Evidence for parietal reward prediction errors using great grand average meta-analysis

Harry J. Stewardson, Thomas D. Sambrook

PII: S0167-8760(20)30064-7

DOI: https://doi.org/10.1016/j.ijpsycho.2020.03.002

Reference: INTPSY 11724

To appear in: *International Journal of Psychophysiology*

Received date: 20 December 2019

Revised date: 26 February 2020

Accepted date: 22 March 2020

Please cite this article as: H.J. Stewardson and T.D. Sambrook, Evidence for parietal reward prediction errors using great grand average meta-analysis, *International Journal of Psychophysiology* (2020), https://doi.org/10.1016/j.ijpsycho.2020.03.002

Evidence for parietal reward prediction errors using great grand average meta-analysis

Harry J. Stewardson and Thomas D. Sambrook, University of East Anglia

Evidence for parietal reward prediction errors using great grand average meta-analysis

Abstract

As a basic principle within the economics of decision-making, reinforcement learning dictates that individuals strive to repeat behaviour that elicits reward, and avoid behaviour that elicits punishment. Neuroeconomics aims to measure reinforcement learning physically in the brain through the use of reward prediction errors: the difference between expected outcome value and actual outcome value following decision-making behaviour. Two electrophysiological components, the frontocentral feedback-related negativity and the more parietal P3, are implicated in outcome processing, but whether these components encode a reward prediction error has been unclear. A source of the unclear literature is likely to be inconsistent quantification of the components. A recent meta-analysis that directly quantified published waveforms rather than using reported effect sizes found strong evidence that the feedback-related negativity encodes a reward prediction error. In the current study, such a meta-analysis was performed on parietal waveforms to establish whether the P3, or parietal areas generally, are sensitive to reward prediction errors. A strong effect was found, both of reward prediction error encoding and simple valence sensitivity at a latency associated with the P3.

Keywords

Great grand average; reward prediction error, feedback related negativity; FRN; P300; P3

Introduction

According to reinforcement learning, an agent seeks to repeat behaviours that are rewarding, and avoid behaviours that are punishing (Thorndike, 1898). With an effective learning mechanism, rewards and punishments are not valued dichotomously, but instead as reward prediction errors (Sutton and Barto, 1998). A reward prediction error (RPE) is the signed difference between the value of an obtained outcome and its prior expected value, and it is this difference that guides future action valuation and thus future behaviour.

Computationally, the value of an RPE is determined by two terms; the valence, or sign, of the prediction error (positive vs. negative, based on whether an outcome is better or worse than expected) and its size (how great the discrepancy in value is between expected and obtained outcomes). Positive RPEs can be the result of either elicited rewards (such as receiving more money for a gamble than expected) or omitted punishments (such as not receiving an electric shock when one was expected to occur). Conversely, negative RPEs are the result of either omitted rewards or incurred punishments. It is important to distinguish between RPE valence and outcome valence; a reward (winning a sum of money, say) can still generate a negative RPE if it was less rewarding than was expected. For reinforcement learning models, it is the valence of the RPE that is relevant and in the present study we use valence to refer to the sign of the RPE not the outcome.

Distinguishing RPE valence is crucial for reinforcement learning, however optimal learning should also be influenced by the size of an RPE: big or unexpected outcomes should drive greater learning than small or expected outcomes. RPE size refers to the amount of "error" in an RPE, with the level of learning adjustment proportional to the difference between an outcome and its expected value. Since the direction of this learning (i.e. whether an action is more or less likely to be repeated) is determined by valence, with large RPEs

having opposite effects depending on whether they are positive or negative, the signature response of a neural RPE encoder consists in a sign x size interaction (Caplin & Dean, 2008).

Feedback related negativity (FRN) is a frontocentral component occurring approximately 240 ms – 340 ms after feedback is received in a task involving either rewards or punishments. It has been proposed as an RPE encoder (Holroyd & Coles, 2002). It is typically quantified by a difference wave created by subtracting the voltage of positive RPEs from negative RPEs. This results in the FRN's characteristic negative peaked difference wave when, in the simple waveforms underlying it, positive RPEs produce a relatively positive-going voltage compared to negative RPEs. The term 'reward positivity' may refer to this relative positive-going voltage of the simple waveforms but also to the claim that the component is sensitive specifically to rewards rather than non-rewards (Proudfit, 2015).

A number of authors have asserted that this component reflects not RPE sign, but size, i.e. how unexpected or surprising an outcome is (Garofalo, Maier, & di Pellegrino, 2014; Hauser et al., 2014; Talmi, Fuentemilla, Litvak, Duzel, & Dolan, 2012) This claim predicts a flat difference wave for the negative – positive contrast (provided each outcome is equally likely) and instead a difference wave for a small – large contrast. A meta-analysis by Sambrook and Goslin (2015) showed both these effects to be present in the interval associated with the FRN. Importantly however, there existed also the signature sign x size interaction indicating the encoding of an RPE proper, i.e. a signed quantitative term complying with reinforcement learning theory (Sutton and Barto, 1998).

The P3, or P300 is a parietally distributed, positive going ERP, occurring 250 ms – 500 ms after presentation of task-relevant information. The P3 has been widely studied in decision-making research, and is associated with engagement of attention, the processing of novelty and expectation, and cognitive workload (Polich, 2003). As with the FRN, its status with regard to RPE encoding has been unclear. An influential "independent coding model"

(Yeung & Sanfey, 2004) has claimed that the P3 is sensitive only to RPE size, with the FRN sensitive only to RPE valence. It is certainly the case that the P3 is usually represented by larger amplitudes following larger outcomes regardless of their valence, (Bellebaum, Kobza, Thiele, & Daum, 2010; Gu et al., 2011; Kreussel et al., 2012; Toyomaki & Murohashi, 2005), and unexpected outcomes likewise (Bellebaum & Daum, 2008; Hajcak, Holroyd, Moser, & Simons, 2005; Hajcak, Moser, Holroyd, & Simons, 2007; Wu & Zhou, 2009) and, as such, may indicate RPE *size* encoding. This does not entail coding for valence however. In fact, in a review of the effect of valence on the P3, San Martin (2012) found that sensitivity to valence was also typically found, though the direction of effect was not consistent. Similarly, a review (Glazer, Kelley, Pornpattananangkul, Mittal, & Nusslock, 2018) found equivocal evidence for the P3's status as a valence encoder. In contrast to these main effects, there has been a dearth of research on the P3's status as a full RPE encoder, capturing the RPE sign x size interaction.

The present study addresses this. It uses the great grand average methodology of ERP meta-analysis, in which published effect sizes are disregarded in favour of direct quantification of published waveforms. The method has been extensively validated (Sambrook and Goslin, 2015) and in the case of the FRN was shown to provide superior estimates of effect size than conventional meta-analysis. We have previously argued the benefits of this approach as a solution to inconsistencies in the latency at which ERP components are identified. The present study demonstrates a further advantage of the method, which is its capacity to use data for purposes other than those to which it was originally put. The meta-analysis provided here is based on thirty-three studies, of which only a small number specifically addressed the P3.

In this study, the principal effect under investigation is an RPE encoding at parietal sites. However, we also present effect sizes for a simple dichotomous coding of valence and

for prediction error size (sometimes referred to as salience, surprise or unsigned prediction error). We additionally perform a moderator analysis on the RPE effect to establish if it is sensitive to the degree of control over the outcome that the task appears to afford.

Method

Inclusion and Exclusion Criteria

For inclusion, a study needed to manipulate the independent variables of prediction error valence (either positive or negative) and prediction error size (either small or large). Experiments were included provided that both of these variables were manipulated in a 2 x 2 factorial design. Prediction error size could be manipulated by adjusting either the magnitude of outcomes or their likelihood. Experiments were included if they had intermediary levels of prediction error size, although these were ignored in favour of the most extreme levels to maximise contrasts. Experiments had to provide, as a dependent variable, simple waveforms corresponding to the design above, or difference waves appropriate for the analysis as described below. Importantly, provided that such voltage x time waveforms were presented, no statistics associated with these needed to be reported. Waveforms had to be presented parietally (Pz, CPz or POz) but also frontocentrally (FCz, Fz, Cz) in order to distinguish parietal effects from those of the FRN. The feedback-locked epoch needed to run, at minimum, from -100 to 500 ms.

Regarding participants, experiments either needed a population of healthy adults or a healthy control group if an alternative experimental group was used. Any studies including participants who had been selected based on a screening process were excluded.

Experimental tasks were restricted to those offering monetary stakes (either wins or losses) but could otherwise vary, including passive tasks in which participants knew they could not influence outcomes, guessing tasks in which participants were encouraged to increase the likelihood of desirable outcomes despite no control being really available, and rule implementation tasks in which outcomes were probabilistically linked to a choice rule

that participants had to learn. These three kinds of task also constituted the three levels of the moderator variable "control over outcome" described below. Criteria for inclusion and exclusion were identical to those used by Sambrook and Goslin (2015) with the exception of the additional requirement of a parietal waveform.

Moderator analyses

A sole moderator, control over outcome was operationalised at three levels: "passive", "guessing" and "rule implementation", following Sambrook and Goslin (2015).

Search Strategies

English language journals and books were searched using following databases: PsycInfo, PsycArticles, ERIC, PubMed and Web of Science. Results were compiled in Zotero. The search covered journal titles, abstracts and keywords. Articles were gathered based on three searches. First, articles used in the Sambrook and Goslin (2015) meta-analysis of the FRN were used if they featured parietal waveforms. This yielded twenty studies. Second, the principal search string used in that meta-analysis was used once again but in the time frame 2015 to present, in order to capture articles published subsequently. This string was "feedback negativity" OR "feedback related negativity" OR "feedback error-related negativity" OR "reward positivity" OR "feedback correct related positivity". This search was conducted on the basis that much RPE research is focussed on the FRN and likely to be associated with this term, but that parietal waveforms might nevertheless be provided. This search provided twelve more studies. Third, a mutually exclusive search was performed on articles targeting the P3 but with limiting terms owing to the extensive P3 literature. The

search string was ("P3" OR "P300") AND ("learning" OR "motivational significance" OR "prediction error") AND NOT ("feedback negativity" OR "feedback related negativity" OR "feedback error-related negativity" OR "reward positivity" OR "feedback correct related positivity"). This search yielded one more study. In total, these searches gave a final study sample of thirty-three studies. The set of studies overlapped heavily with the studies used to meta-analyse the frontal, FRN component in Sambrook and Goslin 2015. Given so few additional non-FRN studies of reward prediction error were found by our literature search, we thought it highly unlikely that there would be a significant number of obtainable unpublished studies that had not been located in the generation of that paper.

Coding Procedures

Electronic copies of experiments were accessed and screenshots were taken of grand average waveforms. Where multiple parietal waveforms were available, they were taken in preference order Pz, CPz, POz; where multiple frontocentral waveforms were available they were taken in preference order FCz, Fz, Cz. Screenshots were cropped and enlarged, and the waveforms digitised with Plot Digitizer (http://sourceforge.net/projects/plotdigitizer) by using a mouse to manually lay points along the waveform at approximately 5 ms intervals. A custom Excel macro (available as a supplementary file in Sambrook and Goslin, 2015) linearly interpolated coordinates at 2 ms intervals between the existing manually assigned ones. The digitizing process was performed twice for each waveform and a mean taken in order to provide greater accuracy. Waveforms were then replotted to make sure that they corresponded visually with the original. Further details can be found in Sambrook and Goslin (2015). In the case of one paper (Banis & Lorist, 2012) original data was used, previously obtained from the authors,

since no parietal waveform was presented in the paper. Original data was also used from papers authored by Sambrook.

Statistical Methods

A major strength of the great grand average method of meta-analysis is that it establishes average effect size at all points on the waveform, not merely those reported in individual papers, thus maximising the data used, and eliminating idiosyncrasies in quantification in individual papers. It means however, that study-level variance is unknown, since this is rarely shown in the grand average waveforms we use here as data. Instead of standardised effect sizes, which are the norm in conventional meta-analysis, simple or "raw" effect sizes are used (see Sambrook and Goslin, 2015 for a discussion of the merits of each). Three simple effect sizes were computed, realised in each case as a waveform of effect size, based on differencing of the simple waveforms described in the design earlier. The *valence effect* was computed by a difference wave (large negative RPE - large positive RPE). Large prediction errors only were used in order to maximise the contrast. The *prediction error size effect* was computed by the difference wave (small negative RPE + small positive RPE) - (large negative RPE + large positive RPE). Following Sambrook and Goslin (2015), the *RPE effect* was computed by the difference wave (small negative RPE - small positive RPE) - (large negative RPE - large positive RPE), a difference of difference waves which implements an axiomatic test for RPE encoding (Caplin and Dean, 2008). In the case of the FRN (operationalised as a difference wave), it predicts that the FRN for small outcomes will be of lower amplitude than the FRN for large outcomes, generating a difference of difference waves of non-zero amplitude. An equivalent rationale was used here to establish parietal RPE encoding.

Study weighting in conventional meta-analysis is based either on a study's variance or its sample size. Since variance was unavailable, study size was used instead, implemented in the form of weighted *t* tests. The significance of effect sizes was established by submitting the appropriate difference waves to running univariate weighted one-sample *t* tests (test value = 0) at each 2 ms time point. All figures in the Results section, both of *t* tests and voltage plots show weighted effects. Publication bias was assessed by subjecting funnel plots to the trim and fill procedure of Duval and Tweedie (2000) implemented in R (R Core team) using the metafor package (Viechtbauer, 2010). Owing to a dearth of studies at the "passive" level of the moderator, moderator analysis was performed on just the "guessing" and "rule implementation" levels. This was implemented as a 2 x 2 (moderator x site) ANOVA on RPE effect size at its peak (286 ms frontally and 348 ms parietally).

Results

Effect of valence

Figure 1a shows the valence effect at parietal and frontal sites. A typical frontocentral FRN is found, peaking at 286 ms, and a later parietal effect peaking at 312 ms with an effect size of -2.33 μv. One-sample *t* tests were conducted on these waveforms and are shown in Figure 1b. Between 400 ms and 442 ms a significant valence effect was found only at parietal sites. Results were similar when unweighted *t* tests were performed and also when effects were examined separately for studies using likelihood and magnitude as the prediction error size modulators (Supplementary Figure 1). Trim and fill applied to the parietal effect size at peak revealed evidence of publication bias. After this was removed with the addition of ten imputed studies, the effect size fell to -1.43 μv but remained highly significant (z = -3.42, p =

0.0007). The funnel plot with imputed studies is shown in Supplementary Figure 2. The main effect of prediction error size (i.e. ignoring valence) is shown in Supplementary Figure 3 and simple waveforms for these effects are provided in Supplementary Figures 4 and 5.

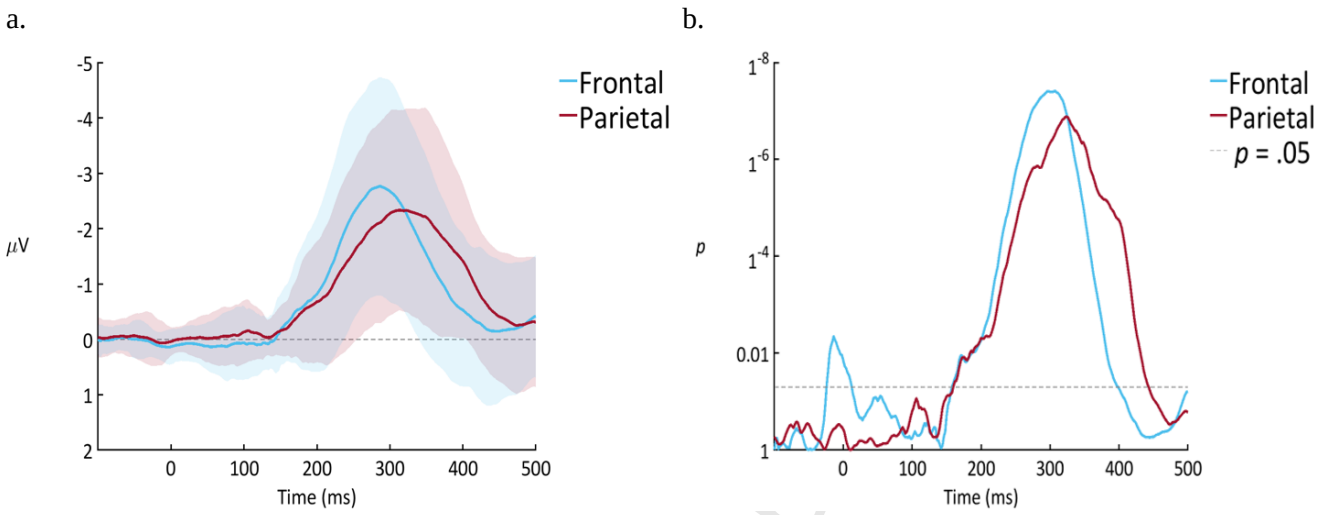


Figure 1. Effect of valence (negative – positive difference wave) at frontal and parietal sites with standard deviations shown in shadow. a. Great grand average voltage. b. Significance of each difference wave under a one-sample *t* test over experiments.

Effect of reward prediction error

Figure 2a shows the RPE effect at parietal and frontal sites. In keeping with Sambrook and Goslin (2015) a frontal RPE effect is found at a latency associated with the FRN, peaking at 286 ms and a later parietal effect peaking at 348 ms with an effect size of -.89 μv. Figure 2b shows the significance of the effect at each site. In the interval 356 ms to 418 ms a significant effect was found only at parietal sites. Again, results were similar when using unweighted *t* tests or analysing likelihood or magnitude modulated studies only (see Supplementary Figure 6). Trim and fill applied to the parietal effect size at peak revealed no evidence of publication bias (funnel plot shown in Supplementary Figure 7). Moderator analysis showed a main effect of moderator: as in Sambrook and Goslin (2015) the RPE effect was stronger in rule implementation than in guessing ($F_{1,28} = 4.24$ $p = .031$, $\sigma^2 = .16$), but no interaction between

this and site ($F_{1,28}$ = .59 p = .45). Supplementary Figure 8 shows moderator effects across the full waveform.

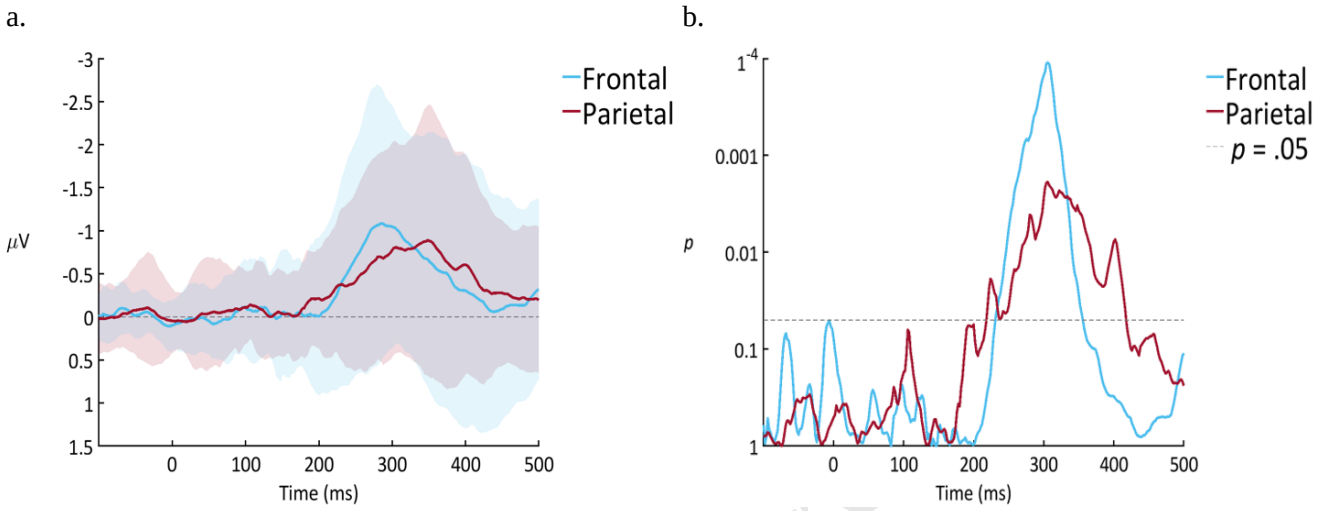


Figure 2. Effect of RPE encoding (small negative – small positive difference wave) – (large negative – large positive difference wave) at frontal and parietal sites with standard deviations shown in shadow. a. Great grand average voltage. b. Significance of each difference wave under a one-sample *t* test over experiments.

Independent components vs. information relay

The parietal effect demonstrated does not necessarily imply an independent process since the underlying generator may simply be receiving information relayed from the generator responsible for the FRN. If this were the case we would expect a strong correlation, across experiments, of the two effects in the intervals in which they were significant. In contrast, a weak correlation between these effects, despite their being strong responses in their own right, would suggest independent processes. Such a correlation, performed across subjects, is commonly used within individual experiments for the purpose of demonstrating common vs. separate processes. It is equally valid when run across experiments however. Fully independent RPE encoders are likely to capture different aspects of this property, which can take forms beyond its basic formulation (as we cover in the Discussion), and different tasks

would thus be expected to variably elicit one or other component if those components are independent.

As an indicator of non-independent processes, correlated activity needs to be taken in the context of the strong temporal and spatial correlations present in any ERP. If an effect at a parietal site constitutes the arrival of information previously held only at a frontal site then we would expect the correlation between parietal and frontal sites, at this time lag, to be greater than the correlation between the parietal site and itself over the same lag. The heat map in Figure 3 shows the difference between these correlations, expressed as a signed Z score. Two masks have been applied, first the parietal-frontal correlation must be greater than the parietal-parietal correlation (since this is signature of a relayed signal described above) and second the parietal-frontal correlation must be significant in its own right (r = .343, N = 33). The rectangle indicates the temporal co-ordinates at which both frontal and parietal sites show a significant RPE effect (as portrayed in Figure 2b), which is where we would expect relatively large parietal – frontal correlations if information were being relayed. These are largely absent, suggesting independent frontal and parietal components.

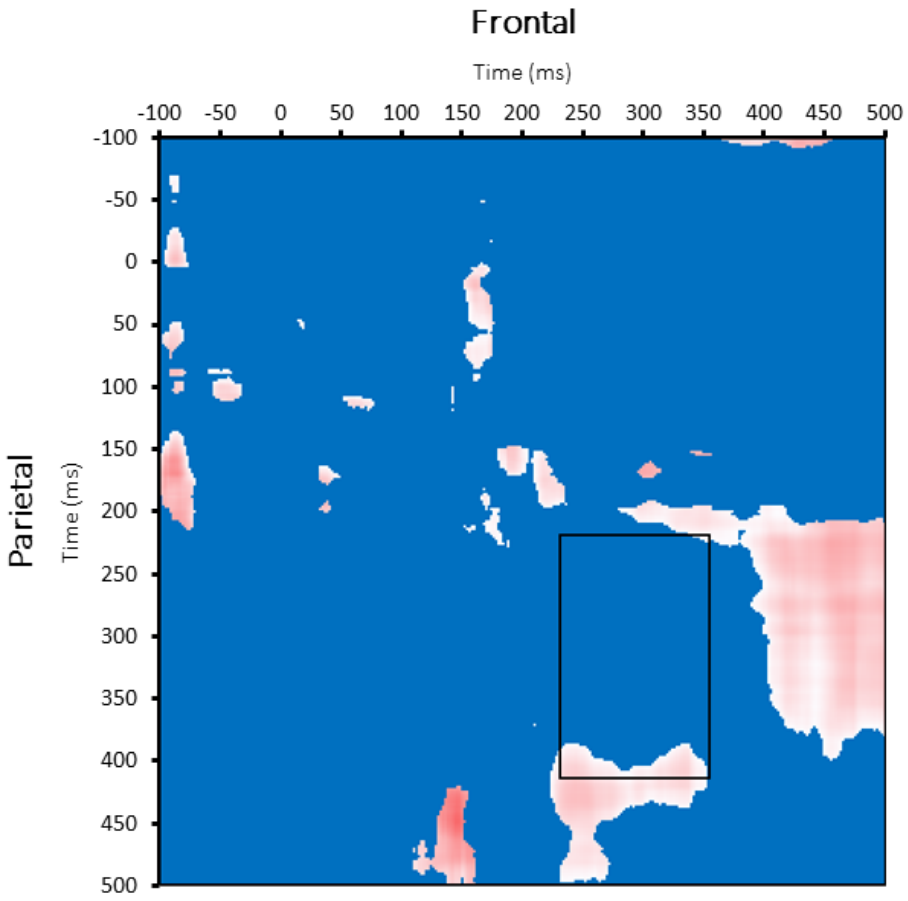


Figure 3. Heat map depicting superiority of parietal-frontal correlations over parietal-parietal correlations at all possible temporal co-ordinates. Regions where parietal-parietal correlations are superior are masked with blue (for example the diagonal from top left to bottom right where the parietal-parietal correlation must be 1). Regions where the parietal-frontal correlation is non-significant are also masked out. The rectangle encloses co-ordinates where both frontal and parietal effect are in operation (see Figure 2): if a signal is relayed between sites, correlations should be strong at these co-ordinates.

Discussion

This meta-analysis has established the reality of a parietal encoding both of valence and RPE, with better outcomes in both cases associated with a relative positivity in voltage. The existence of these parietal effects has been unresolved in the literature, with the comprehensive review of San Martin (2012), concluding a parietal valence effect was likely,

but that its direction was unclear. Unfortunately, such a conclusion is very possible when attempting to collate effects taken from inconsistent intervals of the feedback-locked waveform since it comprises a series of peaks and troughs and small shifts in the measurement window can reverse the polarity of effects. The great grand averaging method used here reveals the underlying direction of effect and its latency.

One benefit of a GGA meta-analysis is that it may indicate an appropriate interval in which to measure a component in the future, as Sambrook and Goslin (2015) provided for the FRN. This is somewhat problematic in the present case because the parietal effect of RPE encoding overlaps closely with the much stronger, and same-signed, frontal effect. Indeed, we cannot rule out the possibility that the earlier portion of the parietal effect is entirely due to volume conduction from the frontal effect. However, this cannot be the case in the interval in which the parietal response is stronger. On a pragmatic basis, we therefore recommend the parietal RPE encoding component be quantified as mean activity in the interval 350 ms to 420 ms, this corresponding to the interval in which our meta-analysis found the parietal effect to be stronger than the frontal effect and still significant (as shown in Figure 2b). Future studies that better separate the two components will be used to refine this.

We should clarify that the difference wave approach we have used does not allow us to conclude whether the valence effect arises from sensitivity (in opposing directions) to both good and bad outcomes, or from sensitivity to just one valence. Nor does it inform us whether the RPE effect arises from a sensitivity to prediction error size in both negative RPEs and positive RPEs, i.e. reflects a full bivalent coding across the range of possible prediction errors. The significant difference of difference waves could be generated by an encoder that is sensitive to the size of positive RPEs but not negative RPEs or vice versa. Difference waves indicate that the brain performs discriminations along the lines of an experiment's conditions but the observed wave is not necessarily representative of neural activity. This is

true whether difference waves are built from great grand averages as in the present case, or from the grand averages found in a single experiment. Establishing the contribution of positive RPEs or negative RPEs to the observed RPE effect depends on the interpretation of simple, single condition waveforms, however this is inherently problematic owing to the multiple components contributing to such waveforms (Luck, 2014) and problems arising from interpreting single condition waveforms in the specific context of the FRN have been illustrated by Sambrook and Goslin (2016). While this question can be addressed using methods for decomposing component overlap, these methods lie beyond the current study which thus restricts itself to difference waves.

The parietal effects, particularly in the case of the RPE effect, lie within the range of activity typically ascribed to the P3 component. An influential model (Yeung and Sanfey, 2004) claims that in reinforcement learning tasks such as those used in this study, the P3 codes only for outcome magnitude (unsigned prediction error size) while the FRN codes only for outcome valence. This claim is not supported by the present study. Nevertheless, we would be cautious in ascribing to the P3 the role of encoding all three properties of valence, RPE, and unsigned prediction error size. First, it must be noted that there is too little spatial information in our meta-analysis to know whether the observed parietal effect is centred at a site representative of the P3 (e.g. Pz). Second, it must be borne in mind that the P3 is a large sustained potential evoked by many tasks. It appears to be broadly implicated in a cluster of related operations: context updating, surprising (e.g. oddball) events and salient and task relevant information processing (Courchesne, Hillyard, & Galambos, 1975; Donchin & Coles, 1988; Donchin, Tueting, Ritter, Kutas, & Heffley, 1975). In practice however, given its large amplitude, the P3 probably has widespread sources (Lutzenberger, Elbert, & Rockstroh. 1987), and so multiple components doubtless occupy the P3 interval. For this reason, not all effects observed in that interval need be attributed to that component. In fact it

would be highly problematic to label the P3 solely as an RPE encoder since, as shown in this study, (Supplementary Figure 3), by Sambrook and Goslin (2015, fig. 7), by Yeung and Sanfey (2004) and by many others (San Martin, 2012) there is a strong parietal encoding of absolute, unsigned prediction error size and this property is well aligned with the P3's sensitivity to oddball and salient stimuli. Since absolute prediction error size is orthogonal to both valence and RPE it is not possible for all these codings to be carried out by a single neural generator. The temporally overlapping presence of all these codings at the scalp implies temporally overlapping but spatially separate neural generators of the effect, and we thus limit our conclusions to there being a 'parietally expressed RPE encoder'. Note that we can be confident that the RPE effect does not simply arise from overlapping components coding for prediction error valence and prediction error size separately because these would sum rather than interact in the manner characteristic of an axiomatic RPE encoder.

Notwithstanding the difficulty of assigning experimental effects to existing components, an important finding is the unequivocal demonstration of RPE encoding beyond the FRN. Since this does not appear to be merely the relaying of earlier frontocentral activity, it raises the question of what a second RPE encoding achieves. Previous studies have highlighted a role for the P3 in behavioural change (Chase, Swainson, Durham, Benham, & Cools, 2011; Zhang et al., 2013). This is the ultimate purpose of RPE computation: to update action values in the light of experience and promote optimal choice. It is thus possible that the parietal effect is a representation of RPEs in the context of action value updating. If this were the case however, it might be expected that the parietal RPE would be more dependent than the frontal RPE on the task being clearly controllable. However, no site x controllability interaction was found.

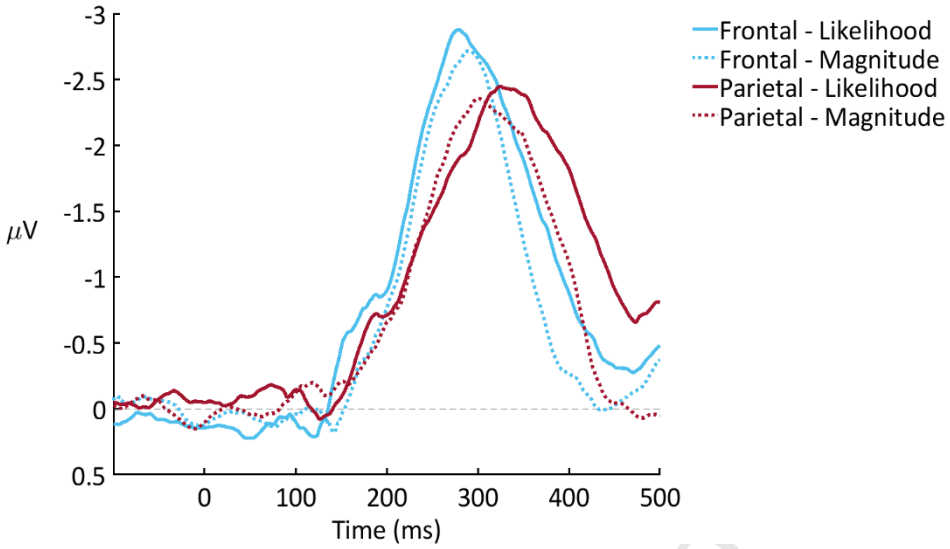
Alternatively, the parietal effect may reflect the encoding of a different kind of RPE. There has been a recent growth in studies investigating the neural correlates of model-free vs.

model-based reinforcement learning. Model-free learning, as indexed in the tasks represented in this meta-analysis, merely entails the actor maintaining action values and updating them with RPEs. In model-based learning, the actor explicitly models environmental contingencies, updating these in the light of unexpected outcomes. RPEs are not generally needed for model-based reinforcement learning. Nevertheless, research has provided evidence for neural encoding of model-based RPEs (Daw et al., 2011; Sambrook, Hardwick, Wills, & Goslin, 2018), and the parietal effect shown here may constitute such an encoding. It is not possible to test this hypothesis without a design that incorporates opportunities for both kinds of learning and which is furthermore sensitive to which kind of learning a participant is engaged in. In fact, the experiments used here do not support model-based learning insofar as they are "structureless" and a model-free process of tracking action values could not be improved upon. Nevertheless, we might expect model-based RPE encoding to occur insofar as participants are likely to have generated ad-hoc models to explain the pattern of outcomes they witness, using these to generate expectations and consequent prediction errors.

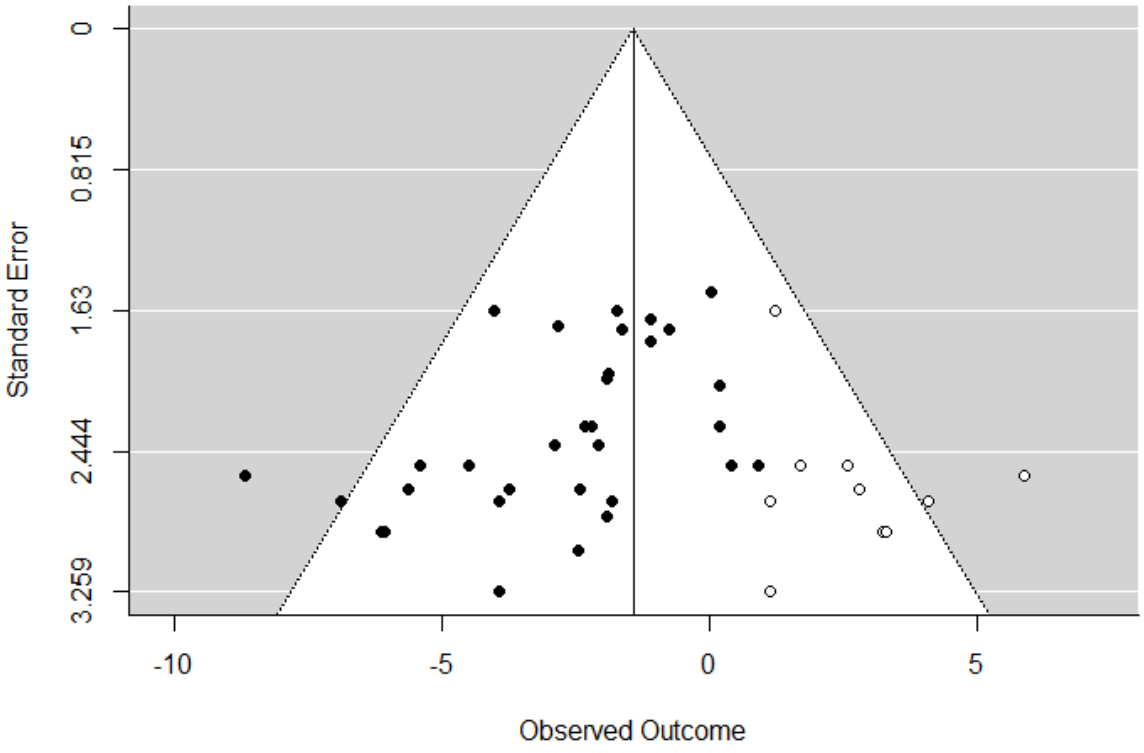
A further possibility that must be considered is that the parietal RPE encoding (and indeed the frontocentral one) is not an RPE in the strict computational sense at all, but some other construct that correlates with this quantity. Model-free learning can be thought of as a simple mathematical model of propositional reasoning (Mitchell, Houwer, & Lovibond, 2009). Alternatively, RPE effects may be better explained by episodic memory formation (Gershman & Daw, 2017; Vikbladh, Shohamy, & Daw, 2017). In both these cases, computational models employing quite different psychological constructs can lead to the same predictions as those from reinforcement learning, making them difficult to distinguish. Rather than reflecting methodological limitations, this may aptly describe the underlying

processes however, since model-based reinforcement learning is likely to be continuous with general cognition (Chater, 2009).

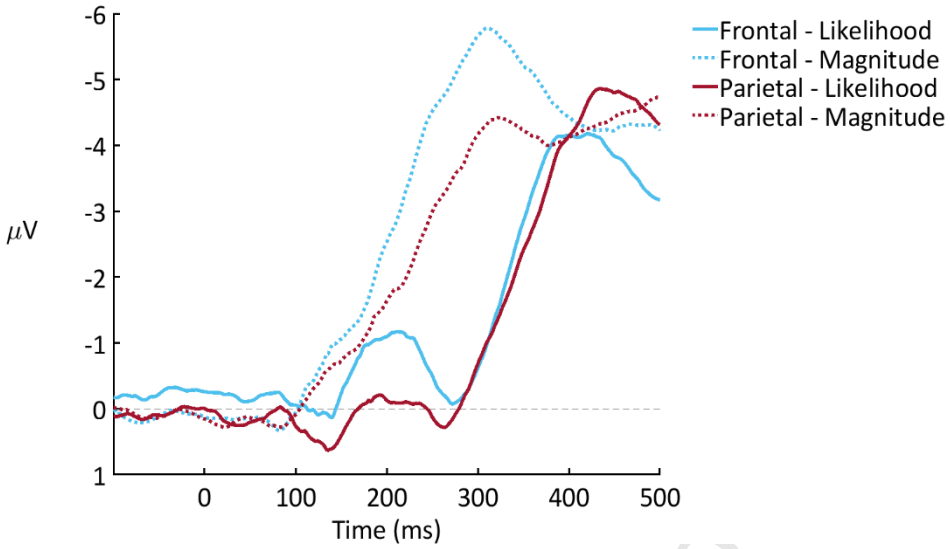
The present study presents evidence for the parietal encoding of RPEs. It demonstrates sensitivity not merely to whether outcomes are good or bad, but the degree to which this is so, thus complying with an axiomatic test for RPE encoding. It uses the great grand averaging technique to "repurpose" scientific articles that had as their principal aim the study of the FRN. While it is not the first study to assess encoding of parietal RPEs (and certainly not the first to assess the P3's response to valence), it enjoys the benefits of a large sample size and incorporates sufficient spread in tasks and other experimental details to ensure the demonstrated effect is robust. By clarifying the presence and time-course of parietal RPE encoding it is hoped that future studies will be able to use this information to better interpret feedback-locked ERPs in order to begin to reveal the circuitry underlying human reinforcement learning.



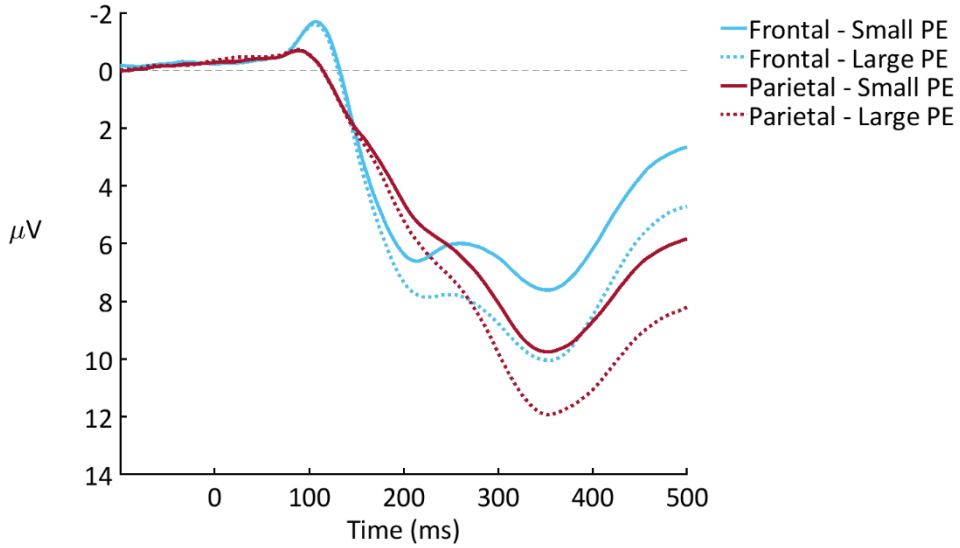
Supplementary Figure 1. Effect of valence (Figure 1a) broken down by whether likelihood or magnitude is used to modulate prediction error size.



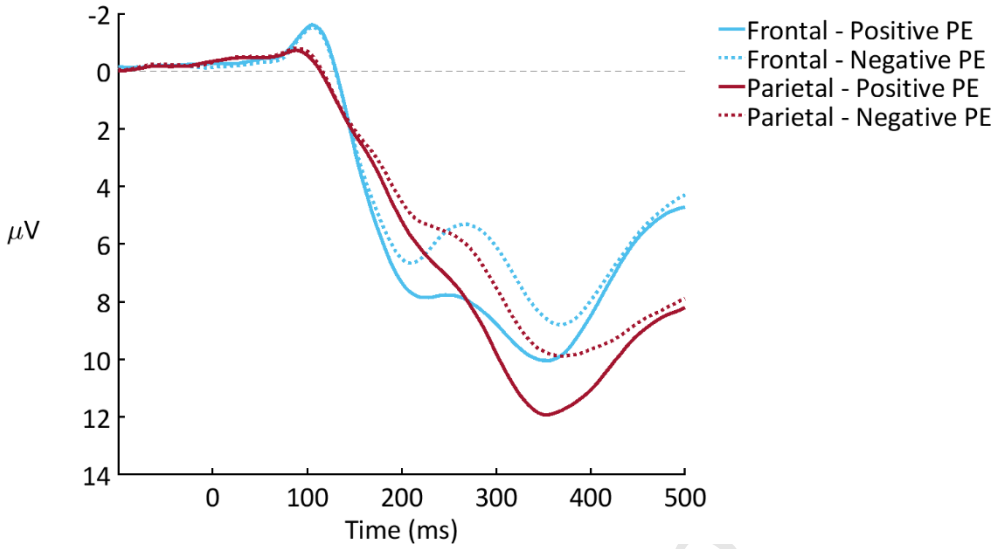
Supplementary Figure 2. Funnel plot for the parietal effect of valence at its peak of 312 ms. Imputed studies resulting from trim and fill are shown in white.



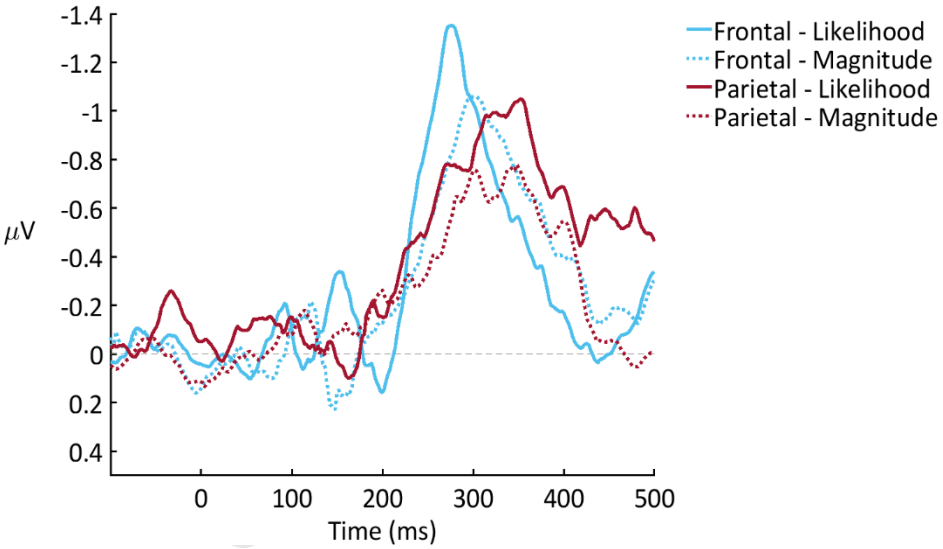
Supplementary Figure 3. Effect of prediction error size broken down by whether likelihood or magnitude is used as its modulator.



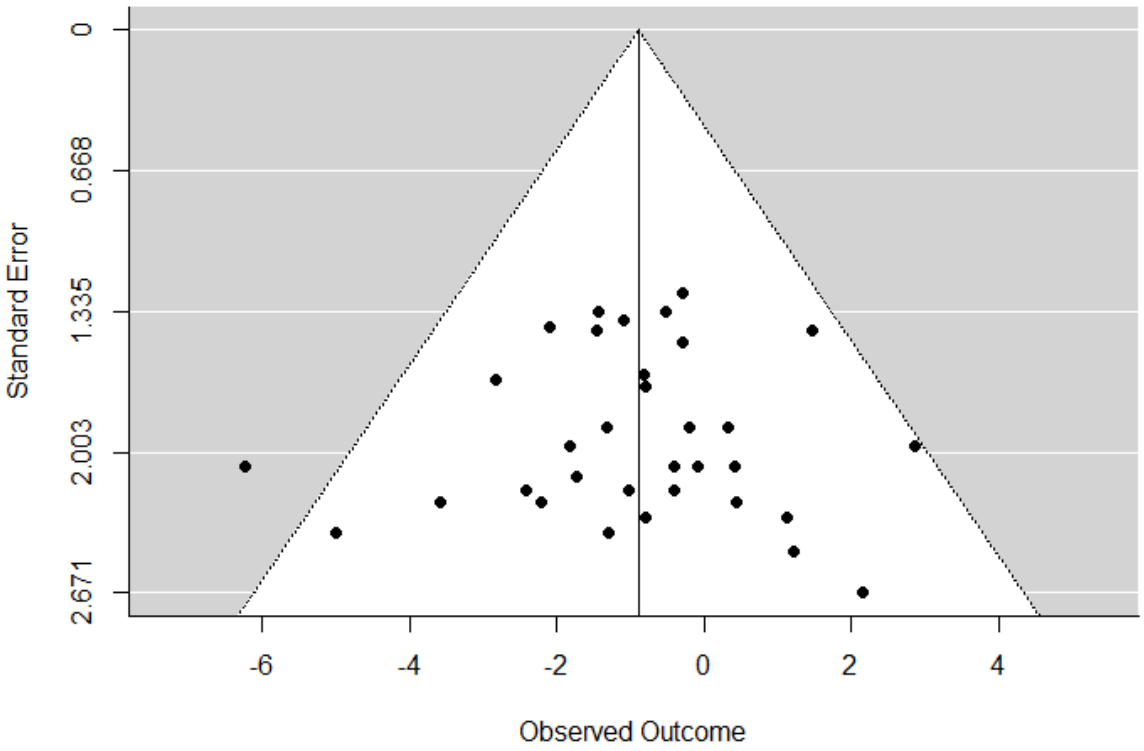
Supplementary Figure 4. Simple ERPs for prediction error size, for rewarding outcomes only.



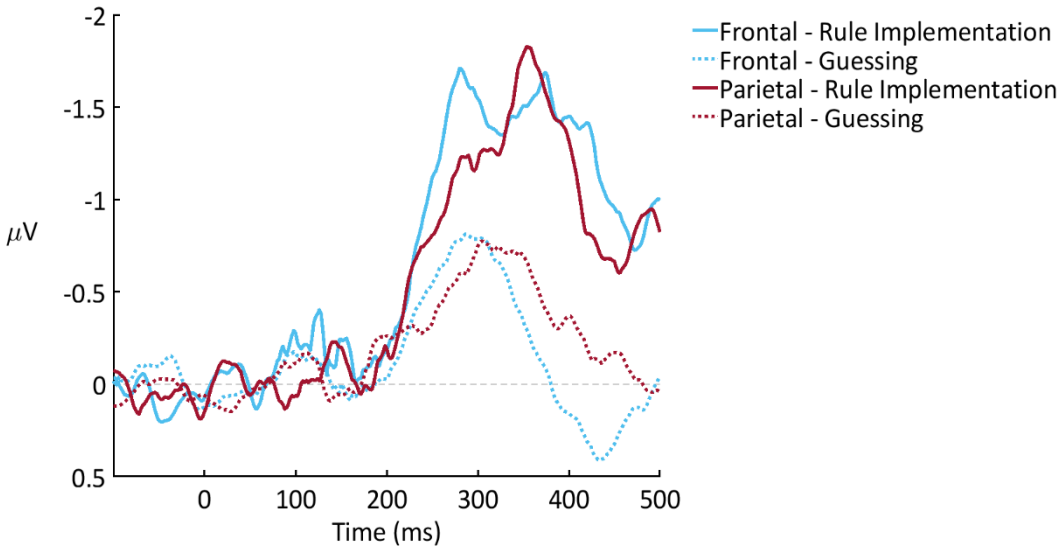
Supplementary Figure 5. Simple ERPs for prediction error valence, for large outcomes only.



Supplementary Figure 6. Effect of RPE (Figure 2a) broken down by whether likelihood or magnitude is used to modulate prediction error size.



Supplementary Figure 7. Funnel plot for the parietal effect of RPE at its peak of 354 ms. Trim and fill reveals no publication bias.



Supplementary Figure 8. Effect of RPE (Figure 2a) broken down by moderator level.

Supplementary Table 1. Studies used for the meta-analysis.

Experiment	*N*	*Modulator*	*Frontal Site/s*	*Parietal Site/s*	*Moderator Level*
Bellebaum and Daum (2008)	17	Likelihood	FC3, FCz, FC4	P3, Pz, P4	Rule Implementation
Bellebaum et al. (2011)	18	Likelihood	Cz	Pz	Rule Implementation
Hajcak et al. (2005). Expt 1	17	Likelihood	Fz	Pz	Guessing
Hajcak et al. (2005). Expt 2	12	Likelihood	Fz	Pz	Guessing
Hajcak et al. (2007)	17	Likelihood	Fz	Pz	Guessing
Hu et al. (2018)	22	Likelihood	FCz	CPz	Guessing
Kreussel et al. (2012)	24	Likelihood	Fz	Pz	Rule Implementation
Liao et al. (2011)	15	Likelihood	Fz	Pz	Rule Implementation
Pfabigan et al. (2011)	20	Likelihood	FCz	Pz	Guessing
Salim et al. (2015)	39	Likelihood	Fz	Pz	Passive
Sambrook and Goslin (2016)	42	Likelihood	FCz	Pz	Guessing
Sambrook and Goslin Unpublished	48	Likelihood	FCz	Pz	Guessing
Walentowska et al. (2016)	30	Likelihood	Fz, FCz	CPz, Pz, P1, P2	Guessing

Wu and Zhou (2009)	16	Likelihood	FCz	Pz	Guessing
Bellebaum et al. (2010)	15	Magnitude	Fz	Pz	Rule Implementation
Banis and Lorist (2012)	32	Magnitude	FCz	Pz	Guessing
Gu et al. (2011)	24	Magnitude	Fz	CPz	Guessing
Kamarajan et al. (2009)	48	Magnitude	FCz	Pz	Guessing
Kreussel et al. (2012)	24	Magnitude	Fz	Pz	Rule Implementation
Luo and Qu (2013)	18	Magnitude	FCz	Pz	Guessing
Meadows et al. (2016)	19	Magnitude	FCz	Pz	Guessing
Pfabigan et al. (2015)	31	Magnitude	Fz	Pz	Rule Implementation
Sato et al. (2005)	18	Magnitude	Fz	Pz	Guessing
Schuermann et al. (2012)	20	Magnitude	FCz	CPz	Passive
Sambrook and Goslin (2014)	55	Magnitude	Fz	Pz	Passive
Sambrook and Goslin (2016)	45	Magnitude	FCz	Pz	Guessing
Van den Berg et al. (2011)	42	Magnitude	Fz	Pz	Guessing
Wei et al. (2018)	22	Magnitude	FCz	Pz	Guessing
Wischnewski and Schutter (2018)	20	Magnitude	Fz, FC1, FC2, Cz	Cz, CP1, CP2, Pz	Guessing
Wu and Zhou (2009)	16	Magnitude	FCz	Pz	Guessing
Yu and Zhou	20	Magnitude	Fz	Pz	Guessing

(2006)					
Yu and Zhou (2008)	14	Magnitude	Fz	Pz	Guessing
Zheng and Liu (2015)	43	Magnitude	FCz	Pz	Guessing

References

Banis, S., & Lorist, M. M. (2012). Acute noise stress impairs feedback processing. *Biological Psychology*, *91*(2), 163–171. psyh. https://doi.org/10.1016/j.biopsycho.2012.06.009

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience*, *27*(7), 1823–1835. https://doi.org/10.1111/j.1460-9568.2008.06138.x

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*(12), 2874–2883. psyh. https://doi.org/10.1093/cercor/bhq038

Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2011). Processing of expected and unexpected monetary performance outcomes in healthy older subjects. *Behavioral Neuroscience*, *125*(2), 241–251. https://doi.org/10.1037/a0022536

Caplin, A., & Dean, M. (2008). Axiomatic methods, dopamine and reward prediction error. *Current Opinion in Neurobiology*, *18*(2), 197–202. https://doi.org/10.1016/j.conb.2008.07.007

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*(4), 936–946. https://doi.org/10.1162/jocn.2010.21456

Chater, N. (2009). Rational and mechanistic perspectives on reinforcement learning. *Cognition*, *113*(3), 350–364. https://doi.org/10.1016/j.cognition.2008.06.014

Courchesne, E., Hillyard, S. A., & Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and Clinical Neurophysiology*, *39*(2), 131–143. https://doi.org/10.1016/0013-4694(75)90003-6

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*(3), 357–374. https://doi.org/10.1017/S0140525X00058027

Donchin, E., Tueting, P., Ritter, W., Kutas, M., & Heffley, E. (1975). On the independence of the CNV and the P300 components of the human averaged evoked potential. *Electroencephalography & Clinical Neurophysiology*, *38*(5), 449–461. psyh. https://doi.org/10.1016/0013-4694(75)90187-X

Duval, S., & Tweedie, R. (2000). Trim and Fill: A Simple Funnel-Plot–Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Garofalo, S., Maier, M. E., & di Pellegrino, G. (2014). Mediofrontal negativity signals unexpected omission of aversive events. *Scientific Reports*, *4*, 4816. https://doi.org/10.1038/srep04816

Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, *68*(1), 101–128. https://doi.org/10.1146/annurev-psych-122414-033625

Glazer, J. E., Kelley, N. J., Pornpattananangkul, N., Mittal, V. A., & Nusslock, R. (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology*, *132*(Part B), 184–202. psyh. https://doi.org/10.1016/j.ijpsycho.2018.02.002

Gu, R., Lei, Z., Broster, L., Wu, T., Jiang, Y., & Luo, Y. (2011). Beyond valence and magnitude: A flexible evaluative coding system in the brain. *Neuropsychologia*, *49*(14), 3891–3897. psyh. https://doi.org/10.1016/j.neuropsychologia.2011.10.006

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*(2), 161–170. https://doi.org/10.1111/j.1469-8986.2005.00278.x

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*(6), 905–912. https://doi.org/10.1111/j.1469-8986.2007.00567.x

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage*, *84*, 159–168. https://doi.org/10.1016/j.neuroimage.2013.08.028

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*(4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679

Hu, X., Xu, Z., Li, Y., & Mai, X. (2018). The impact of trust decision-making on outcome processing: Evidence from brain potentials and neural oscillations. *Neuropsychologia*, *119*, 136–144. psyh. https://doi.org/10.1016/j.neuropsychologia.2018.07.036

Kamarajan, C., Porjesz, B., Rangaswamy, M., Tang, Y., Chorlian, D. B., Padmanabhapillai, A., Saunders, R., Pandey, A. K., Roopesh, B. N., Manz, N., Stimus, A. T., & Begleiter, H. (2009). Brain signatures of monetary loss and gain: Outcome-related potentials in a single outcome gambling task. *Behavioural Brain Research*, *197*(1), 62–76. https://doi.org/10.1016/j.bbr.2008.08.011

Kreussel, L., Hewig, J., Kretschmer, N., Hecht, H., Coles, M. G. H., & Miltner, W. H. R. (2012). The influence of the magnitude, probability, and valence of potential wins and losses on the amplitude of the feedback negativity. *Psychophysiology*, *49*(2), 207–219. psyh. https://doi.org/10.1111/j.1469-8986.2011.01291.x

Liao, Y., Gramann, K., Feng, W., DeáK, G. O., & Li, H. (2011). This ought to be good: Brain activity accompanying positive and negative expectations and outcomes. *Psychophysiology*, *48*(10), 1412–1419. psyh. https://doi.org/10.1111/j.1469-8986.2011.01205.x

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Luo, Q., & Qu, C. (2013). Comparison enhances size sensitivity: Neural correlates of outcome magnitude processing. *PloS One*, *8*(8), e71186. https://doi.org/10.1371/journal.pone.0071186

Lutzenberger, W., Elbert, T. H., & Rockstroh, B. (1987). A brief tutorial on the implications of volume conduction for the interpretation of the EEG. *Journal of Psychophysiology*, *1*, 81-89.

Martín, R. S. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, *6*. psyh. https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2012-33137-001&authtype=sso&custid=s8993828&site=ehost-live

Meadows, C. C., Gable, P. A., Lohse, K. R., & Miller, M. W. (2016). The effects of reward magnitude on reward processing: An averaged and single trial event-related potential study. *Biological Psychology*, *118*, 154–160. psyh. https://doi.org/10.1016/j.biopsycho.2016.06.002

Mitchell, C. J., Houwer, J. D., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*(2), 183–198. https://doi.org/10.1017/S0140525X09000855

Pfabigan, Daniela M., Seidel, E.-M., Paul, K., Grahl, A., Sailer, U., Lanzenberger, R., Windischberger, C., & Lamm, C. (2015). Context-sensitivity of the feedback-related negativity for zero-value feedback outcomes. *Biological Psychology*, *104*, 184–192. psyh. https://doi.org/10.1016/j.biopsycho.2014.12.007

Pfabigan, Daniela Melitta, Alexopoulos, J., Bauer, H., Lamm, C., & Sailer, U. (2011). All about the money—External performance monitoring is affected by monetary, but not by socially conveyed feedback cues in more antisocial individuals. *Frontiers in Human Neuroscience*, *5*. psyh. https://doi.org/10.3389/fnhum.2011.00100

Polich, J. (2003). Theoretical Overview of P3a and P3b. In J. Polich (Ed.), *Detection of Change: Event-Related Potential and fMRI Findings* (pp. 83–98). Springer US. https://doi.org/10.1007/978-1-4615-0294-4_5

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology*, *52*(4), 449-459.

R Core Team. (2013). R: A language and environment for statistical computing.

Salim, M. A. M., van der Veen, F. M., van Dongen, J. D. M., & Franken, I. H. A. (2015). Brain activity elicited by reward and reward omission in individuals with psychopathic traits: An ERP study. *Biological Psychology*, *110*, 50–58. psyh. https://doi.org/10.1016/j.biopsycho.2015.07.001

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*(1), 213–235. https://doi.org/10.1037/bul0000006

Sambrook, Thomas D., & Goslin, J. (2014). Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors. *Neuropsychologia*, *61*, 1–10. psyh. https://doi.org/10.1016/j.neuropsychologia.2014.06.004

Sambrook, Thomas D., & Goslin, J. (2016). Principal components analysis of reward prediction errors in a reinforcement learning task. *NeuroImage*, *124*, 276–286. https://doi.org/10.1016/j.neuroimage.2015.07.032

Sambrook, Thomas D., Hardwick, B., Wills, A. J., & Goslin, J. (2018). Model-free and model-based reward prediction errors in EEG. *NeuroImage*, *178*, 162–171. https://doi.org/10.1016/j.neuroimage.2018.05.023

Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *NeuroReport*, *16*(4), 407.

Schuermann, B., Endrass, T., & Kathmann, N. (2012). Neural correlates of feedback processing in decision-making under risk. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00204

Talmi, D., Fuentemilla, L., Litvak, V., Duzel, E., & Dolan, R. J. (2012). An MEG signature corresponding to an axiomatic model of reward prediction error. *NeuroImage*, *59*(1), 635–645. https://doi.org/10.1016/j.neuroimage.2011.06.051

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements*, *2*(4), i–109. https://doi.org/10.1037/h0092987

Toyomaki, A., & Murohashi, H. (2005). The ERPs to feedback indicating monetary loss and gain on the game of modified "rock–paper–scissors". *International Congress Series*, *1278*, 381–384. https://doi.org/10.1016/j.ics.2004.11.032

Van den Berg, I., Franken, I. H. A., & Muris, P. (2011). Individual differences in sensitivity to reward: Association with electrophysiological responses to monetary gains and losses. *Journal of Psychophysiology*, *25*(2), 81–86. https://doi.org/10.1027/0269-8803/a000032

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Vikbladh, O., Shohamy, D., & Daw, N. (2017). *Episodic Contributions to Model-Based Reinforcement Learning*. 2.

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*(7), 1020–1033. psyh. https://doi.org/10.1111/psyp.12651

Wei, S., Zheng, Y., Li, Q., Dai, W., Sun, J., Wu, H., & Liu, X. (2018). Enhanced neural responses to monetary rewards in methamphetamine use disordered individuals compared to healthy controls. *Physiology & Behavior*, *195*, 118–127. psyh. https://doi.org/10.1016/j.physbeh.2018.08.003

Wischnewski, M., & Schutter, D. J. L. G. (2018). Dissociating absolute and relative reward- and punishment-related electrocortical processing: An event-related potential study. *International Journal of Psychophysiology*, *126*, 13–19. psyh. https://doi.org/10.1016/j.ijpsycho.2018.02.010

Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, *1286*, 114–122. https://doi.org/10.1016/j.brainres.2009.06.032

Yeung, N., & Sanfey, A. G. (2004). Independent Coding of Reward Magnitude and Valence in the Human Brain. *The Journal of Neuroscience*, *24*(28), 6258–6264. psyh. https://doi.org/10.1523/JNEUROSCI.4537-03.2004

Yu, R., & Zhou, X. (2006). Brain responses to outcomes of one's own and other's performance in a gambling task. *NeuroReport*, *17*(16), 1747. https://doi.org/10.1097/01.wnr.0000239960.98813.50

Yu, R., & Zhou, X. (2008). To Bet or Not to Bet? The Error Negativity or Error-related Negativity Associated with Risk-taking Choices. *Journal of Cognitive Neuroscience*, *21*(4), 684–696. https://doi.org/10.1162/jocn.2009.21034

Zhang, D., Gu, R., Wu, T., Broster, L. S., Luo, Y., Jiang, Y., & Luo, Y. (2013). An electrophysiological index of changes in risk decision-making strategies. *Neuropsychologia*, *51*(8), 1397–1407. psyh. https://doi.org/10.1016/j.neuropsychologia.2013.04.014

Zheng, Y., & Liu, X. (2015). Blunted neural responses to monetary risk in high sensation seekers. *Neuropsychologia*, *71*, 173–180. psyh. https://doi.org/10.1016/j.neuropsychologia.2015.04.002

Highlights

- A frontocentral component, the FRN, is believed to code reward prediction error.
- The P3 has also been implicated in reward prediction error but findings are mixed.
- A meta-analysis using great grand averages addressed this question.
- Evidence was found for a separate parietal encoding of reward prediction error.