomes. The FRN, measured as the difference wave, is shown as the solid black line. Data were pooled across Fz, FCz, Cz, and CPz. Following convention, negative voltage is plotted upwards

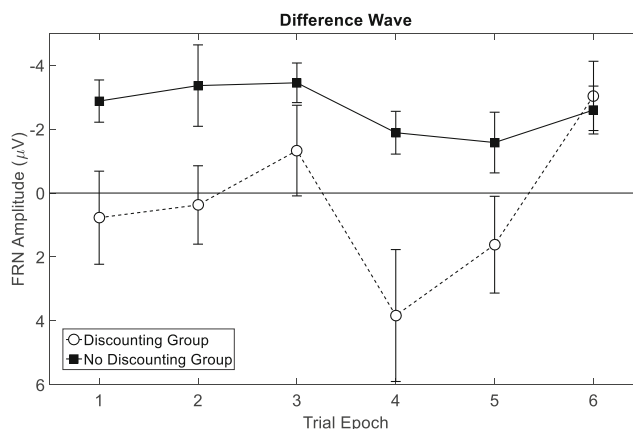


Fig. 5 Mean FRN amplitude in each block for the no discounting (solid line) and discounting (dashed line) groups. Error bars show the standard error of the mean. Note that negative voltages are plotted upward, following convention

of the task. The analysis revealed a significant main effect of group, $F(1, 27) = 8.43, MS_e = 43.67, p = 0.007, \eta_p^2 = 0.24$, reflecting a stronger, more negative FRN amplitude for the no discounting group compared to the discounting group. We also observed a significant main effect of trial epoch, $F(5, 135) = 4.68, MS_e = 12.04, p = 0.001, \eta_p^2 = 0.15$, as well as an interaction between the two factors, $F(5, 135) = 2.39, MS_e = 12.04, p = 0.04, \eta_p^2 = 0.08$, suggesting that changes in the FRN followed a different pattern across the discounting and no discounting groups.

We followed up these results by first investigating the interaction via separate one-way ANOVAs for each group on the trial epoch data. For the no discounting group, there was no effect of trial epoch on FRN amplitude, $F(5, 85) = 0.91, MS_e = 11.54, p = 0.48$, reflecting a consistent level of feedback processing for the duration of the task. This interpretation was further supported by a series of one-sample *t* tests comparing FRN amplitude in each epoch against 0. For each epoch, FRN amplitude in the no discounting group was significantly less than 0, $t_s(17) < -2.64, ps < 0.017, r^2_s > 0.29$. The lone exception was trial epoch 5, which was not significantly different from 0, $t(17) = -1.67, p = 0.11$. For the discounting group, there was an effect of trial epoch on FRN amplitude, $F(5, 50) = 4.82, MS_e = 12.91, p = 0.001, \eta_p^2 = 0.33$; however, the amplitude differences did not follow any systematic pattern. One-sample *t* tests comparing the FRN amplitude in each epoch against 0 revealed no significant differences for the first five trial epochs, $t_s(10) < 1.86, ps > 0.09$. For the sixth trial epoch, FRN amplitude was significantly less than 0, $t(10) = -2.81, p = 0.02, r^2 = 0.44$, but the lack of a consistent overall pattern leads us to believe this difference to be spurious.

Taken together, the results of the difference wave analysis suggest that people in the discounting group processed both correct and error feedback to a similar extent at each trial epoch. The lack of differential responding to error feedback resulted in the absence of an FRN. By contrast, for the no

discounting group, error feedback elicited a stronger neural response compared to correct feedback at all stages of learning, resulting in a consistently negative FRN.

The FRN results dovetail with the model-based analysis, which showed rapid onset of feedback discounting for the discounting group. For the discounting group, learning, as indexed by changes in association weights, did not progress beyond the very first trial blocks, as feedback was rapidly discounted. One potential interpretation of the lack of an FRN for this group is that feedback may have been largely ignored after the very early stages of learning (cf. Luque et al. 2012). By contrast, model-based analysis of the no discounting group showed consistent changes in association weights throughout the task. Importantly, feedback discounting was linked to *better* task performance (i.e., a higher proportion of correct responses). This means that feedback discounting, and the absence of a clear FRN for the discounting group cannot be viewed as reflecting task disengagement or failure to learn about feedback contingencies, both of which would lead to random responding. Instead, discounting seems to reflect an adaptive way to ignore uninformative and persistent error feedback once the cue-outcome contingencies had been acquired by the learner. We return to the question of when people may or may not discount feedback in the “Discussion.”

Component Wave Analysis To gain further understanding of how correct and error feedback was processed by individuals in the discounting and no discounting groups, we also analyzed changes in the amplitudes of the component waves elicited by correct and error feedback during the task. Although the FRN is conventionally studied via analysis of the difference wave, as we have done above, some studies have reported differential patterns in the ERP response to correct and error feedback (e.g., Bellebaum and Daum 2008).

We first consider changes in the neural response to correct feedback (Fig. 6). A group (2: discounting vs. no discounting) × trial epoch (1–6) between-within ANOVA on the ERP amplitudes evoked by correct feedback revealed only a significant effect of trial epoch, $F(5, 135) = 5.80, MS_e = 8.40, p < 0.001, \eta_p^2 = 0.18$, reflecting the tendency for component wave amplitudes to reduce as the task progressed. Although Fig. 6 suggests a weaker response to correct feedback for the discounting group, the main effect of group failed to reach significance, $F(1, 27) = 2.74, p = 0.11$. The interaction was not significant, $F(5, 135) < 0.43, p = 0.83$, suggesting a common pattern of change across groups.

An analogous group (2: discounting vs. no discounting) × trial epoch (1–6) between-within ANOVA was conducted for the component wave amplitudes for error trials (Fig. 7). For these data, there was a significant effect of trial epoch, $F(5,$

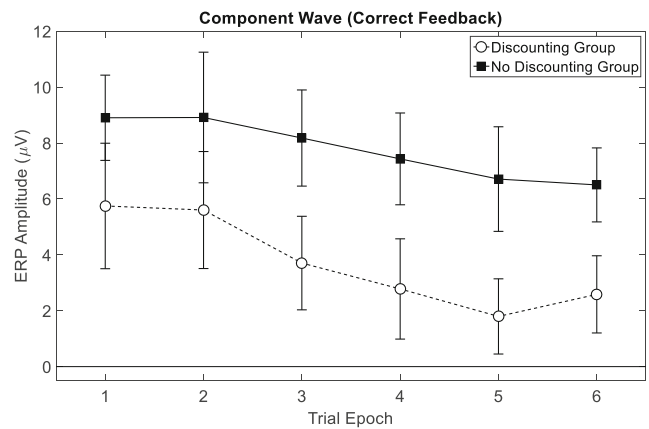


Fig. 6 Changes in the amplitude of ERPs for trials with correct feedback for the discounting (dashed line) and no discounting (solid line) groups. Error bars show the standard error of the mean. Note positive voltages are plotted upwards for displaying component wave data

135) = 9.62, $MS_e = 8.76, p < 0.001, \eta_p^2 = 0.26$, which was qualified by an interaction with group, $F(5, 135) = 3.46, MS_e = 8.76, p = 0.006, \eta_p^2 = 0.11$. As was the case with the correct feedback trials, there was no main effect of group, $F(1, 27) = 0.27, p = 0.61$. We followed up the interaction by conducting separate one-way ANOVAs on the amplitude data across trial epochs. For the no discounting group, the effect of trial epoch was not significant, $F(5, 85) = 1.39, p = 0.24$, implying a consistent neural response to error feedback throughout the task. For the discounting group, there was a significant effect of trial epoch, $F(5, 50) = 8.58, MS_e = 9.96, p < 0.001, \eta_p^2 = 0.46$, which suggested a decrease in ERP amplitude elicited by error responses over the course of the task (see Fig. 7). We tracked changes in the error component wave amplitude across consecutive trial epochs via paired samples t tests, which revealed significant decreases in amplitude across trial epochs four to five, $t(10) = -2.42, p = 0.04, r^2 = 0.37$, and

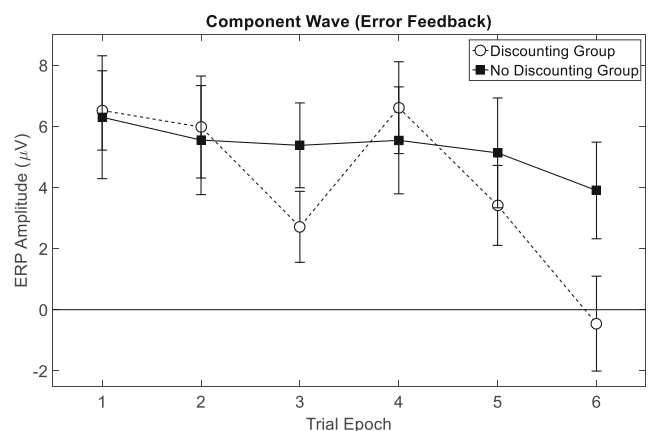


Fig. 7 Changes in the amplitude of ERPs for trials with error feedback for the discounting (dashed line) and no discounting (solid line) groups. Error bars show the standard error of the mean. Note positive voltages are plotted upwards for displaying component wave data

five to six, $t(10) = -2.43$, $p = 0.04$, $r^2 = 0.37$. There was a marginal reduction in amplitude across trial epochs two to three, $t(10) = -2.22$, $p = 0.05$, $r^2 = 0.33$. Surprisingly, there was a significant increase in amplitude between epochs three and four, $t(10) = -2.46$, $p = 0.03$, $r^2 = 0.38$, which is not readily interpretable. On balance, for the discounting group, we suggest that the amplitudes of component waves elicited by error feedback tended to reduce over the course of the task.

Summary of EEG Analysis The results of the EEG analyses show clear differences in the way corrective feedback was processed by the discounting and no discounting groups. For the no discounting group, there was little change in the FRN over the course of the task, as measured by the standard difference wave. Though there was some evidence for a reduction in ERP amplitudes associated with receiving correct feedback, the same pattern did not hold for error feedback. Overall, it appears that while the neural response to receiving correct feedback decreased over time, error feedback elicited a stable sustained response throughout the task. These changes are consistent with the model-based analysis showing that people in the no discounting group learned from feedback in a similar way, and to a similar extent, for the duration of the task. The diminished neural response elicited by correct feedback is in line with what would be expected on the basis of incremental learning of the cue-outcome contingencies across trials. As learning progresses, correct predictions become more likely and are associated with less prediction error, as characterized by standard error-driven learning (Eq. 4) and relevant theory (e.g., Holroyd and Coles 2002; Kruschke and Johansen 1999; Rescorla and Wagner 1972). However, error feedback elicited a more consistent response in the ERP data, as shown by the relatively stable component wave amplitude in Fig. 7, suggesting that these participants were still actively relying on feedback to guide future performance (cf. Cohen et al. 2007).

For the discounting group, the difference wave analysis suggested that there was virtually no differential processing of correct and error feedback over the course of the task. This is unsurprising in light of the model-based analysis, which showed that these participants rapidly discounted feedback received after the very earliest learning blocks. If people in the discounting group were no longer relying on feedback for learning purposes, it follows that the neural response to both correct and error feedback would be similar. The lack of a clear FRN for the discounting group, as measured by the standard difference wave, is consistent with this expectation. Importantly, for both correct and error feedback trials, we observed reductions in ERP amplitudes elicited by both kinds of feedback, consistent with feedback discounting. Overall, the EEG results for this group converge with the model-based analysis of choice behavior reported earlier.

Discussion

Feedback discounting has been proposed as a general strategy for dealing with the continued presence of prediction error in learning environments defined by probabilistic stimulus-outcome contingencies (Kruschke and Johansen 1999). While the idea of feedback discounting is an appealing one from an effort minimization perspective—ignoring errors that one can do nothing about is simpler than adjusting behavior in vain—evidence for feedback discounting is surprisingly sparse. Behavioral studies have attempted to identify discounting via blocked trial designs (e.g., Edgell and Morrissey 1987), cognitive modeling (Kruschke and Johansen 1999), or some combination of the two (Craig et al. 2011). While successful, a limitation of these studies is that the presence of feedback discounting must be inferred solely on the basis of either task performance or interpretation of cognitive model parameters. We sought converging evidence for feedback discounting by combining model-based analysis of choice behavior with measurement of an established electrophysiological correlate of feedback processing, the FRN (Miltner et al. 1997).

Our model-based analysis of behavior revealed striking individual differences in task performance that differed depending on whether a feedback discounting mechanism was needed to account for people's choice behavior. Specifically, we found that people who were not identified as discounting feedback exhibited behavior consistent with probability matching. By contrast, people who were identified as discounting feedback produced a response profile that was more in line with maximizing. The differences in feedback discounting resulted in different learning dynamics, as revealed by changes in RASHNL's association weights. For the no discounting group, learning proceeded at a fairly consistent rate for the duration of the task. For the discounting group, however, learning was restricted to the earliest blocks of trials, before being rapidly suppressed. Because we used single-cue stimuli, following the design of Craig et al. (2011), feedback discounting in our task could not be attributed to sequential learning effects, such as blocking.

Critically, we were able to relate aspects of performance revealed by the model-based analysis to differences in the FRN across the discounting and no discounting groups. The no discounting group showed a sustained FRN that was consistent with a fixed learning rate and ongoing processing of feedback throughout the task. For the discounting group, we failed to find clear evidence of the FRN, and instead found that the electrophysiological response to both correct and error feedback steadily diminished over the course of the task. The EEG results for the discounting group lined up with the results of the model-based analysis, which showed that the effects of feedback discounting on learning were substantial, and affected performance from very early on in the task. On the whole,

our study provides novel support for the idea of feedback discounting, and further shows that the effects of discounting modulate the FRN in a fairly direct way. Before discussing the implications of our results, we first address some potential limitations of the current study.

Limitations

One potential concern about our study is the absence of a pronounced FRN for the discounting group. In particular, one could ask whether the differences in the FRN between the discounting and no discounting groups could be attributed to differences in task engagement or effort. We do not believe that the groups were differentially engaged in the task for two reasons. First, we note that participants in the discounting group performed well on the task, achieving rapid changes in accuracy in the early stages of learning, and implementing a response strategy that approached the rational maximizing profile. The adaptive changes in choice behavior for the discounting group militate against the idea that these participants were disengaged or failing to attend to the task. Second, our analysis of the component waves showed that, for the discounting group, there were gradual reductions in the electrophysiological response elicited by both correct and error feedback over the course of the task, consistent with feedback discounting. Taken together, the absence of a pronounced FRN for the discounting group is consistent with the model-based analysis of behavior, suggesting that these participants rapidly learned to associate stimuli with category outcomes, before discounting subsequent feedback, rather than disengaging from the task.

A second potential concern about the current study is the presence of multiple response strategies in the data. Specifically, the way feedback discounting aligned closely with the response strategies people used to complete the task (i.e., probability matching for the no discounting group, and maximizing for the discounting group) raises the question of whether the differences in the FRN we observed were due to differences in feedback discounting per se or a more global response strategy. This is a complex issue due to the lack of studies investigating how feedback discounting relates to factors such as response strategy. To our knowledge, ours is the first study to identify such a relationship, and it strongly suggests that feedback discounting and response strategy might often be naturally related. Feedback discounting makes sense in relation to maximizing, as a learner is only required to discern whether $P(A) > 0.5$ for each cue in order to accurately implement the strategy, which can be achieved with relatively few observations (Navarro and Newell 2014). Once the learner has determined which outcome is favored by each cue, there is little need to process additional feedback. It is potentially unsurprising, then, that we found feedback discounting to be linked with maximizing. We suspect, however, that feedback

discounting is ultimately dissociable from response strategy. For example, in their analysis of group-level performance, Craig et al. (2011) found model-based support for feedback discounting alongside clear behavioral evidence of probability matching. While it is possible that their group-level analysis obscures individual differences that are consistent with the ones we observed in our study, the close coupling of discounting and variation in response strategy need not occur. In the context of RASHNL, for example, feedback discounting affects learning of cue-outcome associations (Eqs. 4 and 5), whereas choice behavior based on those associations is controlled by the response scaling parameter at the decision stage (Eq. 3). It follows that feedback discounting, which affects how cue-outcome knowledge is learned, is logically and theoretically distinct from the decision processes that operate on what is learned through feedback. More research is clearly needed to fully understand how feedback discounting may (or may not) relate to response strategy. For now, we note that feedback discounting—identified via analysis of behavioral data alone—appears sufficient to reveal clear group differences in the FRN, and that this may often appear alongside a maximizing strategy.

Theoretical Implications

Individual Differences in Feedback Discounting Our results revealed clear individual differences in feedback discounting. This is an interesting finding because previous research investigating feedback discounting has focused on group-level performance (e.g., Craig et al. 2011; Kruschke and Johansen 1999). An open question is why such individual differences exist, given that all learners were required to learn the same set of cue-outcome contingencies. We suggest three likely sources of individual variation in feedback discounting, based on recent findings from the category learning literature: Variability in working memory capacity (WMC), preference for “rule-oriented” learning, and people’s assumptions about the stability of task contingencies.

To the extent that feedback discounting is linked to adoption of a maximizing strategy, individual differences in WMC may play an important role. For example, Rakow et al. (2010) found that higher WMC was associated with higher rates of maximizing in different prediction tasks. Conversely, Schulze and Newell (2016) recently showed that maximizing occurs less often when the strategy is made more difficult to implement (i.e., by randomly varying the mapping of response keys to task outcomes), indicative of a link between availability of cognitive resources and response strategy. In a related vein, Lewandowsky (2011) modeled the effect of higher WMC as conferring additional “rehearsals” when processing corrective feedback—effectively increasing the number of opportunities a learner has to strengthen the association between a cue and an outcome on a given trial. The relationship between WMC and

improved learning outcomes has been repeatedly shown in a variety of category learning tasks (e.g., Craig and Lewandowsky 2012, 2013; Kalish et al. 2017; Lewandowsky et al. 2012; Sewell and Lewandowsky 2012). If higher WMC supports more rapid learning, it is possible that in probabilistic environments, it might also promote discovery of the maximizing strategy and alleviate reliance on corrective feedback. Illuminating the link between feedback discounting and WMC is an important target for future research.

Another factor that might influence feedback discounting is variability in people's preferences to adopt association-based or abstraction-based approaches to learning (e.g., McDaniel et al. 2014). Whereas association-based learners attempt to relate individual exemplars with their respective outcomes, abstraction-based learners attempt to uncover regularities that explain exemplar-outcome relationships that hold across *multiple* stimuli. These preferences have been shown to be independent of WMC (Little and McDaniel 2015). Given that our task involved systematic relationships between line orientation and feedback probability (viz. probabilities either increased or decreased as line orientation changed), it is possible that some people attempted to extract a rule relating the different stimuli to their respective category outcomes. It is possible that once the key regularities relating stimulus orientation to category outcome were discovered, people would have no need to process further feedback. By contrast, people who learn by associating individual stimuli with category outcomes might be more likely to continually process feedback, in order to ensure precise estimation of the different feedback contingencies. We acknowledge that rule-based or hypothesis-testing approaches to learning might be viewed as permitting immediate (and complete) discounting of feedback once a suitable rule has been discovered. Indeed, we agree that this might be viewed as a somewhat degenerate form of feedback discounting that operates differently to feedback discounting on association-based learning. However, we believe that feedback discounting operates in a similar manner, regardless of whether it is tied to a rule-based or association-based approach to learning. From a modeling perspective, Craig et al. (2011) showed that feedback discounting, as implemented in our study, was wholly compatible with rule-based approaches to learning. They showed that a modified version of Erickson and Kruschke's (1998) ATRIUM model, which only had access to a single rule representation, produced similar patterns of discounting behavior as RASHNL and other exemplar-based models. In a different vein, Navarro et al. (2016) analyzed performance in an observe-or-bet task, and found that people had a tendency to periodically observe trial outcomes as a means of "checking" that their response strategy remained valid. Viewed in terms of rule-based responding, this sort of checking behavior is analogous to feedback discounting. After a suitable rule has been discovered, a stronger tendency to discount feedback would translate to less

frequent, but occasional, processing of corrective feedback in order to confirm that the rule remained valid. The rate at which this sort of occasional confirmatory processing of feedback declines in people applying rule-based strategies would present as a non-zero estimate of the ρ parameter in RASHNL.

A third potential source of individual differences in feedback discounting is based on people's prior beliefs about the stability of the learning environment. Schiffer et al. (2017) showed that expectations about the stability of cue-outcome contingencies modulated the FRN in a manner consistent with discounting feedback. When the learning environment was expected to be stable, people produced a weaker FRN even in the early stages of the task, much like we observed in the current study. People's expectations about the stability of feedback contingencies in probabilistic environments were recently investigated in detail by Navarro et al. (2016). They found that people's choice and information-seeking behavior in an observe-or-bet task were consistent with the belief that cue-outcome contingencies were more likely to change than to remain the same over the course of the task. A consequence of this belief is that people's responding will deviate from maximizing, which can result in very poor outcomes if environmental contingencies ever change (Navarro and Newell 2014). Navarro et al. found that, even in static environments, people typically begin a task under the assumption that contingencies are volatile before shifting toward the optimal maximizing strategy (see also Rakow et al. 2010). We suspect that the belief in a changing environment was also held by participants in our study—particularly those in the no discounting group—as progressively ignoring feedback runs counter to the belief that outcome contingencies may change. If an individual believes that cue-outcome contingencies may suddenly change, discounting feedback would guarantee a slower response to the change.

Individual variation in WMC, preferred learning style, and prior beliefs about task contingencies all have sensible theoretical connections to feedback discounting. While our study was not equipped to identify the causes of variation in feedback discounting, there are a number of prospects for future research investigating factors that might predict people's discounting behavior. Combining such an individual difference approach with measurement of neural activity is a particularly interesting avenue for further exploration. In addition, the form of the discounting function—how rapidly feedback discounting begins—could be studied in more detail by having participants complete multiple runs through a probabilistic learning environment such as ours, but with fewer blocks per run. Pooling data across common time points from different experimental runs would allow for more precise characterization of the onset of feedback discounting, which would be useful for further exploring the relationship between changes in the FRN and components of successful cognitive models. We leave these as targets for future research.

Feedback Discounting Unrelated to Learning Although we identified adaptive discounting of feedback in our task, which supported better learning, it is possible that people might discount feedback for other motivational reasons. For example, Bellebaum and Daum (2008) found that non-learners in their task failed to produce an FRN in the later stages of the task. Although this is usually associated with *better* task performance (i.e., acquiring the relevant cue-outcome contingencies), Bellebaum and Daum's non-learner participants failed to show this. A natural account for why non-learners do not show an FRN is perhaps the most extreme form of feedback discounting: simply "giving up" on the task and disengaging. The idea that unmotivated individuals might be disinclined to process feedback has some support from studies where individuals passively watched another person perform a learning task. In these scenarios, both passive observers and active learners show an FRN, but the amplitude is diminished for the passive observers (e.g., Bellebaum et al. 2010; Holroyd et al. 2009). The motivational facet of the FRN was highlighted by Gehring and Willoughby (2002). An interesting avenue for future research would be to investigate whether the difference in FRN for passive and active learning might also relate to stronger feedback discounting for passive observers. Dissociating the immediate relevance of learning (i.e., enhancing one's performance on a task) from the informational value of learning (i.e., supporting predictive accuracy) would clarify the role of motivation in determining feedback discounting.

Conclusions

We conducted a novel investigation of the relationship between feedback discounting and the FRN. We identified feedback discounting via model-based analysis of individual choice behavior, and related the performance of people who did and did not discount feedback to changes in the FRN during a probabilistic category learning task. Our results provided novel evidence of individual differences in feedback discounting that had not previously been identified in the literature. We also showed striking correspondence between feedback discounting, people's response strategies, and the amplitude of the FRN. We found feedback discounting to be associated with a rational maximizing-like strategy, which, in turn, produced no clear FRN. By contrast, people who did not discount feedback tended to respond via probability matching. These participants showed continuing processing of feedback information for the duration of the task. Analysis of the component ERPs elicited by correct and error feedback for both groups of participants supported the idea that discounting reflects a progressive weakening of the neural response to outcome feedback. For people who did not discount feedback, the neural response to receiving correct feedback reduced slightly over the course of the task, but remained steady for

error feedback, signifying the continuing importance of both feedback and prediction error in shaping performance. Our study showcases the utility of combining model-based analysis of behavior with neural measures of information processing. By seeking converging evidence from both domains, we are able to develop a more complete understanding of otherwise elusive psychological phenomena.

Acknowledgements We thank Kashmira Daruwalla and Maggie Webb for assistance during data collection.

Funding Information This research was supported by Australian Research Council Discovery Early Career Researcher Awards to David Sewell (DE140100772) and Stefan Bode (DE140100350).

Compliance with Ethical Standards

The study was approved by the Human Research Ethics Committee at the University of Melbourne and was conducted in accordance with the Declaration of Helsinki.

References

- Arbel, Y., & Wu, H. (2016). A neurophysiological examination of quality of learning in a feedback-based learning task. *Neuropsychologia*, *93*, 13–20.
- Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience*, *27*, 1823–1835.
- Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2010). It was not MY fault: event-related brain potentials in active and observational learning from feedback. *Cerebral Cortex*, *20*, 2874–2883.
- Blair, M., & Homa, D. L. (2005). Integrating novel dimensions to eliminate category exceptions: when more is less. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *31*, 258–271.
- Bland, A. R., & Schaefer, A. (2011). Electrophysiological correlates of decision making under varying levels of uncertainty. *Brain Research*, *1417*, 55–66.
- Bode, S., Bennett, D., Stahl, J., & Murawski, C. (2014). Distributed patterns of event-related potentials predict subsequent ratings of abstract stimulus attributes. *PLoS One*, *9*, e109070.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Castellan, N. J. (1973). Multiple-cue probability learning with irrelevant cues. *Organizational Behavior & Human Performance*, *9*, 16–29.
- Chase, H. W., Swainson, R., Durham, L., Benham, L., & Cools, R. (2011). Feedback-related negativity codes prediction error but not behavioral adjustment during probabilistic reversal learning. *Journal of Cognitive Neuroscience*, *23*, 936–946.
- Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *NeuroImage*, *35*(2), 968–978.
- Cohen, M. X., Wilmes, K. A., & van de Vijver, I. (2011). Cortical electrophysiological network dynamics of feedback learning. *Trends in Cognitive Sciences*, *15*, 558–566.
- Craig, S., & Lewandowsky, S. (2012). Whichever way you choose to categorize, working memory helps you learn. *Quarterly Journal of Experimental Psychology*, *65*, 439–464.

- Craig, S., & Lewandowsky, S. (2013). Working memory supports inference learning just like classification learning. *Quarterly Journal of Experimental Psychology*, *66*, 1493–1503.
- Craig, S., Lewandowsky, S., & Little, D. R. (2011). Error discounting in probabilistic category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *37*, 673–687.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21.
- Edgell, S. E. (1983). Delayed exposure to configural information in non-metric multiple-cue probability learning. *Organizational Behavior & Human Decision Processes*, *32*, 55–65.
- Edgell, S. E., & Morrissey, J. M. (1987). Delayed exposure to additional relevant information in nonmetric multiple-cue probability learning. *Organizational Behavior & Human Decision Processes*, *40*, 22–38.
- Edwards, W. (1961). Probability learning in 1000 trials. *Journal of Experimental Psychology*, *62*, 385–394.
- Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia*, *46*, 521–539.
- Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*, 107–140.
- Falkenstein, M., Hohnsbein, J., Hoonmann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography & Clinical Neurophysiology*, *78*, 447–455.
- Forstmann, B. U., Wagenmakers, E. J., Eichele, T., Brown, S., & Serences, J. T. (2011). Reciprocal relations between cognitive neuroscience and formal cognitive models: opposites attract? *Trends in Cognitive Sciences*, *15*, 272–279.
- Friedman, D., & Massaro, D. W. (1998). Understanding variability in binary and continuous choice. *Psychonomic Bulletin & Review*, *5*, 370–389.
- Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*, 2279–2282.
- Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385–390.
- Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, *71*, 148–154.
- Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: the feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*, 905–912.
- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709.
- Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., & Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cognitive, Affective, & Behavioral Neuroscience*, *9*, 59–70.
- Ichikawa, N., Siegle, G. J., Dombrowski, A., & Ohira, H. (2010). Subjective and model-estimated reward prediction: association with the feedback-related negativity (FRN) and reward prediction error in a reinforcement learning task. *International Journal of Psychophysiology*, *78*, 273–283.
- Kalish, M. L., Newell, B. R., & Dunn, J. C. (2017). More is generally better: higher working memory capacity does not impair perceptual category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *43*, 503–514.
- Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In R. M. Church & B. A. Campbell (Eds.), *Punishment and aversive behavior* (pp. 279–296). New York: Appleton-Century-Crofts.
- Krigolson, O. E., Pierce, L. J., Holroyd, C. B., & Tanaka, J. W. (2009). Learning to become an expert: reinforcement learning and the acquisition of perceptual expertise. *Journal of Cognitive Neuroscience*, *21*, 1834–1841.
- Kruschke, J. K., & Johansen, M. K. (1999). A model of probabilistic category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *25*, 1083–1119.
- Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *Journal of Experimental Psychology: General*, *135*, 162–183.
- Lewandowsky, S. (2011). Working memory capacity and categorization: individual differences and modeling. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *37*, 720–738.
- Lewandowsky, S., Yang, L.-X., Newell, B. R., & Kalish, M. L. (2012). Working memory does not dissociate between different perceptual categorization tasks. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *38*, 881–904.
- Little, D. R., & Lewandowsky, S. (2012). Multiple-cue probability learning. In N. Seel (Ed.), *Encyclopedia of the learning sciences* (pp. 2386–2388). New York: Springer.
- Little, J. L., & McDaniel, M. A. (2015). Individual differences in category learning: memorization versus rule abstraction. *Memory & Cognition*, *43*, 283–297.
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 1–14.
- Love, B. C., & Gureckis, T. M. (2007). Models in search of a brain. *Cognitive, Affective, & Behavioral Neuroscience*, *7*, 90–108.
- Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 103–189). New York: Wiley.
- Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge: MIT Press.
- Luft, C. D. B. (2014). Learning from feedback: the neural mechanisms of feedback processing facilitating better performance. *Behavioral Brain Research*, *261*, 356–368.
- Luque, D., López, F. J., Marco-Pallares, J., Cámara, E., & Rodríguez-Fomells, A. (2012). Feedback-related brain potential activity complies with basic assumptions of associative learning theory. *Journal of Cognitive Neuroscience*, *24*, 794–808.
- McDaniel, M. A., Cahill, M. J., Robbins, M., & Wiener, C. (2014). Individual differences in learning and transfer: stable tendencies for learning exemplars versus abstracting rules. *Journal of Experimental Psychology: General*, *143*, 668–693.
- Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a “generic” neural system for error detection. *Journal of Cognitive Neuroscience*, *9*, 788–798.
- Navarro, D. J., & Newell, B. R. (2014). Information versus reward in a changing world. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th annual conference of the cognitive science society* (pp. 1054–1059).
- Navarro, D. J., Newell, B. R., & Schulze, C. (2016). Learning and choosing in an uncertain world: an investigation of the explore-exploit dilemma in static and dynamic environments. *Cognitive Psychology*, *85*, 43–77.
- Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. H. (2004a). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience & Biobehavioral Reviews*, *28*, 441–448.
- Nieuwenhuis, S., Yeung, N., Holroyd, C. B., Schurger, A., & Cohen, J. D. (2004b). Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. *Cerebral Cortex*, *14*, 741–747.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

- Pietschmann, M., Simon, K., Endrass, T., & Kathmann, N. (2008). Changes of performance monitoring with learning in older and younger adults. *Psychophysiology*, *45*, 559–568.
- Potts, G. F., Martin, L. E., Kamp, S.-M., & Donchin, E. (2011). Neural response to action and reward prediction errors: comparing the error-related negativity to behavioral errors and the feedback-related negativity to reward prediction violations. *Psychophysiology*, *48*, 218–228.
- Rakow, T., Newell, B. R., & Zougkou, K. (2010). The role of working memory in information acquisition and decision making: lessons from the binary prediction task. *Quarterly Journal of Experimental Psychology*, *63*, 1335–1360.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
- Sailer, U., Fischmeister, F. P. S., & Bauer, H. (2010). Effects of learning on feedback-related brain potentials in a decision-making task. *Brain Research*, *1342*, 85–93.
- Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*, 213–235.
- Schiffers, A.-M., Siletti, K., Waszak, F., & Yeung, N. (2017). Adaptive behaviour and feedback processing integrate experience and instruction in reinforcement learning. *NeuroImage*, *146*, 626–641.
- Schulze, C., & Newell, B. R. (2016). Taking the easy way out? Increasing implementation effort reduces probability maximizing under cognitive load. *Memory & Cognition*, *44*, 806–818.
- Sewell, D. K., & Lewandowsky, S. (2012). Attention and working memory capacity: insights from blocking, highlighting, and knowledge restructuring. *Journal of Experimental Psychology: General*, *141*, 444–469.
- Shanks, D. R., Tunney, R. J., & McCarthy, J. D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making*, *15*, 233–250.
- Stahl, J. (2010). Error detection and the use of internal and external error indicators: an investigation of the first-indicator hypothesis. *International Journal of Psychophysiology*, *77*, 43–52.
- Takeda, Y., Yamanaka, K., & Yamamoto, Y. (2008). Temporal decomposition of EEG during a simple reaction time task into stimulus- and response-locked components. *NeuroImage*, *39*, 742–754.
- Walsh, M. M., & Anderson, J. R. (2011). Modulation of the feedback-related negativity by instruction and experience. *Proceedings of the National Academy of Science*, *108*, 19048–19053.
- Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews*, *36*, 1870–1884.
- Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, *24*, 6258–6264.