Effects of feedback reliability on feedback-related brain activity: A feedback valuation account

Benjamin Ernst¹ · Marco Steinhauser¹

Published online: 6 April 2018
© Psychonomic Society, Inc. 2018

Abstract

Adaptive decision making relies on learning from feedback. Because feedback sometimes can be misleading, optimal learning requires that knowledge about the feedback's reliability be utilized to adjust feedback processing. Although previous research has shown that feedback reliability indeed influences feedback processing, the underlying mechanisms through which this is accomplished remain unclear. Here we propose that feedback processing is adjusted by the adaptive, top-down valuation of feedback. We assume that unreliable feedback is devalued relative to reliable feedback, thus reducing the reward prediction errors that underlie feedback-related brain activity and learning. A crucial prediction of this account is that the effects of feedback reliability are susceptible to contrast effects. That is, the effects of feedback reliability should be enhanced when both reliable and unreliable feedback are experienced within the same context, as compared to when only one level of feedback reliability is experienced. To evaluate this prediction, we measured the event-related potentials elicited by feedback in two experiments in which feedback reliability was varied either within or between blocks. We found that the fronto-central valence effect, a correlate of reward prediction errors during reinforcement learning, was reduced for unreliable feedback. But this result was obtained only when feedback reliability was varied within blocks, thus indicating a contrast effect. This suggests that the adaptive valuation of feedback is one mechanism underlying the effects of feedback reliability on feedback processing.

Keywords Feedback validity · Feedback reliability · Decision making · Feedback-related negativity · P3

Learning from feedback is essential for optimal decision making. However, feedback is not always guaranteed to be valid, and thus can be misleading. If the same decision consistently leads to the same feedback (as in deterministic learning tasks), this feedback is highly reliable. However, if the same decision leads to varying feedback (as in probabilistic learning tasks), this feedback is less reliable. Advance knowledge of the reliability of feedback should allow humans to adjust their feedback processing in order to minimize the negative impact of invalid feedback and maximize the long-term outcome. Prior research has provided evidence that feedback processing is affected by feedback reliability. Unreliable feedback has been shown to elicit weaker feedback-related brain activity than reliable feedback (Ernst & Steinhauser, 2017; Schiffer, Siletti, Waszak, & Yeung, 2017; Walentowska, Moors, Paul,

& Pourtois, 2016). Although it has been proposed that this reflects the involvement of top-down control—that is, a modulation of experience-driven learning by high-level cognitive processes—the underlying mechanism through which this is accomplished remains unclear. In the present study, we introduce the idea that a core mechanism underlying these effects is the adaptive, top-down valuation of feedback. We propose that explicit knowledge about the reliability of feedback causes a devaluation of unreliable feedback relative to reliable feedback, which results in a reduced impact of unreliable feedback on learning.

Research on feedback in decision making has mainly considered two components of the human event-related potential (ERP) as quantifying distinct aspects of feedback processing: the feedback-related negativity (FRN; Miltner, Braun, & Coles, 1997) and the feedback P3 (for a review, see San Martín, 2012). The FRN refers to a negativity at fronto-central electrodes that peaks around 250–300 ms after feedback onset and is more pronounced after negative than after positive feedback (for reviews, see San Martín, 2012; Walsh & Anderson, 2012). The reinforcement-learning theory of the error-related negativity (Holroyd & Coles, 2002) assumes that

✉ Benjamin Ernst
Benjamin.Ernst@ku.de

¹ Allgemeine Psychologie, Catholic University Eichstätt-Ingolstadt, D-85072 Eichstätt, Germany

this component reflects a negative reward prediction error that is generated when the outcome expectancy associated with feedback (i.e., the feedback's value) is smaller than a prior outcome expectancy derived from the stimulus and response. In recent years, this theory has experienced modifications—for example, by attributing the amplitude difference between negative and positive feedback to a positivity on positive feedback (the reward positivity; e.g., Holroyd, Pakzad-Vaezi, & Krigolson, 2008). However, the link between this effect and the reward prediction error has remained a central assumption, and converging evidence suggests that fronto-central valence effects (FRN or reward positivity) can indeed be used as an index of the use of feedback in reinforcement learning (San Martín, 2012; Walsh & Anderson, 2012).

In contrast, the feedback P3 is a positivity that peaks at posterior electrodes between 300 and 600 ms after feedback onset and belongs to the broader class of posterior P3 components. This activity is typically assumed to reflect controlled or endogenous processing, since it can be tied to attention and working memory updating (Donchin & Coles, 1988; Polich, 2007). A recent theory of the P3 suggests that this attentional response is affected by the motivational significance of a stimulus (Ferdinand, Mecklinger, Kray, & Gehring, 2012; Nieuwenhuis, 2011; Nieuwenhuis, Aston-Jones, & Cohen, 2005), which is in line with studies showing that the feedback P3 is associated with reward magnitude and, thus, feedback value (Bellebaum, Poleszki, & Daum, 2010; Yeung & Sanfey, 2004; for a review, see San Martín, 2012).

Effects of feedback reliability on these components have already been observed in a number of previous studies (Ernst & Steinhauser, 2017; Schiffer et al., 2017; Walentowska et al., 2016). These experiments have in common that participants had to repeatedly choose among a set of stimuli that were associated with either monetary reward or loss. By utilizing feedback, participants had to learn via trial and error which choice was more likely to be associated with a reward—that is, which response was correct for a given set of stimuli. Crucially, the feedback's reliability varied between blocks, and participants were explicitly informed of this. In blocks with unreliable feedback, the feedback was highly likely to be invalid, occasionally indicating that a correct response had been incorrect and vice versa (i.e., more probabilistic feedback), whereas in blocks with reliable feedback, the feedback was highly likely to be valid (i.e., more deterministic feedback). Whereas a pronounced fronto-central valence effect emerged after reliable feedback, the effect was reduced or absent after unreliable feedback in most (Ernst & Steinhauser, 2017; Schiffer et al., 2017; Walentowska et al., 2016, Exp. 1), but not all (Walentowska et al., 2016, Exp. 2), experiments. This suggests that prior knowledge about feedback reliability affected the reinforcement-learning system. Similarly, the feedback P3 was reduced after unreliable feedback (Schiffer et al., 2017; Walentowska et al., 2016),

suggesting that prior knowledge also affected controlled feedback processing.

Mechanisms underlying the effects of feedback reliability

So far, it is still unclear how this modulation of reinforcement learning is realized. One plausible mechanism could be a *top-down control* process that gates or suppresses the processing of unreliable feedback (e.g., Ernst & Steinhauser, 2017; Walentowska et al., 2016). Advance knowledge about feedback reliability could result in feedback processing being configured in such a way that the generation of reward prediction errors following unreliable feedback stimuli is reduced or abolished. This account has received support from the finding that instructions can affect reinforcement learning through top-down control (e.g., Doll, Hutchison, & Frank, 2011; Doll, Jacobs, Sanfey, & Frank, 2009; Li, Delgado, & Phelps, 2011). These studies have shown that explicit knowledge about contingencies can bias learning from feedback, possibly via altered activity in the basal ganglia. Specifically, Li et al. suggested that the dorsolateral prefrontal cortex dynamically adjusts responses to feedback in the basal ganglia according to the usefulness of the feedback. In line with this suggestion, Schiffer et al. (2017) interpreted the reduced fronto-central valence effect for unreliable feedback as indicating that the perceived informativeness of feedback can affect prediction error processing.

However, top-down suppression is not the only mechanism that could explain effects of feedback reliability. These results are also compatible with an alternative account in which knowledge about feedback reliability affects feedback processing via adaptive *feedback valuation*. More specifically, we propose that the subjective value of feedback, which reflects the outcome expectation induced by the feedback, is lower for unreliable than for reliable feedback. This idea is based on evidence that representations of stimulus values in the dopaminergic reward system can be adjusted in accordance with context information (e.g., uncertainty, costs, or possible alternatives; Plassmann, O'Doherty, Shiv, & Rangel, 2008; Rushworth & Behrens, 2008), as well as with goals and instructions (e.g., De Araujo, Rolls, Velazco, Margot, & Cayeux, 2005; Grabenhorst & Rolls, 2010; Grabenhorst, Rolls, & Bilderbeck, 2007; Hare, Camerer, & Rangel, 2009). In fact, the existence of top-down valuation of feedback can already be derived from the observation that feedback stimuli without any intrinsic value (e.g., colors: Ernst & Steinhauser, 2015; tones: Miltner et al., 1997) can evoke valence effects in feedback-related brain activity after a positive or negative valence has been assigned them via instruction. Explicit knowledge of feedback reliability could already affect this feedback valuation at the time of

instruction, resulting in different values for reliable and unreliable feedback stimuli. As a result, unreliable positive (or negative) feedback stimuli might have a lower positive (or negative) subjective value than the corresponding reliable feedback stimuli, similar to a copper coin having a lower value than a gold coin.¹ Because the computation of prediction error is based on outcome expectations derived from feedback value, this devaluation of unreliable feedback would result in attenuated positive and negative reward prediction errors that would manifest in a reduced fronto-central valence effect.

A testable prediction that can be derived from this account is that the valuation should take into account the presence or absence of possible alternative outcomes. Numerous studies have shown that instrumental learning of a response is affected not only by the absolute value of its associated outcome, but also by its relative value—that is, by the presence of alternative outcomes with different values (for an overview, see Flaherty, 1999). A classic example of such a *contrast effect* is the finding that rats reduced their running speed toward a given number of food pellets when they had previously experienced a larger number of pellets, but increased their running speed toward the same amount when they had previously experienced a lower one (Crespi, 1942; Flaherty, 1982; Maxwell, Calef, Murray, Shepard, & Norville, 1976). Comparable effects have been obtained for feedback-related brain activity. The fronto-central valence effect for a given outcome has been shown to depend on the magnitudes of the alternative outcomes (Bellebaum et al., 2010; Holroyd, Larsen, & Cohen, 2004). Activity to the same outcome (0 cents) differed depending on the context created by alternative outcomes (either gains or losses) indicated by an instruction at the beginning of a block (Holroyd et al., 2004) or by a cue at the beginning of a trial (Bellebaum et al., 2010). On the basis of our feedback valuation account, we predicted that similar contrast effects should be obtained for the effects of feedback reliability. Unreliable feedback should be devalued more strongly if participants also experience more reliable feedback in a given context, whereas the unreliable feedback should be devalued to a lesser degree in the absence of reliable feedback.

The present study

To test this prediction, we considered ERPs in a simple decision task that we have employed in previous studies (Ernst & Steinhauser, 2012, 2015, 2017). This task differs from a typical reinforcement-learning paradigm insofar as the items are presented only twice. In a learning phase, participants had to

guess which of two stimuli was correct (i.e., was associated with reward) and were provided with immediate feedback. In a test phase, they could earn a reward by responding to the same stimulus pair. Of interest were only the ERPs to the feedback during the learning phase. The advantage of this learn–test paradigm is that the measured feedback-related brain activity was not influenced by prior learning and previously acquired outcome expectations. In this way, we could exclude that any fronto-central valence effect to feedback on a given trial was influenced by the reliability of the feedback for this stimulus pair on previous trials. Feedback reliability was manipulated on two levels: It could be either reliable (i.e., the predicted reward in the test phase on 100% of trials) or unreliable (i.e., the predicted reward on 75% of trials). In contrast to the previous studies, the feedback's reliability was not given by instruction prior to a block of trials, but rather was signaled by the feedback itself: While one feature of the feedback stimulus indicated valence (positive, negative), a different stimulus feature indicated feedback reliability (reliable, unreliable). This allowed us to manipulate feedback reliability across trials and to present trials with different feedback reliabilities within a block. Moreover, assigning different levels of feedback reliability to different feedback stimuli should enable participants to devalue unreliable feedback stimuli already during instruction.

In Experiment 1, reliable and unreliable feedback was randomly intermixed within each block of trials. We predicted a strong contrast effect in this case. Because participants were confronted with two levels of feedback reliability, the unreliable feedback should be strongly devalued relative to the reliable feedback. This should lead to a smaller fronto-central valence effect on trials with unreliable feedback, and possibly also to a reduced feedback P3 and impaired learning. In Experiment 2, we used the same feedback stimuli but presented the reliable and unreliable feedback in different blocks. Here we predicted that the unreliable feedback would be less devalued because, within the context of these blocks, no alternative level of feedback reliability was expected or experienced. This should lead to a smaller difference in fronto-central valence effects between reliable and unreliable feedback in this experiment. Crucially, these predictions are unique for our devaluation account. If the effects of feedback reliability were driven by top-down suppression of feedback processing, we would expect to find even larger effects of feedback reliability in Experiment 2. Because feedback reliability was already predictable before the feedback was presented when feedback reliability was blocked, this should facilitate the recruitment of top-down control and, thus, should lead to larger effects than in Experiment 1, in which the feedback reliability was unpredictable before feedback onset.

¹ Please note that “feedback value” in our account is used in accordance with temporal difference-learning frameworks and directly corresponds to the (reward) value associated with a feedback stimulus in previous studies (Holroyd & Coles, 2002; Schultz, Dayan, & Montague, 1997). It should not be equated with learning rates or action values.

Experiment 1

Method

Participants Forty participants (29 female, 11 male) between 18 and 33 years of age (mean = 21.5 years) with normal or corrected-to-normal vision participated in this experiment. The participants were recruited at the Catholic University of Eichstätt and received course credit or a base fee of €8 per hour, as well as a performance-dependent bonus (mean: €2.45, range: €0–€4.90). The data from five participants (all female) were excluded from further analysis because their test performance was close to chance, and the data of another female participant were excluded because of excessive artifacts in the electroencephalogram (EEG). As a result, the data from 34 participants (23 female, 11 male) were included in the analysis. The study was conducted in accordance with institutional guidelines, and informed consent was acquired from all participants.

Stimulus material We used a set of 384 Chinese and Chinese-looking characters as test stimuli. Each character had a height of 2.3° visual angle and a width of 1.8° visual angle at a viewing distance of 70 cm, and was presented in white color on a black background. For each participant, the stimuli were randomly divided into 32 sets of six items, with each item consisting of a pair of characters. Within each item, each character was presented at 0.5° of visual angle to the left or right of screen center. The position of each character (left or right) was randomly determined for each presentation.

The feedback stimuli in learning trials were blue or yellow circles (see Fig. 1d) with a diameter of 2.5°, presented centrally on the screen. The saturation of the stimuli indicated the reliability of the feedback: The colors of reliable feedback stimuli were more saturated (RGB: yellow, 255, 255, 0; blue, 61, 83, 255) than the colors of unreliable feedback stimuli (RGB: yellow, 240, 240, 186; blue, RGB: 136, 183, 255).

Design and procedure Each item was presented only once in a learning trial and once in a test trial. In learning trials, participants had to guess which of the two characters (left or right) would be associated with a reward in the subsequent test trial. The procedure of a learning trial is illustrated in Fig. 1a. After presentation of an item, participants had up to 5 s to choose one character by pressing one of two keys on a standard keyboard: Pressing “X” with left index finger indicated a choice of the character on the left side, and pressing “M” with the right index finger indicated a choice of the character on the right side. Following the response (or after 5 s), the item disappeared and was replaced by a fixation cross that was presented for 400–600 ms (uniformly distributed). Then the feedback stimulus was presented for 600 ms. If no response was given within the 5-s time limit, the word VERPASST (“miss”) was presented instead of the feedback.

Participants were informed that a miss was always associated with a loss of 10 cents. Following a blank screen presented for 2,000 ms, a new trial began or (if the end of a block had been reached) the test phase started with a prompt. This prompt informed about the upcoming test phase and disappeared upon a self-paced keypress.

Participants were instructed that the feedback color indicated the correctness of the response. The color–valence mapping (e.g., blue is correct) was counterbalanced across participants. Participants were further informed that feedback stimuli with paler colors indicated that this feedback was invalid in 25% of cases, whereas the feedback provided by stimuli with clear colors was never invalid. Feedback reliability was randomized over the experiment—that is, each block contained a mixture of reliable and unreliable feedback. In this way, the feedback reliability could be known only at the moment the feedback stimulus was presented.

In the test phase, the character pairs from the preceding learning phase were presented again in randomized order. Moreover, although the same character pairs were used as in the learning phase, the spatial positions (left/right) of the two characters were switched relative to the learning phase for half of the items. This was done to ensure that participants learned the stimuli rather than the response side. The procedure for the test trials (see Fig. 1b) was the same as that for the learning trials, with two exceptions: First, the feedback consisted not of a colored circle but of a number indicating the number of cents gained (“+ 3”) or lost (“− 3”) in a trial. This feedback was always valid; that is, this amount was added to a participant’s bonus money. Second, the blank screen following feedback was presented only for 1,000 ms until a new trial or the next learning phase began.

Each participant was tested in an individual session. After the electrode cap had been fitted, participants were seated comfortably in a dimly lit, noise-proofed room. They were made familiar with the task by completing three training blocks. The main part of the experiment consisted of 32 blocks, each of which comprised a learning phase and a test phase, in which six items were first learned and then tested (see Fig. 1c). The order of items was randomized within each phase. Between phases and blocks was a short, self-paced interruption in which information about the upcoming phase or block was presented, allowing participant to take a short pause and prepare for the next phase or block. Moreover, a longer break occurred after 12 and after 24 blocks.

Electrophysiological recordings The EEG was recorded using a BIOSEMI Active-Two system (BioSemi, Amsterdam, The Netherlands) with 32 Ag–AgCl electrodes for channels Fp1, AF3, F3, FC1, FC5, C3, T7, CP1, CP5, P3, P7, P9, PO3, O1, Oz, Pz, CPz, Fp2, AF4, AFz, Fz, F4, FC2, FC6, FCz, Cz, C4, T8, CP6, CP2, P4, P8, and O2, as well as for the left and

Fig. 1 **a, b** Sequences of events in a typical trial during the learning phase (**a**) and the test phase (**b**). **c** Each block consisted of a learning phase that included six stimulus pairs that were tested in the subsequent test phase. **d** Feedback valence and feedback reliability were indicated by the color and

saturation of the feedback stimulus, respectively. Please note that the assignment of colors to feedback valences was counterbalanced across participants

right mastoid. The common mode sense and driven right leg electrodes were used as the reference and ground electrodes, respectively. The vertical and horizontal electrooculogram (EOG) was recorded from electrodes above and below the right eye and at the outer canthi of both eyes. All electrodes were re-referenced offline to the averaged mastoids. The EEG and EOG were recorded continuously at a sampling rate of 512 Hz.

Data analysis The EEG data from learning trials were analyzed using EEGLAB version 11 (Delorme & Makeig, 2004) and custom routines written in MatLab 8 (The Mathworks, Natick, MA). The data were band-pass-filtered excluding activity below 0.5 Hz and above 35 Hz. Epochs were extracted ranging from 200 ms before to 800 ms after feedback onset, and baseline activity was removed by subtracting the average voltage from an interval between 100 and 0 ms before feedback onset. Epochs for which activity at any electrode (except frontal electrodes AF7, Fp1, Fpz, Fp2, and AF8) was more than 100 μ V above or below the baseline-corrected mean were excluded. Ocular artifacts were corrected using independent component analysis (Delorme & Makeig, 2004). Finally, the epochs were averaged separately for each condition of interest.

We quantified fronto-central activity by calculating the mean amplitudes for each condition and each participant in

the time range of the FRN peak (250–300 ms; e.g., Walentowska et al., 2016) at electrode FCz. Because previous studies had frequently used peak-to-peak measures (e.g., Ernst & Steinhauser, 2012, 2015, 2017; Yeung & Sanfey, 2004), we additionally quantified this activity by determining the amplitude of the FRN peak (most negative peak within 200 and 400 ms after feedback onset) and then subtracting the amplitude of the P2 peak (most positive peak between 150 ms after feedback onset and the FRN peak), separately for each condition and each participant at electrode FCz. If there was no negative peak (i.e., if the minimum was at the borders of the 200- to 400-ms time window), the FRN amplitude was scored as 0 μ V. For an additional analysis of the P2, we quantified this component for each condition and each participant as the mean amplitude in the time range around the P2 peak (175–225 ms) at FCz. To control for possible differences in P3 latency, we quantified the P3 as the maximum peak in the ERP in the time range of 300–800 ms after stimulus onset at Pz (Pontifex, Hillman, & Polich, 2009; Sailer, Fischmeister, & Bauer, 2010). To identify a clear P3 peak, we first applied a 4-Hz low-pass filter (Foti, Weinberg, Bernat, & Proudfit, 2015) before averaging the ERPs over participants and conditions for the P3 analysis. Each of these ERP measures was analyzed using an analysis of variance (ANOVA) with repeated measures on the variables feedback reliability (reliable, unreliable) and feedback valence (positive, negative).

We quantified learning performance as the percentage of correct choices in the test phase. A choice was defined as “correct” when the feedback in the learning phase indicated that it would be rewarded in the test phase, and as “incorrect” when the feedback indicated that it would lead to a loss. Please note that, with this definition of “correct” and “incorrect,” it was possible in trials with unreliable feedback for a choice in the test phase to be registered as “correct” although the participant lost money. If we had used the actual outcome in the test phase to define correctness, our results would have been biased by the presence of trials with invalid feedback in the unreliable-feedback condition. As with the ERPs, learning performance was analyzed using an ANOVA with repeated measures on the variables feedback reliability (reliable, unreliable) and feedback valence (positive, negative).

Results

Learning performance

As expected, we obtained main effects of feedback reliability, $F(1, 33) = 40.6, p < .001$, and feedback valence, $F(1, 33) = 39.0, p < .001$, but no interaction, $F < 1$. The mean performance for each condition is depicted in Fig. 2. After negative feedback ($M = 68.6\%, SD = 13.7$), learning was impaired relative to positive feedback ($M = 79.0\%, SD = 11.4$). More importantly, unreliable feedback ($M = 70.3\%, SD = 12.7$) impaired performance relative to reliable feedback ($M = 77.4\%, SD = 12.4$). These behavioral results clearly corroborate that learning from feedback was impaired when the feedback was unreliable.

ERP results

Fronto-central valence effect Our analysis of fronto-central amplitudes during the learning phase revealed no main effect of feedback valence, $F(1, 33) = 0.58, p = .45$, but a main effect of feedback reliability, $F(1, 33) = 9.94, p < .001$. Importantly, this main effect was qualified by a significant interaction between the two variables, $F(1, 33) = 8.80, p < .01$ (see Fig. 3a, c, e). Subsequent planned contrasts confirmed that whereas reliable feedback yielded a significant fronto-central valence effect, with more negative amplitudes after negative ($M = -1.86 \mu V, SD = 4.51$) than after positive ($M = -0.53 \mu V, SD = 4.98$) feedback, $t(33) = 2.17, p < .05$, this effect was almost inverted after unreliable feedback (positive feedback: $M = -2.94 \mu V, SD = 5.06$; negative feedback: $M = -2.19 \mu V, SD = 4.78$), $t(33) = 1.89, p = .07$. Additional contrasts revealed that although there was no significant amplitude difference between reliable and unreliable negative feedback, $t < 1$, reliable positive feedback was associated with a less pronounced negativity than was unreliable positive feedback, $t(33) = 3.88, p < .0001$.

Fig. 2 Behavioral performance in the test phase as a function of feedback valence and feedback reliability in the learning phase of Experiment 1. Please note that “correct” performance was defined as choosing the stimulus that had been indicated to be correct in the learning phase (which was not the same as being rewarded in the test phase). FB = feedback

In addition, we also analyzed the fronto-central valence effect using a peak-to-peak difference measure reflecting the difference between the negative N2 peak and the preceding P2 peak. Although there was no main effect of feedback valence, $F(1, 33) = 2.09, p < .16$, we did obtain a main effect of feedback reliability, $F(1, 33) = 10.9, p < .01$, as well as a marginally significant interaction between these variables, $F(1, 33) = 2.93, p = .09$. Planned contrasts revealed a marginally significant valence effect for reliable feedback (positive feedback: $M = -6.85 \mu V, SD = 3.21$; negative feedback: $M = -7.84 \mu V, SD = 4.15$), $t(33) = 1.97, p = .06$, which was absent for unreliable feedback (positive feedback: $M = -8.57 \mu V, SD = 4.76$; negative feedback: $M = -8.45 \mu V, SD = 4.33$), $t(33) = 0.31, p = .76$. Although these results are only marginally significant, they reveal a pattern similar to that for the mean amplitudes.

The previous analysis suggests that peak-to-peak measures are less sensitive to the effects of feedback reliability on the fronto-central valence effect. Visual inspection showed that this is due to the fact that the P2 is already influenced by feedback reliability. Indeed, analyzing the mean amplitudes in the P2 time window revealed that, although there was no main effect of feedback valence, $F < 1$, there was a main effect of feedback reliability, $F(1, 33) = 5.93, p < .05$, which was qualified by a significant interaction, $F(1, 33) = 7.42, p < .05$. Contrasts revealed that, on the one hand, no significant valence effect was obtained for reliable feedback, $t(33) = 1.62, p = .12$, but unreliable negative feedback led to a slightly larger P2 than did unreliable positive feedback; this difference, however, reached only marginal significance, $t(33) = 1.81, p = .08$. On the other hand, while negative feedback led to equally pronounced P2 amplitudes in both feedback reliability conditions (reliable feedback: $M = 3.68 \mu V, SD = 4.42$; unreliable feedback: $M = 3.66 \mu V, SD = 4.31$), $t < 1$, reliable positive feedback led to a significantly larger P2 amplitude ($M = 4.46 \mu V, SD = 4.65$) than did unreliable positive feedback ($M = 3.04 \mu V, SD = 3.64$),

Fig. 3 **a, b** Grand average feedback-locked ERPs at FCz (**a**) and Pz (**b**) in the learning phase of Experiment 1, as a function of feedback valence and feedback reliability. **c, d** Topographies of the difference waveforms for the typical time windows of the fronto-central valence effect (250–300 ms; **b**) and the P3 (300–600 ms; **d**) in Experiment 1. The upper row shows the topographies for the difference between positive and negative feedback for both feedback reliability conditions. The lower row shows

the topographies for the difference between reliable and unreliable feedback for both feedback valence conditions. **e** Mean amplitudes in the time interval 250–300 ms at FCz as a function of feedback valence and feedback reliability. **f** Maximum amplitudes in the time interval 300–800 ms at Pz as a function of feedback valence and feedback reliability. Bars represent standard errors of the means. FB = feedback

$t(33) = 0.31, p = .76$. This pattern suggests that our manipulations was already influencing earlier feedback processing.

P3 amplitude As we expected, the analysis of P3 amplitudes revealed a significant main effect of feedback reliability, $F(1, 33) = 5.08, p < .05$, as well as a significant main effect of feedback valence, $F(1, 33) = 4.74, p < .05$, but no significant interaction, $F < 1.56$. As is depicted in Fig. 3b and f, P3 amplitudes were attenuated by unreliable feedback ($M = 9.06 \mu V, SD = 4.53$) relative to reliable feedback ($M = 9.98$

$\mu V, SD = 4.68$) and by negative feedback ($M = 9.11 \mu V, SD = 4.43$) relative to positive feedback ($M = 9.94 \mu V, SD = 4.70$). This result demonstrates that unreliable feedback leads to a reduced P3 amplitude.

Discussion

In this first experiment we manipulated feedback reliability within blocks and found, in line with our expectations, that unreliable feedback was associated with impaired learning, a

reduced feedback P3, and a strongly reduced fronto-central valence effect. The latter result particularly is in line with the proposed devaluation account, which assumes that the two feedback stimuli for unreliable feedback were devalued relative to the feedback stimuli for reliable feedback, thus leading to a decreased reward prediction error. However, this pattern could also reflect that processing of unreliable feedback was suppressed by a top-down control process. Because the feedback reliability was not predictable, in this case one would have to additionally assume that top-down control can be flexibly triggered by the feedback stimulus itself.

In the next experiment, we aimed to directly discriminate between the two accounts by holding feedback reliability constant within blocks. For this experiment, the feedback valuation account would predict that the effects of feedback reliability on the fronto-central valence effect should be reduced. Under these conditions, experiencing only one level of feedback reliability should lead to decreased devaluation of the unreliable feedback relative to reliable feedback, due to the absence of a contrast effect. In contrast, the top-down control account would predict that varying feedback reliability across blocks would lead to an increased effect of feedback reliability. This prediction can be derived because feedback reliability was now predictable before stimulus onset, thus facilitating the recruitment of top-down control.

Experiment 2

Method

Participants Thirty-seven participants (34 female, three male) between 18 and 26 years of age (mean 21.4 years) with normal or corrected-to-normal vision participated in Experiment 2. The participants were again recruited at Catholic University of Eichstätt, but we ensured that they had not participated in the first experiment. Again, participants received either course credit or a base fee of €8 per hour, as well as a performance-dependent bonus (mean: €2.97, range: €0.70–€4.90). The data from three female participants were excluded from further analysis due to excessive artifacts in the EEG. As a result, the data from 34 participants (31 female, three male) were analyzed. This experiment was conducted in accordance with institutional guidelines, and informed consent was acquired from all participants.

Design and procedure We utilized the same stimuli, tasks, design, and procedure as in Experiment 1, with two exceptions: First, feedback reliability remained the same within each block of trials. That is, in each learning phase, the feedback was either always reliable or always unreliable. Participants were informed about the reliability at the beginning of each learning phase. The feedback reliability switched

after every fourth block, and half of the participants started with reliable feedback, whereas the other half started with unreliable feedback. Reliability was additionally indicated by the saturation of the feedback color, as in the previous experiment. Second, we asked participants to write a short statement after the experiment about how they had dealt with unreliable feedback. We included this question for Experiment 2 because the majority of the participants in Experiment 1 had claimed during debriefing that they had ignored the information about feedback reliability and had learned similarly from both unreliable and reliable feedback.

Electrophysiological recordings and data analysis The EEG data for Experiment 2 were collected using 64 Ag–AgCl electrodes (instead of 32 electrodes, as in Exp. 1). To ensure comparability between the experiments, we analyzed data only from the 32 electrode positions used in Experiment 1.

Results

Learning performance

We conducted the same analysis of learning performance as in Experiment 1, and again obtained main effects of feedback reliability, $F(1, 33) = 23.8$, $p < .001$, and feedback valence, $F(1, 33) = 50.8$, $p < .001$. However, we also obtained a marginally significant interaction between the two variables, $F(1, 33) = 3.10$, $p = .09$, indicating that the effect of feedback reliability on performance was more pronounced for negative than for positive feedback (see Fig. 4). However, even for positive feedback, performance was significantly impaired for unreliable as compared to reliable feedback, $t(33) = 2.40$, $p < .05$ [same contrast for negative feedback: $t(33) = 4.13$, $p < .0001$]. These results not only replicated the main effects found in Experiment 1, they also stand in contrast to the statements provided by the participants. Although the results showed that unreliability attenuated their learning from

Fig. 4 Behavioral performance in the test phase as a function of feedback valence and feedback reliability in the learning phase for Experiment 2. Bars represent standard errors of the means

feedback, 79.4% of the participants stated that they had ignored this information and had not treated reliable and unreliable feedback differently. Interestingly, all effects reported here and in the following analysis were also obtained when we analyzed only the data from this subset of participants.

ERP results

Fronto-central valence effect As is suggested by the ERPs shown in Fig. 5, unreliable feedback did not have the same

effect on the fronto-central valence effect in Experiment 2 as in the previous experiment. This observation is confirmed by our analysis of fronto-central amplitudes, which yielded a marginally significant main effect of feedback reliability, $F(1, 33) = 3.71, p = .06$, and a main effect of feedback valence, $F(1, 33) = 10.5, p < .005$, but in contrast to Experiment 1, no significant interaction, $F < 1$. Planned contrasts showed significant fronto-central valence effects for both reliable feedback (positive feedback: $M = -1.19 \mu V, SD = 4.68$; negative feedback: $M = -2.61 \mu V, SD = 4.24$), $t(33) = 2.90, p < .01$,

Fig. 5 a, b Grand average feedback-locked ERPs at FCz (a) and Pz (b) in the learning phase of Experiment 2, as a function of feedback valence and feedback reliability. c, d Topographies of the difference waveforms for the typical time windows of the fronto-central valence effect (250–300 ms; b) and the P3 (300–600 ms; d) in Experiment 2. The upper row shows the topographies for the difference between positive and negative feedback for both feedback reliability conditions. The lower row shows

the topographies for the difference between reliable and unreliable feedback for both feedback valence conditions. (e) Mean amplitudes in the time interval 250–300 ms at FCz as a function of feedback valence and feedback reliability. f Maximum amplitudes in the time interval 300–800 ms at Pz as a function of feedback valence and feedback reliability. Bars represent standard errors of the means. FB = feedback

and unreliable feedback (positive feedback: $M = -2.08 \mu\text{V}$, $SD = 4.81$; negative feedback: $M = -3.36 \mu\text{V}$, $SD = 4.41$), $t(33) = 2.37$, $p < .05$.

We again conducted an additional peak-to-peak analysis. In line with the previous analysis of mean amplitudes, we obtained a main effect of feedback valence, $F(1, 33) = 14.0$, $p < .001$, but neither a main effect of feedback reliability nor an interaction between the variables, $F_s < 1$. Negative feedback led to a larger negativity (reliable: $M = -9.03 \mu\text{V}$, $SD = 3.65$; unreliable: $M = -9.20 \mu\text{V}$, $SD = 3.39$) than did positive feedback (reliable: $M = -8.03 \mu\text{V}$, $SD = 3.33$; unreliable: $M = -8.06 \mu\text{V}$, $SD = 2.75$). Moreover, the analysis of the P2 peak revealed a main effect of feedback valence, $F(1, 33) = 5.21$, $p < .05$, but no effect of feedback reliability, $F(1, 33) = 1.81$, $p = .19$, or any interaction, $F < 1$. Positive feedback led to a more pronounced P2 amplitude ($M = 3.72 \mu\text{V}$, $SD = 4.97$) than did negative feedback ($M = 2.85 \mu\text{V}$, $SD = 5.14$).

P3 amplitude As in Experiment 1, the analysis of P3 amplitudes again revealed significant main effects of feedback valence, $F(1, 33) = 4.83$, $p < .05$, and feedback reliability, $F(1, 33) = 7.78$, $p < .01$, but no significant interaction, $F < 1$ (see Fig. 5b, d, and f). Again, positive feedback led to a larger P3 amplitude ($M = 8.63 \mu\text{V}$, $SD = 4.14$) than did negative feedback ($M = 7.52 \mu\text{V}$, $SD = 3.33$), and unreliable feedback attenuated the P3 amplitude ($M = 7.66 \mu\text{V}$, $SD = 3.62$) as compared to reliable feedback ($M = 8.50 \mu\text{V}$, $SD = 3.52$).

Combined results of both experiments

Learning performance In a final stage, we analyzed the two experiments in a single analysis, to determine whether the observed differences were statistically robust. To this end, we conducted mixed-model analyses on all dependent variables, with the additional between-subjects variable experiment (Exp. 1, Exp. 2). For the proportions of correct responses during the test phase, we again obtained main effects of feedback reliability, $F(1, 66) = 61.4$, $p < .001$, and feedback valence, $F(1, 33) = 91.4$, $p < .001$, but additionally observed a significant three-way interaction, $F(1, 33) = 4.38$, $p < .05$. This confirmed that the interaction between feedback valence and feedback reliability observed in Experiment 2 had been absent in Experiment 1. Separate analyses for positive and negative feedback revealed a significant interaction between feedback reliability and experiment for positive feedback, $F(1, 66) = 5.27$, $p < .05$, but not for negative feedback, $F(1, 66) = 0.67$, $p = .42$. Further testing revealed that learning performance after unreliable positive feedback was lower in Experiment 1 than in Experiment 2, $t(33) = 2.32$, $p < .05$, but there was no such difference for reliable positive feedback, $t < 1$. This indicates that unreliable feedback impaired learning from positive feedback more strongly in Experiment 1 than in

Experiment 2, whereas feedback reliability affected learning from negative feedback similarly in both experiments.

ERP results For the fronto-central valence effect, we obtained main effects of feedback valence, $F(1, 66) = 8.45$, $p < .01$, and feedback reliability, $F(1, 66) = 13.0$, $p < .01$, as well as a significant interaction between the two variables, $F(1, 66) = 5.58$, $p < .05$. Importantly, these effects were qualified by a significant three-way interaction, $F(1, 66) = 4.32$, $p < .05$, indicating that the effect of feedback reliability on fronto-central valence effect indeed differed significantly between experiments.

Because the effect of feedback reliability on the fronto-central valence effect in Experiment 1 was less robust for the peak-to-peak measures, possibly due to the effects of reliability on the P2, it is not surprising that, for the peak-to-peak analysis, the three-way interaction failed to reach significance, $F(1, 66) = 1.49$, $p = .23$. We only observed reliable effects of feedback valence, $F(1, 66) = 13.1$, $p < .001$, and feedback reliability, $F(1, 66) = 6.58$, $p < .05$. However, the latter effect was qualified by a significant interaction with experiment, $F(1, 66) = 4.66$, $p < .05$, indicating that only in Experiment 1 was fronto-central activity affected by feedback reliability.

For the P2, we observed a significant main effect of feedback reliability, $F(1, 66) = 5.58$, $p < .05$, and a marginally significant effect of feedback valence, $F(1, 66) = 3.52$, $p = .07$. Despite the interaction between feedback valence and feedback reliability in Experiment 1, there was neither a significant three-way interaction, $F(1, 66) = 2.53$, $p = .12$, nor a significant interaction between experiment and feedback valence, $F(1, 66) = 2.42$, $p = .13$. This analysis does not support the initial impression that feedback reliability and feedback valence had differential effects on the P2 in each experiment.

Finally, for the P3, we found significant main effects of feedback valence, $F(1, 66) = 9.65$, $p < .01$, and feedback reliability, $F(1, 66) = 11.5$, $p < .001$, but no main effect of the experiment variable, $F(1, 66) = 2.25$, $p = .14$, and no significant interactions involving this variable, $F_s < 1.2$. This confirms our observation that the effects of feedback reliability on the P3 did not differ between experiments.

Discussion

The results of Experiment 2 differed in one crucial respect from those in Experiment 1. In line with the prediction of the feedback valuation account, the effects of feedback reliability on the fronto-central valence effect were now absent. This suggests that a contrast effect contributed to the pattern of results in Experiment 1, which provides evidence that the adaptive valuation of feedback contributed to the pattern of results observed in Experiment 1. This increased effect of feedback reliability on the fronto-central valence effect in Experiment 1 was mirrored by learning effects, at least for

learning from positive feedback. Here, unreliable feedback impaired learning more strongly in Experiment 1 than in Experiment 2. In contrast, the feedback P3 was similarly reduced for unreliable as compared to reliable feedback in both experiments.

General discussion

In the present study, we used a simple learn–test paradigm to investigate the mechanisms underlying the effects of feedback reliability on the ERP correlates of feedback processing. Our central hypothesis was that the reduction of the fronto-central valence effect (representing an FRN or a reward positivity) following unreliable feedback (Ernst & Steinhauser, 2017; Schiffer et al., 2017; Walentowska et al., 2016) reflects the adaptive devaluation of unreliable feedback, which influences the generation of prediction errors during reinforcement learning. From this account, we derived the prediction that the reduction of the fronto-central valence effect should be stronger if participants experience both unreliable and reliable feedback within the same context than if only one level of reliability is experienced. Such a contrast effect would imply that the valuation of feedback varies with the presence of alternative outcomes within a given context. This prediction was fully confirmed. Whereas a strong effect of feedback reliability on the fronto-central valence effect was obtained when reliable and unreliable feedback varied across trials (Exp. 1), it was fully absent when only one level of reliability was experienced within a block of trials (Exp. 2). Because feedback reliability mainly influenced positive feedback, the effects in Experiment 1 presumably reflect a modulation of the reward positivity (Holroyd et al., 2008). We interpret this finding as reflecting the consequences of top-down control on reinforcement learning through modulation of the value of a specific type of feedback and not as a direct correlate of the activity of the top-down control process itself. Because top-down control devaluated unreliable feedback, subsequent reward prediction errors were attenuated, as reflected by fronto-central ERP activity (Holroyd & Coles, 2002).

Our results cannot be explained by an alternative account that assumes that the processing of unreliable feedback in the dopaminergic reward system is gated or suppressed via top-down control (Ernst & Steinhauser, 2017; Walentowska et al., 2016). Such an account would have predicted that a stronger effect of feedback reliability on the fronto-central valence effect should have been observed in Experiment 2. In this experiment, feedback reliability was fully predictable before feedback on an individual trial was even provided. This should have facilitated the recruitment of top-down control. In Experiment 1, in which the gating or suppression of a prediction error had to wait until the feedback color was evaluated, the effect of feedback reliability should have been strongly

decreased. Our results clearly contradicted this prediction. Please note that the devaluation account does not assume that devaluation has to occur on each trial. Rather, the assignment of subjective values to the feedback stimuli could have already occurred, on the basis of instructions at the beginning of the experiment.

Although top-down gating or suppression cannot explain our present findings, this does not imply that top-down gating or suppression cannot affect reinforcement learning at all. It may well be that under different experimental conditions, these mechanisms additionally contribute to the effects of feedback reliability (e.g., Doll et al., 2011). For instance, when feedback stimuli with strong intrinsic value (e.g., a happy face; see Walentowska et al., 2016) are used, devaluation might be ineffective, which necessitates additional mechanisms in order to avoid learning from invalid feedback. A further mechanism could be the modulation of attention to feedback according to either feedback reliability or feedback value. Indeed, feedback reliability in Experiment 1 also modulated the P2 amplitude, possibly reflecting that feedback's valence and reliability also affected early attentional orienting (Luck & Hillyard, 1994). Thus, whereas the devaluation of unreliable feedback could affect learning directly via an altered prediction error, it could additionally imply that less attention is allocated to unreliable and devalued feedback stimuli, a mechanism that should be investigated in future studies.

The question emerges why we did not find an effect of feedback reliability on the fronto-central valence effect at all when feedback reliability was held constant, whereas previous studies had found an effect under blocked reliability conditions (Ernst & Steinhauser, 2017; Schiffer et al., 2017; Walentowska et al., 2016). This could be due to differences in the experimental paradigms. In the present paradigm, participants never knew whether the feedback on an individual trial was valid or invalid, because we used a learn–test paradigm in which each stimulus was presented only once for learning. In Schiffer et al. (2017), participants repeatedly made decisions for the same stimuli. After some learning, participants could know whether a given feedback was presumably invalid. Likewise, in our previous study (Ernst & Steinhauser, 2017) we used a learn–test paradigm, but one in which the participants received information about the validity of the preceding feedback on each trial. In both studies, participants might have devalued the feedback immediately after trials with invalid feedback relative to the aggregated feedback. This could explain why some devaluation was also obtained in these tasks. However, we predict that, for these paradigms, mixing reliable and unreliable feedback should also lead to a stronger effect than holding feedback reliability constant across blocks of trials.

Although the effect of feedback reliability on the fronto-central valence effect was absent in Experiment 2, we still

observed effects of feedback reliability on both the feedback P3 and learning. The differential effects of feedback reliability on the fronto-central valence effect and the P3 suggest that feedback value influences the underlying systems in different ways. Although the generation of prediction error in the reinforcement-learning system seems to be more affected by the *relative* value of feedback, attentional feedback processing, as indicated by the feedback P3, appears to be sensitive to its *absolute* value. This might reflect that the latter system is less sensitive to the reward assigned to a (feedback) stimulus and is more sensitive to the stimulus's informational value. Unreliable feedback is less informative, and thus less motivationally significant, because it conveys less evidence about the correctness of a response (Schiffer et al., 2017). The importance of the information conveyed by a stimulus for the generation of the P3 has already been noted in early P3 theories (e.g., Johnson, 1986; Johnson & Donchin, 1978) and is a central aspect in current theories (e.g., O'Connell, Dockree, & Kelly, 2012). Specifically, "motivational significance" (Nieuwenhuis, 2011; Nieuwenhuis et al., 2005) or "task relevance" (Ferdinand et al., 2012), as major determinants of the P3 amplitude, can be interpreted as the amount of information conveyed by a stimulus that can be used to inform the direction and intensity of future behavior.

Feedback reliability affected learning in both experiments, with unreliable feedback leading to generally reduced test performance. This could imply that learning in our paradigm is driven not only by reinforcement learning (as reflected by the fronto-central valence effect) but also by more explicit learning (possibly reflected by the feedback P3). The observation that the P3 predicts learning in our learn–test paradigm to a higher degree than the fronto-central valence effect is in line with the findings from a previous study (Ernst & Steinhauser, 2012). However, at least for positive feedback, the feedback reliability effect on learning was smaller in Experiment 2 than in Experiment 1, which mirrors the smaller feedback reliability effect on the fronto-central valence effect in Experiment 2. This suggests that reinforcement learning contributes more strongly to learning from positive than from negative feedback in our paradigm.

Our results contribute to the literature on stimulus valuation in the dopaminergic reward system and demonstrates that the well-known dynamic representation of relative stimulus values can also be observed for feedback stimuli (Bellebaum et al., 2010; Holroyd et al., 2004). It has previously been established that activity in the orbitofrontal cortex (OFC) represents not only the absolute value of stimuli (e.g., Padoa-Schioppa & Assad, 2008), but also their relative value, with reward-related OFC activity being adjusted relative to the overall range of available rewards (e.g., Grabenhorst & Rolls, 2011; Kobayashi, de Carvalho, & Schultz, 2010). Furthermore, it has previously been shown that top-down processes can modulate OFC activity in accordance with goals

and instructions (e.g., De Araujo et al., 2005; Grabenhorst & Rolls, 2010; Grabenhorst et al., 2007; Hare et al., 2009; Plassmann et al., 2008) and that this adjustment can occur rapidly (e.g., Hornak et al., 2004; Kringelbach & Rolls, 2003). Although these studies have mostly investigated the valuation of choice options in decision making, our present study ties these findings to feedback processing. Our study demonstrates that prior knowledge about the range of possible feedback reliabilities leads to an adaptive adjustment of the value of feedback stimuli. This value information may be communicated to the striatum (e.g., Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008), where it affects the generation of reward prediction errors upon feedback presentation. Encoded in a phasic dopamine signal, reward prediction errors then modulate activity in the anterior cingulate cortex, as reflected by fronto-central components in feedback-related brain activity (Holroyd & Coles, 2002).

In sum, our study suggests that feedback-related brain activity is affected by feedback reliability via the adaptive valuation of feedback stimuli. Accordingly, these effects follow the principles of value representation in the brain—for example, by being susceptible to contrast effects resulting from the representation of relative stimulus values. Future research could aim to further establish the validity of this account and investigate whether this is the only mechanism underlying the effects of feedback reliability, or whether further mechanisms contribute to these effects.

Author note This research was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG: STE 1708/3-1) to M.S. We are grateful to Johannes Fiedler, Christina Görner, and Sabine Utschick for assistance in conducting the experiments.

References

- Bellebaum, C., Polezzi, D., & Daum, I. (2010). It is less than you expected: The feedback-related negativity reflects violations of reward magnitude expectations. *Neuropsychologia*, 48, 3343–3350.
- Crespi, L. P. (1942). Quantitative variation of incentive and performance in the white rat. *American Journal of Psychology*, 55, 467–517.
- De Araujo, I. E., Rolls, E. T., Velasco, M. I., Margot, C., & Cayeux, I. (2005). Cognitive modulation of olfactory processing. *Neuron*, 46, 671–679.
- 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. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *Journal of Neuroscience*, 31, 6188–6198.
- Doll, B. B., Jacobs, W. J., Sanfey, A. G., & Frank, M. J. (2009). Instructional control of reinforcement learning: A behavioral and neurocomputational investigation. *Brain Research*, 1299, 74–94.
- Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, 11, 357–374, disc. 374–427. <https://doi.org/10.1017/S0140525X00058027>

- Ernst, B., & Steinhauser, M. (2012). Feedback-related brain activity predicts learning from feedback in multiple-choice testing. *Cognitive, Affective, & Behavioral Neuroscience*, 12, 323–336.
- Ernst, B., & Steinhauser, M. (2015). Effects of invalid feedback on learning and feedback-related brain activity in decision-making. *Brain and Cognition*, 99, 78–86.
- Ernst, B., & Steinhauser, M. (2017). Top-down control over feedback processing: The probability of valid feedback affects feedback-related brain activity. *Brain and Cognition*, 115, 33–40.
- Ferdinand, N. K., Mecklinger, A., Kray, J., & Gehring, W. J. (2012). The processing of unexpected positive response outcomes in the medial frontal cortex. *Journal of Neuroscience*, 32, 12087–12092.
- Flaherty, C. F. (1982). Incentive contrast: A review of behavioral changes following shifts in reward. *Animal Learning & Behavior*, 10, 409–440.
- Flaherty, C. F. (1999). Incentive relativity (Vol. 15). Cambridge: Cambridge University Press.
- Foti, D., Weinberg, A., Bernat, E. M., & Proudfit, G. H. (2015). Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. *Clinical Neurophysiology*, 126, 1338–1347.
- Grabenhorst, F., & Rolls, E. T. (2010). Attentional modulation of affective versus sensory processing: Functional connectivity and a top-down biased activation theory of selective attention. *Journal of Neurophysiology*, 104, 1649–1660.
- Grabenhorst, F., & Rolls, E. T. (2011). Value, pleasure and choice in the ventral prefrontal cortex. *Trends in Cognitive Sciences*, 15, 56–67.
- Grabenhorst, F., Rolls, E. T., & Bilderbeck, A. (2007). How cognition modulates affective responses to taste and flavor: Top-down influences on the orbitofrontal and pregenual cingulate cortices. *Cerebral Cortex*, 18, 1549–1559.
- Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*, 324, 646–648.
- Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *Journal of Neuroscience*, 28, 5623–5630.
- 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, 679–708. <https://doi.org/10.1037/0033-295X.109.4.679>
- Holroyd, C. B., Larsen, J. T., & Cohen, J. D. (2004). Context dependence of the event related brain potential associated with reward and punishment. *Psychophysiology*, 41, 245–253. <https://doi.org/10.1111/j.1469-8986.2004.00152.x>
- Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, 45, 688–697. <https://doi.org/10.1111/j.1469-8986.2008.00668.x>
- Hornak, J., O'Doherty, J., Bramham, J., Rolls, E. T., Morris, R. G., Bullock, P. R., & Polkey, C. E. (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *Journal of Cognitive Neuroscience*, 16, 463–478.
- Johnson, J., Jr (1986). A triarchic model of P300 amplitude. *Psychophysiology*, 23, 367–384.
- Johnson, J., Jr, & Donchin, E. (1978). On how P300 amplitude varies with the utility of the eliciting stimuli. *Electroencephalography and Clinical Neurophysiology*, 44, 424–437.
- Kobayashi, S., de Carvalho, O. P., & Schultz, W. (2010). Adaptation of reward sensitivity in orbitofrontal neurons. *Journal of Neuroscience*, 30, 534–544.
- Kringelbach, M. L., & Rolls, E. T. (2003). Neural correlates of rapid reversal learning in a simple model of human social interaction. *NeuroImage*, 20, 1371–1383.
- Li, J., Delgado, M. R., & Phelps, E. A. (2011). How instructed knowledge modulates the neural systems of reward learning. *Proceedings of the National Academy of Sciences*, 108, 55–60.
- Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31, 291–308.
- Maxwell, F. R., Calef, R. S., Murray, D. W., Shepard, J. C., & Norville, R. A. (1976). Positive and negative successive contrast effects following multiple shifts in reward magnitude under high drive and immediate reinforcement. *Animal Learning & Behavior*, 4, 480–484.
- 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. <https://doi.org/10.1162/jocn.1997.9.6.788>
- Nieuwenhuis, S. (2011). Learning, the P3, and the locus coeruleus–norepinephrine system. In R. B. Mars, J. Sallet, M. Rushworth, & N. Yeung (Eds.), *Neural basis of motivational and cognitive control* (pp. 209–222). Oxford: Oxford University Press.
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus–norepinephrine system. *Psychological Bulletin*, 131, 510–532. <https://doi.org/10.1037/0033-2909.131.4.510>
- O'Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature Neuroscience*, 15, 1729–1735.
- Padoa-Schioppa, C., & Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience*, 11, 95–102.
- Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences*, 105, 1050–1054.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118, 2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>
- Pontifex, M. B., Hillman, C. H., & Polich, J. (2009). Age, physical fitness, and attention: P3a and P3b. *Psychophysiology*, 46, 379–387.
- Rushworth, M. F., & Behrens, T. E. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, 11, 389–397.
- 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.
- San Martín, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, 6:304. <https://doi.org/10.3389/fnhum.2012.00304>
- Schiffer, 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.
- Schultz, W., Dayan, P., & Montague, P. R. (1997) A neural substrate of prediction and reward. *Science*, 275, 1593–1599.
- Walentowska, W., Moors, A., Paul, K., & Pourtois, G. (2016). Goal relevance influences performance monitoring at the level of the FRN and P3 components. *Psychophysiology*, 53, 1020–1033.
- 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. <https://doi.org/10.1016/j.neubiorev.2012.05.008>
- Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, 24, 6258–6264. <https://doi.org/10.1523/JNEUROSCI.4537-03.2004>