

2016-01

Principal components analysis of reward prediction errors in a reinforcement learning task

Sambrook, TD

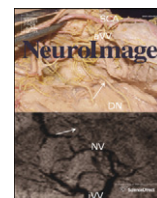
<http://hdl.handle.net/10026.1/9973>

10.1016/j.neuroimage.2015.07.032

NeuroImage

Elsevier BV

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.



Full Length Article

Principal components analysis of reward prediction errors in a reinforcement learning task

Thomas D. Sambrook*, Jeremy Goslin

Cognition Institute, Department of Psychology, Plymouth University, Plymouth PL4 8AA, UK



ARTICLE INFO

Article history:

Received 25 November 2014

Accepted 12 July 2015

Available online 18 July 2015

ABSTRACT

Models of reinforcement learning represent reward and punishment in terms of reward prediction errors (RPEs), quantitative signed terms describing the degree to which outcomes are better than expected (positive RPEs) or worse (negative RPEs). An electrophysiological component known as feedback related negativity (FRN) occurs at frontocentral sites 240–340 ms after feedback on whether a reward or punishment is obtained, and has been claimed to neurally encode an RPE. An outstanding question however, is whether the FRN is sensitive to the size of both positive RPEs and negative RPEs. Previous attempts to answer this question have examined the simple effects of RPE size for positive RPEs and negative RPEs separately. However, this methodology can be compromised by overlap from components coding for unsigned prediction error size, or “salience”, which are sensitive to the absolute size of a prediction error but not its valence. In our study, positive and negative RPEs were parametrically modulated using both reward likelihood and magnitude, with principal components analysis used to separate out overlying components. This revealed a single RPE encoding component responsive to the size of positive RPEs, peaking at ~330 ms, and occupying the delta frequency band. Other components responsive to unsigned prediction error size were shown, but no component sensitive to negative RPE size was found.

© 2015 Elsevier Inc. All rights reserved.

Introduction

A key concept of reinforcement learning is that it is driven by both reward and punishment. When rewarded, actions are more likely to be repeated, when punished, less likely. While reward and punishment might appear qualitatively very different (e.g. food vs. electric shock) reinforcement learning models reconcile them with the underlying principle of the *reward prediction error*, a numerical signed term describing the value of an outcome relative to its expected value. Positive reward prediction errors (+RPEs) indicate better than expected outcomes (i.e. rewards) while negative reward prediction errors (−RPEs) indicate worse than expected outcomes (i.e. punishments). For example, receiving a larger than expected delivery of food would constitute a reward, and would be expressed as a +RPE in a learning model. Likewise, being subjected to a small electric shock when a larger one were expected would also be expressed as a +RPE. Correspondingly, a smaller than expected food delivery or larger than expected electric shock would both constitute punishment, or a −RPE. Expressing rewards and punishments as quantitative terms, differing only by their sign, makes them commensurable and allows calculation of the net value of a course of action that will incur both rewards and punishments, producing powerful reinforcement learning algorithms. However, although potentially powerful, is this “integrated coding” the approach used by the

human brain, or are reward and punishment disassociated into separate systems?

The current evidence leans towards dissociation. For example, fMRI meta-analyses suggest that brain areas that code for both rewards and punishments are the exception rather than the norm (Bartra et al., 2013; Garrison et al., 2013). Similarly, single cell studies show that neurons that fire in response to reward or punishment only are markedly more common than those that raise their firing in response to rewards and reduce it in response to punishments (Kobayashi et al., 2006). Furthermore, it has been argued that the dissociation of reward and punishment is neuro-chemically instantiated, with dopamine coding reward and serotonin coding punishment (Daw et al., 2002).

Studies of an electrophysiological component known as feedback related negativity (FRN) have provided a valuable insight into this debate, as this component has been claimed to represent an RPE used in reinforcement learning (Holroyd and Coles, 2002). The component occurs at frontocentral sites approximately 240–340 ms after feedback, and shows a relative negativity for −RPEs compared to +RPEs. That the response is strictly an RPE, i.e. limited to a comparison of obtained value against expected value without any appraisal of error commission, was shown by Gehring and Willoughby (2002) who demonstrated that a bad outcome produces a negativity even when it is simultaneously revealed that the alternative choice would have produced a worse outcome.

The discrimination of RPE valence that the FRN shows is clearly necessary for an RPE encoder, and indeed, an axiomatic model of RPEs

* Corresponding author. Fax: +44 1752 584814.

E-mail address: tom.sambrook@plymouth.ac.uk (T.D. Sambrook).

(Caplin and Dean, 2008) states this as its first axiom of RPE encoding. The second axiom in this model requires that changes in the FRN's voltage in response to increasing RPE size should not be in the same direction for both +RPEs and –RPEs. For example, if increasing the size of +RPEs makes the FRN more positive then increasing the size of –RPEs should not also make the FRN more positive. This is equivalent to stating that there must be an interaction between RPE valence and RPE size, recently confirmed in a meta-analysis of the FRN (Sambrook and Goslin, 2015).

While this interaction is a key criterion for identifying RPE encoding in the brain, it is the nature of the interaction that is of interest in determining whether processing of reward and punishment is integrated or dissociated. If the FRN represents the activity of an integrated RPE encoder then the interaction should arise because voltage shifts negatively with increasing –RPE size and positively with increasing +RPE size (or vice versa). However, an interaction of RPE size and valence will also be observed if the FRN's voltage is responsive only to the size of *either* +RPEs or –RPEs. If the FRN were to show such a response function this would provide further evidence of the dissociation of the processing of reward and punishment.

Unpacking the interaction of RPE size and valence is thus necessary to answer the question of whether reward and punishment are integrated or dissociated in the human brain. Normally such an unpacking would be achieved by examining simple effects, that is, by examining the extent of FRN sensitivity to RPE size for +RPEs and –RPEs separately. Many papers do report such simple effects, with a recent review of the literature performed by Walsh and Anderson (2012) reporting a significantly greater sensitivity of the FRN to changes in +RPE size compared to –RPE size, i.e. a preferential sensitivity to reward rather than punishment. Such an analysis of simple effects is nevertheless unsafe because the behaviour of single waveforms, on which simple effects are based, is subject to unknown component overlap. This can effect a positive or negative translation on only some of the single waveforms, distorting the pattern of simple effects. An example is given in Fig. 1, which shows an idealised FRN as an integrated encoder of both +RPEs and –RPEs, overlaid with a component that codes for an absolute, or unsigned prediction error. This latter component does not encode reward at all, it responds to the mere salience of the outcome, that is, the prediction error's absolute size. The summed effect of these components is a waveform which appears sensitive only to +RPE size. There is now mounting evidence of such salience encoding components in the same temporal interval as the FRN (Hauser et al., 2014; Talmi et al., 2013) and such components therefore stand to account entirely for the apparent preferential sensitivity of the FRN to +RPEs shown in Walsh and Anderson's review.

Because the true response function of the FRN will remain uncertain as long the component remains overlain by other components, a logical step would appear to be to separate out components using a technique such as principal components analysis (PCA) prior to examining simple effects. Some previous attempts have been made in this regard. For example, Holroyd et al. (2008) concluded that the FRN responded only to +RPEs after spatial PCA revealed a frontocentral factor with this behaviour. However these authors did not follow the spatial step of the PCA with a further temporal step. Given surprisingly high variance explained by this factor (>50%), it is likely that this would have allowed the separation of further frontocentral components. Foti et al. (2011) did conduct a temporospatial PCA designed specifically to resolve the question of the relative sensitivity of the FRN to +RPEs and –RPEs. These authors found a factor whose temporospatial profile resembled the FRN, and which responded only to +RPEs. Crucially however, the experiment employed simple dichotomous good vs. bad feedback, and RPE size was not varied. This meant that the requirements of the axiomatic model of RPE encoding could not be satisfied, as this requires an interaction between RPE size and valence.

This highlights a dichotomy within the literature between the FRN's operationalisation and its proposed function. This component is

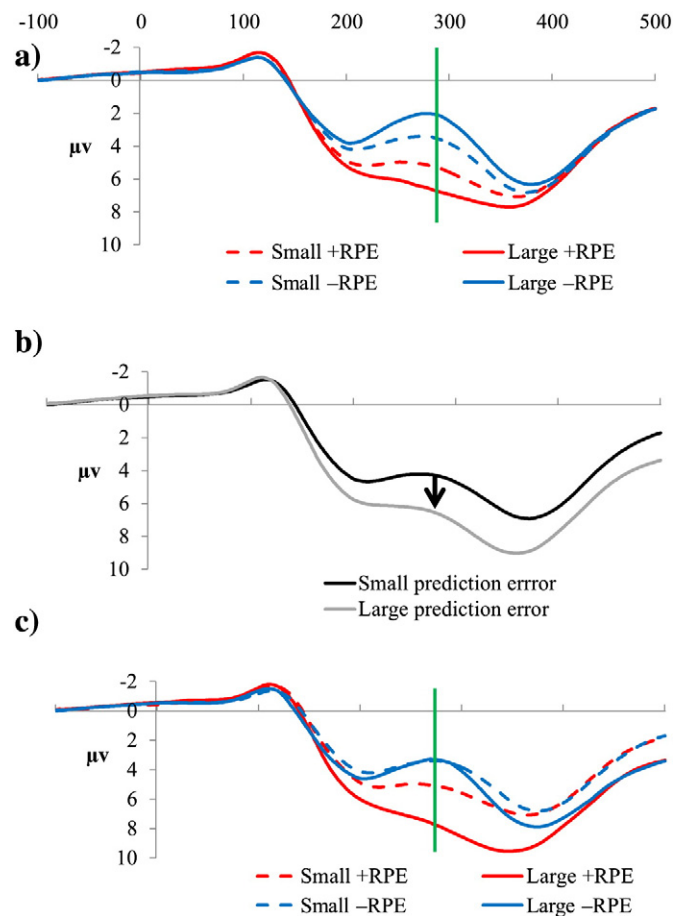


Fig. 1. How component overlap undermines simple effects analysis of the FRN (measured at the green line). a) How the FRN might look in simple waveforms unaffected by component overlap. In this hypothetical case the FRN shows integrated coding, with both –RPEs and +RPEs modulated by their size, on a common utility scale where positive voltages are good. b) a separate component lies in the same interval which simply codes for the absolute size of the prediction error, i.e. its salience. Large prediction errors shift the waveform positively by $\sim 2 \mu\text{V}$. c) The observed ERP waveforms. The effect of the salience encoding component is to shift large –RPEs and large +RPEs positively, producing equivalent voltages for small and large –RPEs while increasing the difference between the voltage of small and large +RPEs. Analysis of simple effects would result in the erroneous conclusion that the FRN responds to the size of +RPEs only.

characterised as being sensitive to the valence of an outcome, with this sensitivity maximal over frontocentral electrodes and in the approximate interval 200–350 ms after feedback is received. These are the necessary and sufficient properties for its identification and, indeed, it is typically operationalised simply as a difference wave of good and bad outcomes. Sensitivity to RPE size does not constitute part of the operationalisation. However, the theoretical claim of Holroyd and Coles that the FRN constitutes an RPE encoder *does* require observation of its response to RPE size, by the terms of the axiomatic model. Thus Foti et al.'s (2011) PCA study successfully isolated the FRN as typically defined but nevertheless cannot be said to have successfully isolated an RPE encoding component. The distinction is far from academic because in our meta-analysis we showed that the interval of the feedback-locked ERP that shows an RPE size \times valence interaction is considerably smaller than the interval showing a main effect of valence (Sambrook and Goslin, 2015). This suggests the presence of multiple valence sensitive components, of which only some may meet the axiomatic requirements for an RPE encoder.

Thus, in this study our aim was to establish the axiomatic verities of the FRN by conducting a temporospatial PCA analysis of this component, similar to that of Foti et al.'s study, but including a factor of RPE size. To

best implement this test, we parametrically varied RPE size across a wide range of values. Then, rather than attempting to include the different levels of RPE size in a standard factorial design, we examined the strength of the correlation between voltage and RPE size at successive sample points. Details of this process are given in the methods below and also in [Sambrook and Goslin \(2014\)](#) where we previously used the technique. Using such a correlation waveform has a number of advantages. First, it harnesses the extra power of parametric designs over factorial ones ([Cohen, 1983](#)). Second, it tests for a monotonic relationship between a continuous independent variable and a component's amplitude, something frequently assumed, but not demonstrated, in factorial designs. Third, this also has the side effect of greatly simplifying the presentation of the results, as the effect of RPE size on voltage can be represented by a single correlation waveform, rather than separate ERP waveforms for each factor (30 levels in the present case). These advantages pertain regardless of whether a following PCA is performed. However, when it is, the PCA is rendered more powerful. This is because when factorial designs are entered into a PCA the relationship between conditions is not specified prior to extraction of components, with each treated separately until later factor analysis. However, ignorance of the structure of the design can result in misallocation of variance during the PCA, especially in noisy data. Replacing a factor (here RPE size) with its *effect size* (here Pearson's r) ensures that this information is available to the PCA, ensures that all extracted factors explain variance in that factor, and no variance is "wasted" extracting factors that are merely obligatory responses to the arrival of a stimulus on a screen.

In the particular case where overlapping components have different frequency characteristics, an additional means by which they may be separated is to analyse the ERP in separate frequency bands. It has been recently proposed, for example, that the FRN might reflect a theta response to $-$ RPEs superimposed on a delta response to $+$ RPEs ([Bernat et al., 2015](#); [Foti et al., 2014](#)), while other recent papers have suggested theta is in fact a salience response ([Hajihosseini and Holroyd, 2013](#); [Mas-Herrero and Marco-Pallarés, 2014](#)). The main analyses of the study are therefore supplemented with analyses conducted on theta and delta separately.

A number of other measures were taken to isolate the FRN and any other RPE encoding components as effectively as possible. Since RPE size is a function of both the magnitude of a reward or punishment and its prior likelihood, an RPE encoding component should be responsive to variations in either of these properties. Thus to ensure we were observing an RPE encoder rather than simply a response to likelihood or magnitude, we manipulated both these properties to ensure they produced comparable effects. We also counterbalanced the *domain* of the outcome, whether it constituted a monetary loss or gain to the participant's bankroll. An outcome of zero, or even a small loss, is still a $+$ RPE if a larger loss was the expected value for the trial. Since the FRN's response to RPEs has been shown to be modulated by the domain in which they occur ([Kreussel et al., 2012](#); [Kujawa et al., 2013](#); [Sambrook et al., 2012](#); [Yu and Zhang, 2014](#); [Zheng et al., 2015](#)) half the trials were performed in the gain domain, with participants attempting to maximise their gain, and half in the loss domain where they attempted to minimise their loss.

It is also important to note that the design ensures against attributing an RPE response to what is really a salience response as a result of correlations between signed and unsigned prediction error values. This concern was raised by [Hauser et al. \(2014\)](#) who used multiple regression to show that an apparent response to valence was in reality a response to salience, arising from correlations between signed and unsigned prediction errors. The present design produced a distribution of $-$ RPEs and $+$ RPEs that was highly symmetrical around a value of zero, ensuring a negligible correlation of signed and unsigned prediction error size. Moreover, and as we detail below, salience and RPE responses are here defined by mutually exclusive criteria: salience responses are indicated by voltage changes in the *same polarity*

for both $+$ RPEs and $-$ RPEs (with respect to increases in RPE size) and RPE responses are shown by any other pattern of sensitivity.

Methods

Participants

The study was approved by the ethics committee of the Faculty of Health and Human Sciences at the University of Plymouth. Eighty seven (23 male) students of the University of Plymouth participated for course credit and an opportunity to win money. All were right handed, under 29 years (mean age = 21.1), had no history of neurological damage or other significant health problems, and were not on medication at the time of the experiment. No other information was recorded. Forty five participants were tested with RPEs manipulated by varying outcome magnitude, and forty two with RPEs manipulated by varying outcome likelihood.

Task

Participants took part in a probabilistic reinforcement learning task. On each trial, they selected one of two keys and were then given reward feedback. They were informed that one key gave on average slightly better feedback (by the terms of the block, see below) than the other and that they should pay attention to the feedback they received so they might learn which key to prefer for the block's duration. Trials were presented in blocks of sixty, and the participant told that at the start of each block (clearly indicated to them) the good key would be randomly re-selected, requiring a new learning episode. In fact, feedback was pseudo-randomly predetermined, and the participant could not influence the outcome by key selection.

For magnitude participants, feedback on a given trial took the form of a number between 23 and 82. Furthermore, each block of sixty trials was designated a gain or loss block. On gain blocks, participants attempted to accrue a points total equal or exceeding 3000 by the end of the block since they received £2 for doing so and nothing if they remained under 3000. In these blocks large numbers were therefore $+$ RPEs and small numbers $-$ RPEs. On loss blocks, participants attempted to *avoid* accruing a points total equaling or exceeding 3000 since they lost £2 for doing so and lost nothing if they remained under. In these blocks large numbers were $-$ RPEs and small numbers $+$ RPEs. In reality, and regardless of the block's domain, the same sixty numbers were presented: each of the values from 23 to 82 inclusive, with this fact concealed by reporting a fictional sum to participants at the block's end, predetermined to exceed and fall under the target equally often for both gain and domain blocks. Thus the magnitude of a number and its RPE sign was reversed between gain and loss blocks (ensuring no confounding of magnitude and valence) and $+$ RPEs and $-$ RPEs were equally prevalent on both loss and gain blocks (ensuring no confound of domain and valence).

Sixteen gain domain blocks were run followed by sixteen loss domain blocks (sixty trials per block), with this order reversed for half the participants. A blocked design was used for the domain variable since the study sought only to control, not study, domain effects and it was believed that alternating domain on a trial by trial basis would confuse some participants and reduce FRNs generally as feedback stimulus–reward associations were continually being reversed. We also expected that blocking domain would reduce its effects generally.

For likelihood participants, feedback on a given trial was binary: either a one or a six. At the start of each block, a guideline probability for sixes was given, which could take the value 25% (low), 50% (medium) or 75% (high). A corresponding target number of sixes for the block was also given: 15, 30 and 45 respectively. On gain blocks, participants attempted to reach or exceed the target number of sixes by the end of the block since they received £2 for doing so and nothing if they did not. In these blocks sixes were thus $+$ RPEs and ones $-$ RPEs.

On loss blocks, participants attempted to avoid reaching the target number of sixes since they lost £2 for doing so and lost nothing if they did not. Once again, in reality neither key was set to have a higher chance of returning sixes, with feedback pseudo-randomly predetermined. Since it was possible that participants could maintain a count of the number of sixes accrued on a block, deceit was not attempted in reporting this number; it was faithfully reported but predetermined to reach or fall short of the target equally often for both gain and loss blocks. As with the magnitude participants, stimulus magnitude, RPE sign and domain were fully counterbalanced.

Fifteen gain domain blocks were run followed by fifteen loss domain blocks (sixty trials per block), with this order reversed for half the participants. Within each domain, five sub-blocks were presented for each of the three baseline probabilities for sixes. The order of these sub-blocks was counterbalanced.

To provide participants with an additional incentive to attend to feedback, at the end of each block they were required to perform an estimation task. Magnitude participants were asked whether they believed they had exceeded or fallen short of 3000, and were awarded £0.20 for the “right” answer (i.e. that which corresponded with the fictional report). Likelihood participants were asked to report which key they believed had performed better, and were awarded £0.10 for the right answer: although neither key was actually set to be better one would typically show a purely incidental superiority over a block.

The experimental task was presented using E-Prime software and is summarised in Fig. 2. Participants were shown a prompt and were required to press the right or left key of a key pad, using either hand. A fixation cross appeared (600–700 ms duration, with a uniform distribution), followed by the outcome stimulus (700 ms duration) and then a blank screen (800 ms duration).

Final earnings from the reinforcement learning task were zero for all subjects. Due to the end of block estimation questions, total earnings averaged £2.12 for magnitude participants and £1.96 for likelihood participants.

Modelling of reward prediction errors

RPEs were modelled as follows. At the beginning of each block, the expected value of the first trial was set to the mathematical expectation

given the parameters for that block (always 52.5 under magnitude modulation and 2.25, 3.5 or 4.75 under likelihood modulation, depending on the baseline probability of sixes on that trial). Thereafter, expected value was computed using a standard Rescorla–Wagner algorithm (Rescorla and Wagner, 1972). For each trial, t , an RPE was computed as the difference between the actual outcome value $R(t)$ and its expected value $V(t)$ on that trial

$$\text{RPE}(t) = R(t) - V(t).$$

The expected value of the next trial $V(t + 1)$ was updated by adding the RPE, weighted by a learning rate, α , set to 0.3. This value was arbitrary, but in keeping with similar studies of reinforcement learning (Metereau and Dreher, 2013; Seymour et al., 2005), and comparable with a value extracted in a recent study of non-reversal probabilistic learning and the FRN (Mas-Herrero and Marco-Pallarés, 2014).

$$V(t + 1) = V(t) + \alpha \cdot \text{RPE}(t)$$

EEG recording

EEG data were collected from 61 Ag/AgCl active electrodes (actiCAP, Brain Products, Gilching, Germany) mounted on an elastic cap and arranged in a standard International 10–20 montage. Electrodes were referenced to the left mastoid and re-referenced off-line to the average of left and right mastoid activity. Vertical eye movement was monitored by a right suborbital electrode, and horizontal eye movement was monitored using an electrode on the right external canthus. Electrode impedances were kept below 20 k Ω . EEGs were amplified using a BrainAmp amplifier (Brain Products), continuously sampled at 500 Hz, and filtered offline with 60 and 50 Hz notch filters followed by a .1 Hz high pass filter and 30 Hz low pass filter. EEGs were time-locked to 100 ms before the onset of the feedback to 700 ms afterward, and then were baseline-corrected using the interval –100–0 ms. Eye movement artefacts were removed using a criterion of a voltage change exceeding 75 μV /200 ms in eye electrodes. Other non-specific artefacts were removed using a criterion of any electrode showing either a voltage change exceeding 40 μV /ms, a voltage value exceeding $\pm 200 \mu\text{V}$ relative to baseline, or activity across the epoch of $< 2.5 \mu\text{V}$. The percentage of trials retained was 85.8%. The resulting mean number of valid trials, and standard errors for each of the conditions, were as follows: magnitude + RPEs 803 (16.01), magnitude – RPEs 795 (16.11), likelihood + RPEs 755 (17.68), and likelihood – RPEs 748.81 (18.34). The minimum number of valid trials for any participant for any of the conditions was 295.

Electrodes which malfunctioned in the course of an experiment were substituted using topographic interpolation (Perrin et al., 1989). This produced an average incidence of topographic interpolation of 1.06%, with no more than three electrodes substituted for any participant.

Data analysis

ERPs and correlation waveforms

ERPs showing the categorical effect of valence, i.e. + RPEs vs. – RPEs, were created by conventional means, that is by averaging all trials of the given valence and then grand averaging those across participants. Analysis of the FRN was performed by comparing mean amplitude for + RPE vs – RPE waveforms at FCz and in the interval 240–340 ms, a period shown to provide an appropriate representation of the FRN in a recent meta-analysis (Sambrook and Goslin, 2015). Subsequent analyses however, were based on the parametric variation of RPEs, with separate correlation waveforms created for trials with – RPEs and + RPEs, following Sambrook and Goslin (2014). A Pearson correlation of RPE

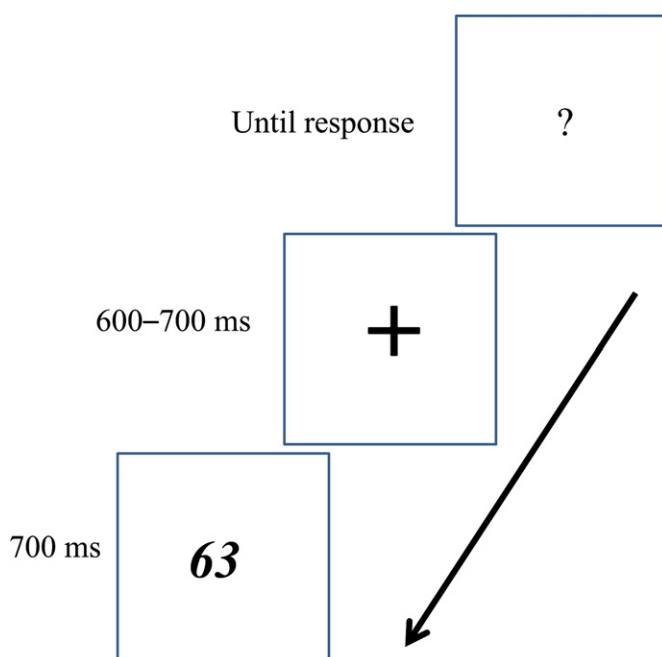


Fig. 2. Summary of a single trial of the task procedure.

size and voltage was performed across all trials performed by a particular participant, with a separate correlation calculated for each sample point from -100 to 700 ms, and for each of the active electrodes. These values of r were then re-plotted against time, generating, for that participant, a waveform in which the y axis represented, not voltage, but the strength of RPE size encoding at a particular time and electrode. The interpretation of such a figure is straightforward: where the waveform is at baseline there is no effect of RPE size on voltage, where there are deviations from baseline this indicates a relationship, suggesting that RPE size is being coded by voltage. Insofar as they show sensitivity to an experimental variable these correlation waveforms are in fact very similar to difference waves and should be interpreted in the same manner. Note that the correlation coefficients are small because they are derived from single trial data.

PCA

PCA was performed on these waveforms using the ERP PCA Toolkit Version 2.41 (Dien, 2010a) using identical means to those used by Foti et al. (2011) and following published guidelines (Dien, 2010b; Dien et al., 2005, 2007). Separate PCAs were performed for magnitude modulated and likelihood modulated RPEs, to ensure putative RPE encoding components such as the FRN were elicited in both cases, as theoretically required. First, a temporal PCA was performed using each time point as a variable and each combination of participant, electrode, and condition (i.e. + RPE vs. – RPE) as observations. This produced 5490 observations for magnitude, and 5124 for likelihood. Factors were retained if they explained more variance than a factor extracted from a null dataset, i.e. passed a parallel test (Horn, 1965) and these were subjected to Promax rotation. Following this step, factors peaking prior to 50 ms post feedback, or factors showing a non-significant response to both + RPEs and – RPEs at the factor's peak, were discarded as likely artefacts, as were factors which explained less variance than these. The remaining temporal factors were then entered into a spatial PCA using Infomax rotation. In this step, electrodes were used as variables, and each combination of participant, condition and temporal factor used as observations. This produced 90 observations for magnitude and 84 for likelihood.

Following the method of Dien et al. (2003), factors were then reconstructed into waveforms using the product of the factor pattern matrix and the standard deviations. These could then be interpreted in the same manner as the correlation waveforms depicted in Fig. 5: indeed, waveforms in Fig. 5 are simply the sum of the factors shown in Fig. 6, minus the residual factors that were extracted by the PCA but then discarded. This demonstrates an important feature of our analysis: because the correlation waveforms entered into the PCA describe sensitivity to RPE size, all factors extracted from these waveforms by the PCA necessarily constitute components sensitive to RPE size, with the degree of sensitivity indicated by the factor's amplitude (denominated in units of r). Since the rationale of the study is to isolate RPE encoding components and then compare their differential sensitivity to the size of + RPEs and – RPEs, this can be simply achieved by comparing the amplitude of the factor over + RPE and – RPE conditions.

For each extracted factor, this comparison was performed as follows. For each participant, the factor's sensitivity to + RPEs and – RPEs was scored by its peak amplitude under those conditions, following Foti et al. (2011). These scores were not directly compared (e.g. with a paired samples t -test) because *relative* sensitivity could not be used to characterise the factor's response function. Instead, the *presence* or *absence* of sensitivity to RPE size, along with its *sign* was established separately for + RPEs and – RPEs using one sample t -tests against a test value of zero. The resulting pattern of significance over the two tests could be used to classify the component into one of four mutually exclusive and exhaustive categories. Thus, significant sensitivity for + RPEs only would indicate a + RPE encoder, for – RPEs only a – RPE encoder. If the factor was sensitive to both RPE valences, but responded in opposite polarities for + RPEs and – RPEs it would be described as an

integrated coder. On the other hand, if a factor responded with the same polarity to both RPE valences, it could not be described as an RPE encoder at all, but rather a salience encoder.¹ In some cases, factors responded in the same polarity but with different strengths to + RPEs and – RPEs, typically stronger for the former. This differential sensitivity notwithstanding, such a factor could not be described as an RPE encoder because it could not distinguish between the two valences, producing an equivalent output for small + RPEs and large – RPEs and thus failing the axiomatic model. Nevertheless, to assess the degree of this asymmetrical response, these salience encoding factors were additionally subjected to a paired-samples t -test to establish the degree of differential sensitivity to + RPEs and – RPEs.

Source localisation

After extraction of factors, source localisation was performed using Robert Oostenveld's FieldTrip (<http://fieldtrip.fcdonders.nl/start>), which was interfaced directly from within the ERP PCA Toolkit. The procedure closely followed Foti et al. (2011). This used a four-shell model to specify a pair of hemispheric dipoles. To reduce the likelihood of finding a solution for a local minimum, a grid scan was performed on the head space to identify the best starting position for the dipoles. An iterative algorithm then moved dipole positions until maximum fit was achieved under a maximum-likelihood estimation algorithm. The entire epoch was used for the fitting process. Stability of the solution was assessed using the ERP PCA toolkit's jack-knife technique, in which the spatial step of the PCA was recomputed as many times as there were participants, with a single participant omitted each time, allowing examination of the extent to which the solution depended on particular participants. Additionally, distributed source analysis was performed using low resolution brain electromagnetic tomography (sLORETA; Pascual-Marqui, 2002). This procedure made no assumptions about the number of sources, simply computing the smoothest possible solution. An unconstrained solution was applied to all grey cortical and hippocampal matter (sLORETA localisation is limited to these regions), with the source voxels only considered if their amplitude exceeded 85% of the global maximum.

Separation into theta and delta frequency bands

Single ERP trials were re-filtered into theta and delta bands using 4–7 Hz and 0.1–4 Hz bandwidth pass filters respectively (48 dB roll off). Separate analyses of the filtered waveforms proceeded as above, with the exception that power, rather than amplitude was used as the dependent variable. The creation of correlation waveforms and their entry into PCA were also performed using the same methodology as that described previously.

Results

Behavioural results

Because the relationship of outcome to key was entirely random, any profitability of one key over the other was transient and the expected values derived from the model were in no way an indicator of differences real long term expectation. However, since participants had been told that in each block one key was programmed to be slightly more favourable than the other one, and were under a monetary incentive, we expected they would track expected values and act on them.

¹ In a previous paper (Sambrook and Goslin, 2014) we investigated response functions using the same rationale of separately testing sensitivity to + RPEs and – RPEs. However, in that study we correlated voltage against RPE utility rather than RPE size. As utility and size are inversely related in – RPEs, and positively related in + RPEs this meant that salience responses in that study were represented by oppositely signed + RPE and – RPE waveforms. In this study we have changed our convention to allow a direct comparison of our data with that of the study of Foti et al. (2011).

This was confirmed. Participants were more likely to choose the key which at that moment in time enjoyed higher expected value (magnitude participants: 71% correct key choices $\chi^2(1) = 14,237.98$, $p < .001$; likelihood participants: 68% correct key choices, $\chi^2(1) = 9573.44$, $p < .001$). As the difference of expected value between the two keys grew, the likelihood of selecting the better key grew accordingly as shown in Fig. 3. In addition, an analysis of key switches showed that these were more likely if the previous trial incurred a –RPE rather than a +RPE (magnitude participants; $\chi^2(1) = 3093.58$, $p < .001$; likelihood participants: $\chi^2(1) = 1843.79$, $p < .001$). Key switches also became more likely as the expected value of the two keys grew closer, as shown in Fig. 3. Collectively, these results suggest that participants persevered with previously favourable keys until their expected value dropped close to the alternative. As such, their behaviour confirmed that they were engaged in reinforcement learning.

ERPs

Grand average waveforms shown in Fig. 4 show the relative negativity for bad outcomes that is characteristic of the FRN, with this effect strongly significant in the interval 240–340 ms (magnitude modulated FRN $t(44) = 5.69$, $p < .001$, likelihood modulated FRN $t(41) = 5.28$, $p < .001$). Fig. 5 provides correlation plots showing the strength of encoding of –RPEs and +RPEs. Axes are plotted in reverse by convention. For the majority of the epoch, the correlation between RPE size and voltage is positive for both +RPEs and –RPEs, indicating that increases in the size of both rewards and punishments produce a more positive feedback-locked waveform. This suggests that the underlying components are salience encoders, not RPE encoders. The exception is as ~330 ms where +RPEs retain a correlation with voltage but –RPEs show no correlation, suggesting that at this point an RPE encoding occurs, but for +RPEs only. This hypothesis was investigated in the PCA analysis below.

PCA components

The temporal step of the PCA resulted in 11 temporal factors for magnitude modulated RPEs and nine for likelihood modulated RPEs. In the spatial step, three factors were extracted under both modulators producing a total of 33 and 27 temporospatial factors for magnitude and likelihood modulated RPEs respectively.

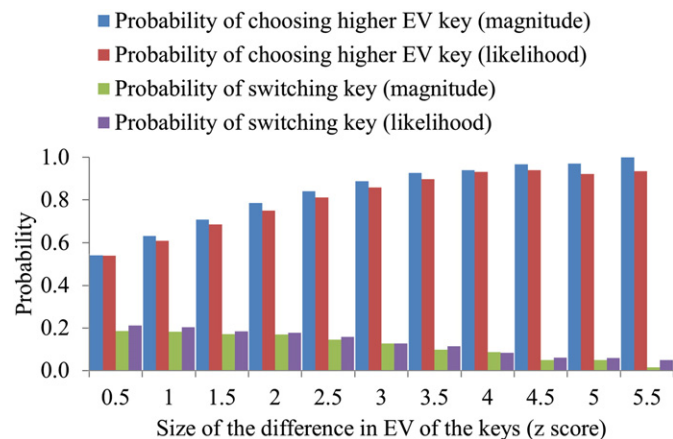


Fig. 3. Behavioural data showing participants' likelihood of selecting the key with higher expected value, and their likelihood of switching keys, as a function of how great the difference in expected value was between keys. This difference was computed as the expected value of the better key divided by that of the poorer key. To allow comparison across magnitude and likelihood modulators this was then transformed to a z score (non-mean-corrected so that in all cases a z score of zero means identical expected values for the two keys).

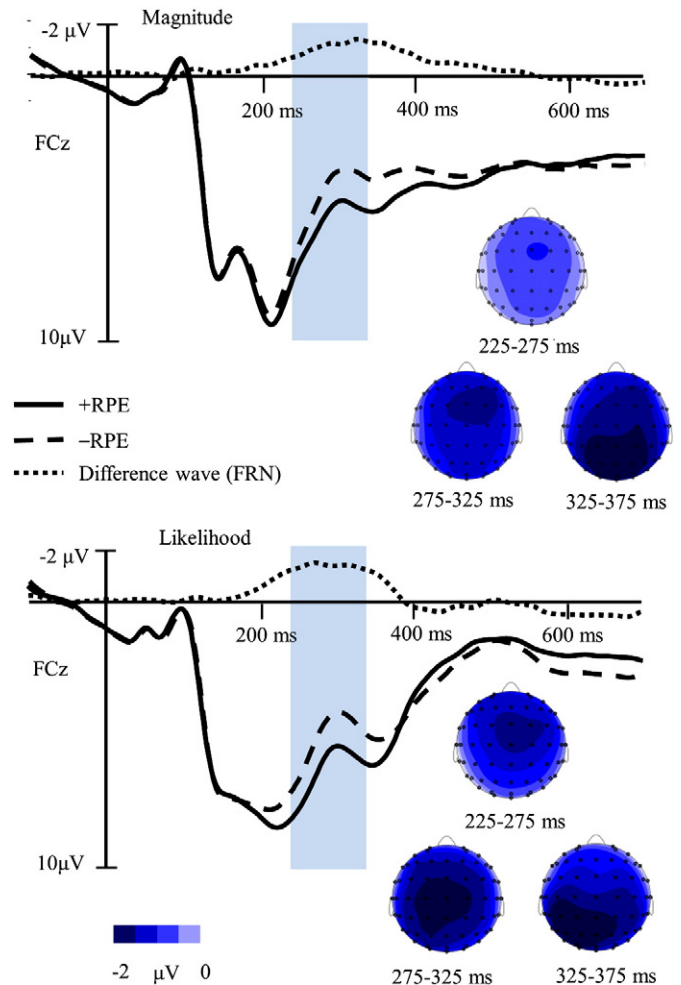


Fig. 4. The FRN under modulation by magnitude and likelihood, measured at FCz. RPE size is disregarded in this plot, the plots simply show the grand average of all +RPEs vs. all –RPEs, with the associated difference wave (–RPE minus +RPE). The shaded area indicates the measurement interval of 240–340 ms. Scalp topographies of the difference wave are shown for the interval of measurement and adjacent intervals.

The first four factors from each temporospatial PCA were selected for plotting and analysis on the basis that they accounted for three times more variance than the remaining factors. These factors are summarised in Table 1 and plotted for +RPEs and –RPEs separately in Fig. 6. Of the remaining factors, the first temporospatial factor for each temporal factor is shown in Supplementary Fig. 1.

Proposed FRN PCA component

Under both modulators, factor TS3/SF1 resembled the temporospatial profile of the FRN (see Table 1). One sample t-tests conducted at the peak of this factor's activity revealed it to be significantly responsive to +RPE size (magnitude $t(44) = 6.71$, $p < .001$; likelihood $t(41) = 7.97$, $p < .001$) but not to –RPE size (magnitude $t < 1$, likelihood $t < 1$). The response function of this factor, and its common occurrence in response to both magnitude and likelihood modulated RPEs was thus consistent with an RPE encoding function, albeit only of +RPEs. The slightly increased latency of the factor's peak under magnitude modulation (344 ms compared to 322 ms for likelihood, see Table 1) may be attributable to the greater complexity of processing a number in a sixty point range compared to the simple dichotomous stimuli of the likelihood experiments.

While a separate analysis of magnitude and likelihood modulated RPEs was necessary to establish their independent generation of the component, having established this, we collapsed likelihood and

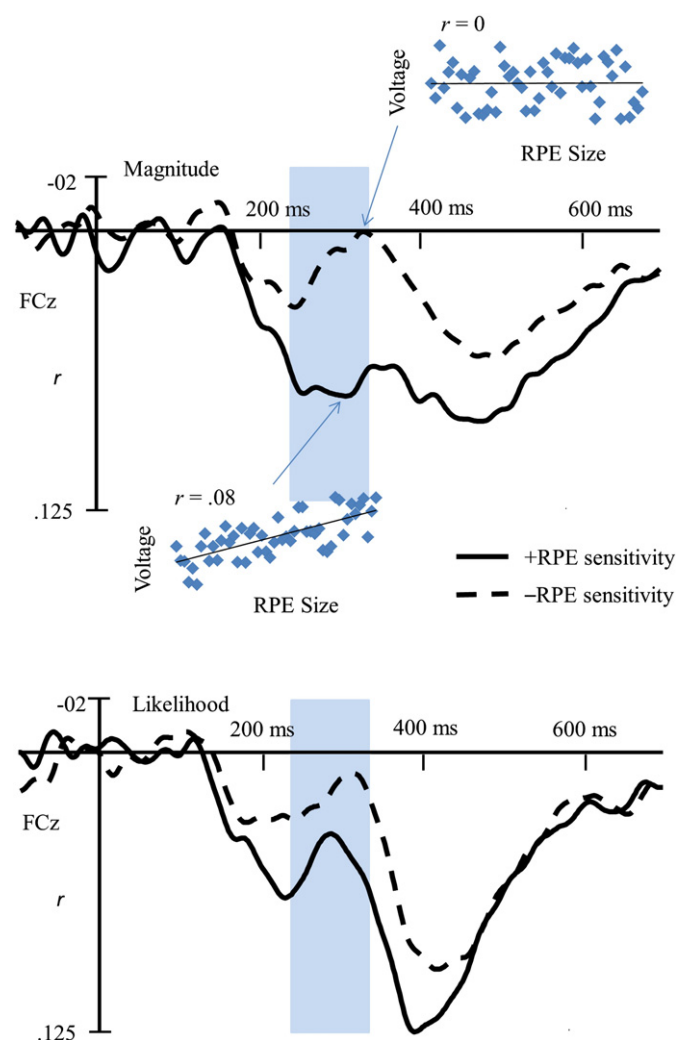


Fig. 5. Grand average of participant correlation waveforms of the strength of the relationship between RPE size and voltage determined by Pearson correlation of voltage and RPE size over single trials on a sample-by-sample basis (note the y axis is denominated in units of r). Correlations are performed separately for +RPEs and –RPEs. Figure insets are used to illustrate the correlation that lies behind a single datum on the correlation waveform: in these insets voltage is shown as a function of RPE size, with their correlation, or its absence illustrated by a regression line. Shading shows the interval 240–340 ms to allow comparison with the FRN in Fig. 4.

magnitude data together and re-extracted the factor to maximise the power of source localisation. This produced a good quality solution (residual variance = .98%) which identified a source in the putamen, shown in Fig. 7. (Talairach co-ordinates: 29.44, –18.27, –3.07). Application of the jack-knife technique showed this solution was extremely stable, with the standard deviation for these co-ordinates being 0.60, 0.54 and 0.42. The dipole was significantly larger in the right hemisphere $t(87) = 77.13$, $p < .001$. Since distributed source

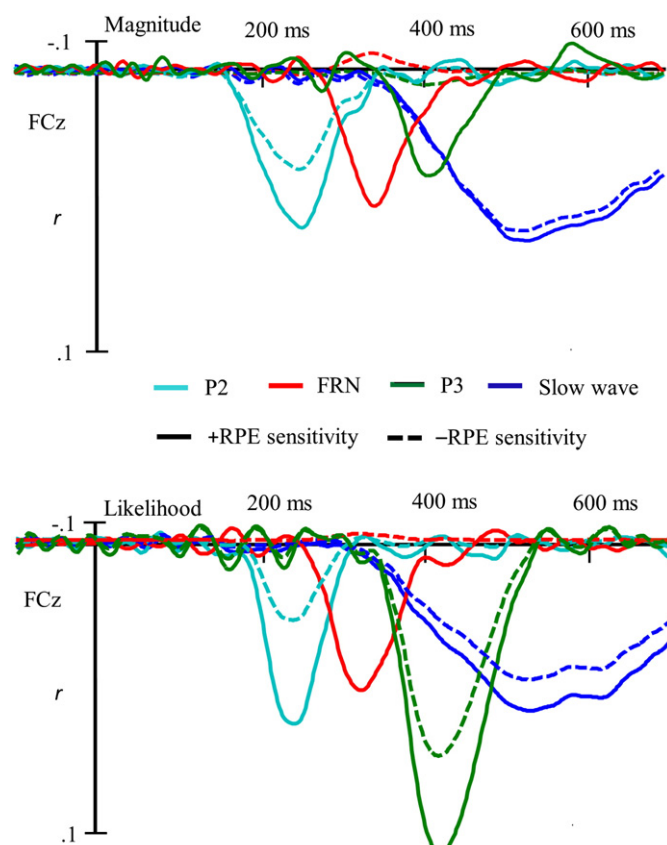


Fig. 6. Factors extracted from correlation waveforms using PCA. Factors have been denominated in units of r and thus reflect the proportion of r they account for in the original undecomposed data (see Fig. 5). Amplitude represents the sensitivity of a factor to RPE size over time, and calculated separately for +RPEs (undashed) and –RPEs (dashed). To allow comparison all factors are plotted at FCz. Factor details are provided in Table 1.

localisation in sLORETA is limited to cortical and hippocampal grey matter, it could not, in principle, reproduce a source in the basal ganglia: nevertheless, activity was clustered in the neighbouring parahippocampal gyrus.

Other PCA components

A late parietal factor with a profile consistent with a slow wave, and showing a salience response function was found under both modulators. It was responsive to both +RPEs (magnitude $t(44) = 11.21$, $p < .001$; likelihood $t(41) = 8.74$, $p < .001$) and –RPEs (magnitude $t(44) = 7.99$, $p < .001$; likelihood $t(41) = 7.49$, $p < .001$), showing a greater response to +RPEs relative to –RPEs under the likelihood but not magnitude modulator (magnitude, $t(44) < 1$; likelihood $t(41) = 12.06$, $p < .001$). A dipole for this factor was found at 31.69, –28.63, –5.32 (hippocampus) which also lay within the region of activation shown by sLORETA.

Table 1
Factor combinations selected for statistical analysis.

Proposed component	Modulator	Temporal loading peak	Spatial peak	Variance explained (%)	Response function	Factor
P2	Magnitude	256	Fz	6.89	Salience	TF2/SF1
	likelihood	240	FC2	5.24	Salience	TF4/SF1
FRN	Magnitude	344	Cz	4.56	+ RPE only	TF3/SF1
	likelihood	322	Cz	6.09	+ RPE only	TF3/SF1
P3	Magnitude	412	Cz	3.38	+ RPE only	TF4/SF1
	likelihood	418	Cz	16.99	Salience	TF2/SF1
Slow wave	Magnitude	532	POz	44.42	Salience	TF1/SF1
	likelihood	528	P1	30.96	Salience	TF1/SF1

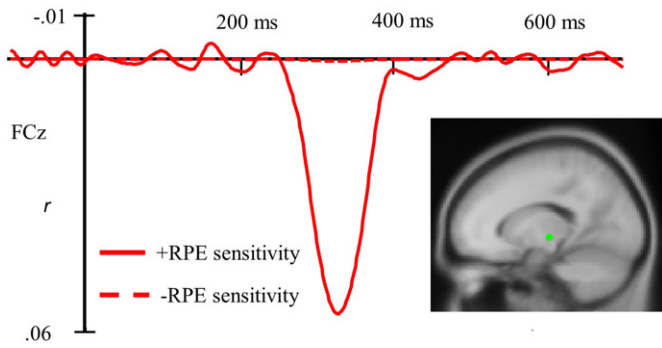


Fig. 7. Factor TF3/SF1, the proposed FRN factor, extracted from a PCA performed on a combined dataset after collapsing together the magnitude and likelihood modulators. The factor's waveform at FCz is shown and its estimated source in the putamen under dipole fitting.

A factor with a time course and amplitude consistent with the P3 was also found, though with a more anterior peak, at Cz. Under the magnitude modulator, the factor was responsive only to +RPEs (+RPEs, $t(44) = 7.66$, $p < .001$; -RPEs $t(44) < 1$). For likelihood, however, this component showed a salience response, having the same polarity of response to both +RPEs ($t(41) = 13.68$, $p < .001$) and -RPEs ($t(41) = 8.70$, $p < .001$), though with a greater response to +RPEs ($t(41) = 12.90$, $p < .001$). As such, its behaviour is not consistent with an RPE encoder. A dipole for this factor was found at 11.07, 0.30, -3.07, with a source in the subcallosal gyrus comprising the nearest grey matter, which also lay within the region of activation shown by sLORETA.

A frontal salience encoding factor was found which has been provisionally dubbed the P2 following Foti et al.'s (2011) classification of a similar factor shown under PCA. Under the magnitude modulator, this had the same polarity of response to both +RPEs ($t(44) = 7.97$, $p < .001$) and -RPEs ($t(44) = 5.53$, $p < .001$) but with greater responsiveness to +RPEs ($t(44) = 3.43$, $p = .001$). The same pattern was observed under the likelihood modulator (+RPEs, $t(41) = 6.40$, $p < .001$; -RPEs, $t(41) = 3.06$, $p = .004$; +RPEs vs. -RPEs, $t = 12.37$, $p < .001$). A dipole for this factor was found in the cerebellum at 29.86, -53.01, -10.35, while sLORETA showed activation in the adjacent grey matter of the fusiform gyrus.

Domain was not an experimental variable in the present study, however to ensure that the results were not particular to either gain or loss domains, additional, separate PCAs were performed on gain and loss trials for both magnitude and likelihood data. This had no effect on the pattern of statistical significance shown by the factors under one sample *t*-tests, suggesting the factors' response functions were unaffected by domain. The factors are shown in Supplementary Fig. 2.

The response function of all factors also remained the same when the learning rate was set to zero, (common in the FRN literature) with the exception that, specifically under likelihood modulation, the FRN component achieved a weak sensitivity to -RPEs in addition to a strong sensitivity to +RPEs (-RPE sensitivity, $t = 2.55$, $p = .013$; +RPE sensitivity, $t = 6.31$, $p < .001$). Factors are presented in Supplementary Table 1.

Theta and delta

Grand average ERPs of theta and delta filtered data are shown in Fig. 8, allowing the contribution of each of these frequencies to the FRN to be assessed. In the interval 240–340 ms, theta power was greater for -RPEs compared to +RPEs (magnitude, $t(44) = 2.99$, $p = .005$; likelihood $t(41) = 3.39$, $p = .002$), while delta power was greater for +RPEs compared to -RPEs (magnitude, $t(44) = 7.42$, $p < .001$; likelihood $t(41) = 6.30$, $p < .001$).

Correlation waveforms for the filtered data are shown in Fig. 9. For delta, these results closely resemble those found in the original broad

frequency data shown in Fig. 5. That is, there is an attenuation in this interval of sensitivity to -RPE size at 320–340 ms, and in the FRN interval overall. A power comparison in the interval 240–340 ms revealed sensitivity to RPE size to be greater for +RPEs than -RPEs (magnitude $t(44) = 4.73$, $p < .001$; likelihood $t(41) = 2.40$, $p = .02$), indicating an interaction between RPE size and valence. In the theta band, the relationship between RPE size and voltage was oscillatory, producing a negativity for large RPEs at 300–350 ms set between two positivities. However, the relationship of RPE size to voltage is the same for both +RPEs and -RPEs, i.e. they can be seen to be in phase oscillation in Fig. 8, with no significant differences in sensitivity to RPE size in the interval 240–340 ms (magnitude $t(44) = .39$, $p = .70$; likelihood $t(41) = 1.78$, $p = .08$). Thus, power within the theta frequencies is indicative of salience encoding.

The correlation waveforms for delta and theta were entered into separate PCAs to separate out overlapping components in those frequency bands. For delta activity, the temporal step of the PCA resulted in 5 temporal factors for both magnitude and likelihood modulated RPEs, each of which was parsed into three temporospatial factors at the spatial stage. Since, in each case, the first of these three temporospatial factors accounted for the bulk of the variance explained at the temporal step, only the first temporospatial factor of each temporal factor was considered. Of these five factors, four were the same factors extracted from the original broad frequency data, shown in Fig. 6. A fifth factor occurring at 90–100 ms showed a salience response under likelihood modulation and a +RPE response under magnitude. For theta activity, the temporal step of the PCA resulted in a single temporal factor extracted for magnitude modulated RPEs and two temporal factors for likelihood, with each of these producing three temporospatial factors at the spatial step. All theta factors showed a salience response. Factors are shown in Supplementary Fig. 3.

Discussion

Our aim in the present study was to disentangle the responses to signed and unsigned prediction errors that are likely to be present in the feedback locked ERP waveform and which have hindered attempts to demonstrate that the FRN axiomatically constitutes an RPE. To ensure that valence and salience were not confounded, we employed a design in which the size of signed and unsigned prediction errors was uncorrelated. We manipulated RPE size as a continuous independent variable to increase the power of our design to reveal a component sensitive to continuous variations in RPE size. We used PCA and time frequency decomposition to reduce the effect of component overlap.

An initial analysis of averaged ERPs revealed a typical FRN showing a relative negativity for -RPEs in the interval 240–340 ms. Separation into theta and delta frequency bands revealed greater power for -RPEs in theta, and greater power for +RPEs in delta, supporting results of a number of recent studies (Bernat et al., 2015; Cavanagh et al., 2010; Foti et al., 2014; Kamarajan et al., 2012; Marco-Pallares et al., 2008; Mas-Herrero and Marco-Pallares, 2014).

Such categorical valence effects, either in the original waveforms or their spectra, do not demonstrate RPE-encoding components however. This requires the demonstration that the component is also sensitive to RPE size, that this interacts with its sensitivity to valence, and that the interaction takes a specific form: either sensitivity to the size of +RPEs only, or -RPEs only, or to both but with the sign of the sensitivity reversed. In contrast, same-signed sensitivity to RPE size for +RPEs and -RPEs would constitute a salience response even if an interaction between RPE size and valence were present since, as noted earlier, such a component could not reliably distinguish good and bad outcomes despite its asymmetrical response to them.

As can be seen in Fig. 5, for the majority of the 700 ms following feedback, sensitivity to the size of RPEs was indeed present, but with the same sign for both +RPEs and -RPEs (i.e. both waveforms showed positive deflections) indicating a salience response. An attenuation of

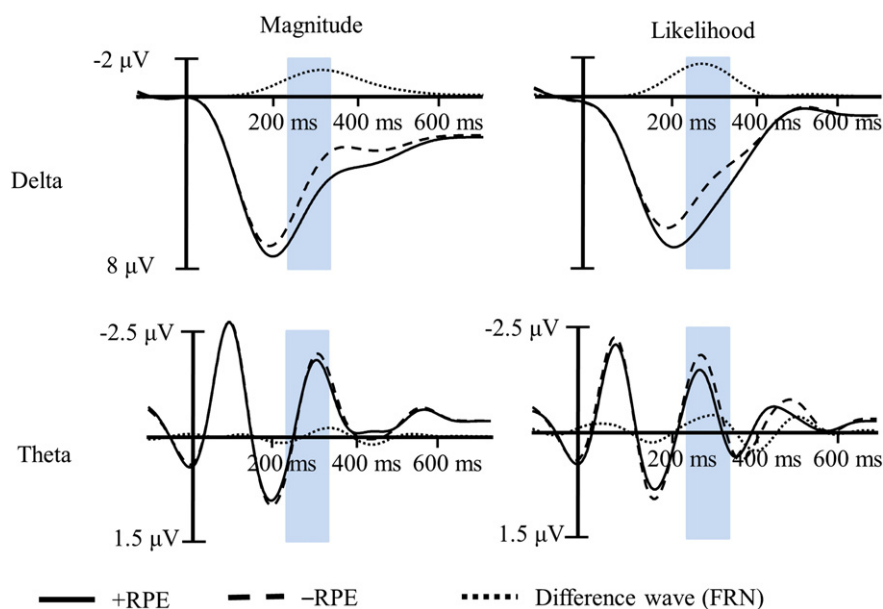


Fig. 8. Theta and delta activity in the interval of the FRN. ERPs show grand averages of single trials filtered by a bandwidth pass of 0.1–4 Hz (delta) or 4–7 Hz (theta); for further details see Fig. 4.

sensitivity to $-RPE$ size at 320–340 ms, however, suggests that at this point $+RPE$ encoding may occur. This effect occurred in the delta frequency, as demonstrated both by the similarity of the delta filtered waveform to the broad frequency parent, and the fact that a significant interaction of RPE size and valence was present in the delta filtered waveform but not the theta. This effect was also shown in a recent study by Cavanagh (2015), although the absence of a $-RPE$ condition in that study means that a salience response cannot be ruled out. In theta, as can be seen in Fig. 9, the oscillations of $+RPE$ and $-RPE$ waveforms are largely in phase, thus sensitivity to RPE size is same-signed at any given time, indicating a salience response. This is in accord with recent studies which have shown that the greater average theta power of $-RPE$ s compared to $+RPE$ s does not entail greater sensitivity to $-RPE$ size (Bernat et al., 2015; Mas-Herrero and Marco-Pallarés, 2014, though see Cavanagh et al., 2010).

While the effect at 320–340 ms was suggestive, PCA allowed the effect to be addressed more robustly. Since PCA is a purely data-driven process, it is resistant to biases in data interpretation, and particularly those resulting from the choice of interval or electrode for analysis (all data are considered equally meaningful at the PCA's outset). The PCA confirmed the existence of the $+RPE$ encoder at 320–340 ms. An additional PCA, performed separately on delta and theta filtered correlation waveforms confirmed that this component occupied the delta frequency. In contrast, only salience encoding components were extracted from the theta correlation waveforms.

A component with a temporospatial profile consistent with the FRN was thus extracted from the present data. According to Holroyd & Coles' original theory, the FRN codes for a temporal difference error carried by dopaminergic midbrain neurons to the anterior cingulate cortex (ACC), where it serves as a teaching signal. In the present study, this

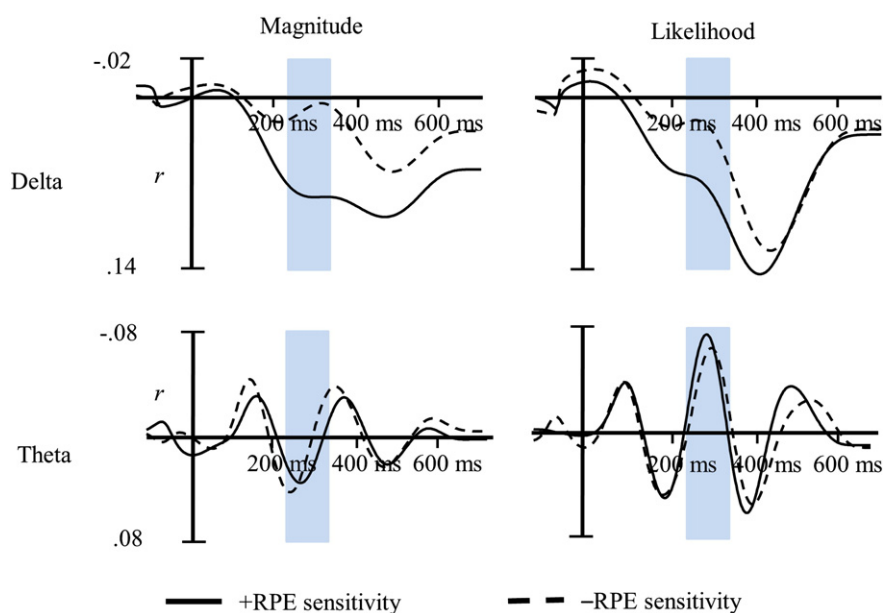


Fig. 9. Correlation waveforms showing the strength of relationship between voltage and RPE size for single trials filtered by a bandwidth pass of 0.1–4 Hz (delta) or 4–7 Hz (theta); for further details see Fig. 5.

component was responsive only to the size of +RPEs, suggesting that only a reward signal is transmitted. This finding is in keeping with the fact that midbrain dopaminergic neurons are strongly suspected to code +RPEs with scaled phasic bursts (Schultz et al., 1997), but because of their low tonic output appear to have little ability to code –RPEs with phasic reductions (Bayer and Glimcher, 2005). Other aspects of Holroyd and Coles' theory are less compatible with the present findings. In particular, the theory claims a close functional link between the FRN and the response-locked error related negativity (ERN), a component arising in response to internal registration of an error. Since the ERN necessarily arises exclusively from –RPEs, the plausibility of a common neural substrate with the FRN is strongly questioned.

The localisation of the FRN to the ACC is also not supported by the present data. While the ACC has been commonly implicated as the source of the FRN, this finding is typically based on ERPs rather than individual components extracted by PCA or a similar process. In contrast, recent source localisation attempts on PCA-extracted components have consistently found a source in the striatum (Carlson et al., 2011; Foti et al., 2011; Martin et al., 2009), as was found in our own study. This raises the possibility that source localisation attempts based upon data that has not been decomposed by PCA may well have been compromised by overlap with some other component. It is notable that the effect of valence in the undecomposed waveforms is strongest at ~300 ms (shown by the difference waves in Fig. 4) which is typical for the FRN (Sambrook and Goslin, 2015), but that this effect nevertheless does not correspond to any extracted factor. Decomposition of the effect reveals it to arise from a combination of TS3/SF1, the factor we have described as the FRN, and the earlier P2 component, which shows a salience response which is somewhat stronger for +RPEs. Clearly, source localisation of the valence effect at 300 ms could thus not give a meaningful result. In contrast, localisation of the proposed FRN component to the ventral striatum, combined with the finding that it codes exclusively for +RPEs, converges with recent fMRI meta-analyses showing the striatum to be strongly +RPE biased (Bartra et al., 2013; Garrison et al., 2013; Liu et al., 2011). In these meta-analyses, the areas of the striatum responding to +RPEs were also relatively large, increasing the plausibility of their producing a scalp detectable voltage, something which has previously been questioned (Cohen et al., 2011).

Our prime motive for using PCA in the present study was to remove components overlapping the FRN. An additional aim was to use this data-driven technique to reveal other components that appeared to encode RPEs. None were found. Instead, numerous salience encoding components were found which were sensitive to absolute, or unsigned prediction error size. While some of these showed a larger response for +RPEs than –RPEs, thus satisfying the axiomatic model's first axiom, they failed its second axiom because voltage became more positive for both +RPEs and –RPEs as RPE size grew. The importance of the second axiom should be clear when one considers that the activity of such components will be the same for large –RPEs and small +RPEs, rendering them unsuitable as RPE encoders. These salience components had latencies, durations and strengths strongly resembling Foti et al.'s (2011) PCA study, and also resembled the behaviour of ERP components shown in other recent studies and reviews (San Martin, 2012; Talmi et al., 2013).

Notably, the study failed to find any –RPE encoding component. This is a surprising finding, given the apparent value that coding the size of –RPEs has for learning. While theta power was greater overall for –RPEs compared to +RPEs, no component sensitive to –RPE size was found that could not better described as a salience encoder. It should be noted that our experiment was specifically designed to provide a fair assessment of the presence of the four response functions (+RPE, –RPE, integrated and salience) in the post feedback waveform. First, the problem of component overlap was directly addressed by PCA. Second, and also as a consequence of the PCA, the waveform's true activity was not distorted by averaging activity over a canonical interval

taken from the literature which might have historically reified a component that was in fact a composite. Third, an appropriate signal for each of the possible RPE encoders was derived from the axiomatic model.

Nevertheless, other studies which have met these criteria have in some cases revealed a –RPE encoder. Using a Pavlovian, rather than operant learning variant of the present task, we have observed sensitivity to –RPE size at ~360 ms albeit as a much fainter signal (Sambrook and Goslin, 2014). Pedroni et al. (2011) also found sensitivity to –RPEs at that time (albeit with reversed polarity) in a design that addressed component overlap with topographic ANCOVA. Donamayor et al. (2011) and Talmi et al. (2012) addressed the problem of component overlap with MEG, with Donamayor et al. finding a –RPE encoding 480–600 ms and no +RPE encoding, and Talmi et al. finding an integrated coding at ~320 ms. The literature is thus inconsistent. This is very likely due the difficulty in disaggregating reward encoding components even in designs that have explicitly set out to do so. Components are reified by the statistical effects of experiments, but given the quantity of neural activity that contributes to those effects it is unlikely that the components so described are not in reality conglomerates of functionally distinct components. This is particularly pernicious in the case of RPE encoders since there are many ways that two quite separate RPE encoders which are analysed as a single unit can produce an aggregate response that corresponds perfectly to some third RPE encoder that is not present at all (Sambrook and Goslin, 2014). The problem is reduced, but not eliminated, by decomposing the waveform into principal components or time frequency spectra and, as such, the present study, and others addressing this problem comprise one step along this road. As such, one emerging result from this literature, consolidated by the present study, is that the encoding of –RPEs is absent from the ERP waveform in the specific time interval associated with the FRN. This suggests that the robust valence effects that have been observed there since Holroyd and Coles' influential paper constitute a response to good, not bad outcomes.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2015.07.032>.

References

- Bartra, O., McGuire, J.T., Kable, J.W., 2013. The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage* 76, 412–427. <http://dx.doi.org/10.1016/j.neuroimage.2013.02.063>.
- Bayer, H.M., Glimcher, P.W., 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47 (1), 129–141. <http://dx.doi.org/10.1016/j.neuron.2005.05.020>.
- Bernat, E.M., Nelson, L.D., Baskin-Sommers, A.R., 2015. Time-frequency theta and delta measures index separable components of feedback processing in a gambling task. *Psychophysiology*.
- Caplin, A., Dean, M., 2008. Axiomatic methods, dopamine and reward prediction error. *Curr. Opin. Neurobiol.* 18 (2), 197–202. <http://dx.doi.org/10.1016/j.conb.2008.07.007>.
- Carlson, J.M., Foti, D., Mujica-Parodi, L.R., Harmon-Jones, E., Hajcak, G., 2011. Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study. *NeuroImage* 57 (4), 1608–1616. <http://dx.doi.org/10.1016/j.neuroimage.2011.05.037>.
- Cavanagh, J.F., 2015. Cortical delta activity reflects reward prediction error and related behavioral adjustments, but at different times. *NeuroImage* 110, 205–216. <http://dx.doi.org/10.1016/j.neuroimage.2015.02.007>.
- Cavanagh, J.F., Frank, M.J., Klein, T.J., Allen, J.J.B., 2010. Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *NeuroImage* 49 (4), 3198–3209. <http://dx.doi.org/10.1016/j.neuroimage.2009.11.080>.
- Cohen, J., 1983. The cost of dichotomization. *Appl. Psychol. Meas.* 7, 249–253.
- Cohen, M.X., Cavanagh, J.F., Slagter, H.A., 2011. Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: temporospatial principal components analysis and source localization of the feedback negativity: commentary. *Hum. Brain Mapp.* 32 (12), 2270–2271. <http://dx.doi.org/10.1002/Hbm.21358>.
- Daw, N.D., Kakade, S., Dayan, P., 2002. Opponent interactions between serotonin and dopamine. *Neural Netw.* 15 (4–6), 603–616.
- Dien, J., 2010a. The ERP PCA toolkit: an open source program for advanced statistical analysis of event-related potential data. *J. Neurosci. Methods* 187 (1), 138–145. <http://dx.doi.org/10.1016/j.jneumeth.2009.12.009>.
- Dien, J., 2010b. Evaluating two-step PCA of ERP data with geomin, infomax, oblimin, promax, and varimax rotations. *Psychophysiology* 47 (1), 170–183. <http://dx.doi.org/10.1111/j.1469-8986.2009.00885.x>.
- Dien, J., Spencer, K.M., Donchin, E., 2003. Localization of the event-related potential novelty response as defined by principal components analysis. *Cogn. Brain Res.* 17 (3), 637–650. [http://dx.doi.org/10.1016/S0926-6410\(03\)00188-5](http://dx.doi.org/10.1016/S0926-6410(03)00188-5).

- Dien, J., Beal, D.J., Berg, P., 2005. Optimizing principal components analysis of event-related potentials: matrix type, factor loading weighting, extraction, and rotations. *Clin. Neurophysiol.* 116 (8), 1808–1825. <http://dx.doi.org/10.1016/j.clinph.2004.11.025>.
- Dien, J., Khoe, W., Mangun, G.R., 2007. Evaluation of PCA and ICA of simulated ERPs: promax vs. infomax rotations. *Hum. Brain Mapp.* 28 (8), 742–763. <http://dx.doi.org/10.1002/Hbm.20304>.
- Donamayor, N., Marco-Pallares, J., Heldmann, M., Schoenfeld, M.A., Munte, T.F., 2011. Temporal dynamics of reward processing revealed by magnetoencephalography. *Hum. Brain Mapp.* 32 (12), 2228–2240. <http://dx.doi.org/10.1002/Hbm.21184>.
- Foti, D., Weinberg, A., Dien, J., Hajcak, G., 2011. Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: temporospatial principal components analysis and source localization of the feedback negativity. *Hum. Brain Mapp.* 32 (12), 2207–2216. <http://dx.doi.org/10.1002/Hbm.21182>.
- Foti, D., Weinberg, A., Bernat, E.M., Proudfit, G.H., 2014. Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. *Clin. Neurophysiol.* 126, 1338–1347.
- Garrison, J., Erdeniz, B., Done, J., 2013. Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neurosci. Biobehav. Rev.* 37 (7), 1297–1310. <http://dx.doi.org/10.1016/j.neubiorev.2013.03.023>.
- Gehring, W.J., Willoughby, A.R., 2002. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295 (5563), 2279–2282.
- Hajihosseini, A., Holroyd, C.B., 2013. Frontal midline theta and N200 amplitude reflect complementary information about expectancy and outcome evaluation. *Psychophysiology* 50 (6), 550–562. <http://dx.doi.org/10.1111/Psyp.12040>.
- 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.
- Holroyd, C.B., Coles, M.G.H., 2002. The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109 (4), 679–709. <http://dx.doi.org/10.1037/0033-295x.109.4.679>.
- 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 (5), 688–697. <http://dx.doi.org/10.1111/j.1469-8986.2008.00668.x>.
- Horn, J.L., 1965. A rationale and test for the number of factors in factor analysis. *Psychometrika* 30 (2), 179–185.
- Kamarajan, C., Rangaswamy, M., Manz, N., Chorlian, D.B., Pandey, A.K., Roopesh, B.N., Porjesz, B., 2012. Topography, power, and current source density of theta oscillations during reward processing as markers for alcohol dependence. *Hum. Brain Mapp.* 33 (5), 1019–1039. <http://dx.doi.org/10.1002/Hbm.21267>.
- Kobayashi, S., Nomoto, K., Watanabe, M., Hikosaka, O., Schultz, W., Sakagami, M., 2006. Influences of rewarding and aversive outcomes on activity in macaque lateral prefrontal cortex. *Neuron* 51 (6), 861–870. <http://dx.doi.org/10.1016/j.neuron.2006.08.031>.
- 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. <http://dx.doi.org/10.1111/j.1469-8986.2011.01291.x>.
- Kujawa, A., Smith, E., Luhmann, C., Hajcak, G., 2013. The feedback negativity reflects favorable compared to nonfavorable outcomes based on global, not local, alternatives. *Psychophysiology* 50 (2), 134–138. <http://dx.doi.org/10.1111/Psyp.12002>.
- Liu, X., Hairston, J., Schrier, M., Fan, J., 2011. Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neurosci. Biobehav. Rev.* 35 (5), 1219–1236. <http://dx.doi.org/10.1016/j.neubiorev.2010.12.012>.
- Marco-Pallares, J., Cucurell, D., Cunillera, T., Garcia, R., Andres-Pueyo, A., Munte, T.F., Rodriguez-Fornells, A., 2008. Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia* 46 (1), 241–248. <http://dx.doi.org/10.1016/j.neuropsychologia.2007.07.016>.
- Martin, L.E., Potts, G.F., Burton, P.C., Montague, P.R., 2009. Electrophysiological and hemodynamic responses to reward prediction violation. *Neuroreport* 20 (13), 1140–1143. <http://dx.doi.org/10.1097/Wnr.0b013e32832f0dca>.
- Mas-Herrero, E., Marco-Pallares, J., 2014. Frontal theta oscillatory activity is a common mechanism for the computation of unexpected outcomes and learning rate. *J. Cogn. Neurosci.* 26 (3), 447–458.
- Metereau, E., Dreher, J.C., 2013. Cerebral correlates of salient prediction error for different rewards and punishments. *Cereb. Cortex* 23 (2), 477–487. <http://dx.doi.org/10.1093/cercor/bhs037>.
- Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find. Exp. Clin. Pharmacol.* 24 (Suppl. D), 5–12.
- Pedroni, A., Langer, N., Koenig, T., Allemand, M., Jancke, L., 2011. Electroencephalographic topography measures of experienced utility. *J. Neurosci.* 31 (29), 10474–10480. <http://dx.doi.org/10.1523/Jneurosci.5488-10.2011>.
- Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current-density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72 (2), 184–187. [http://dx.doi.org/10.1016/0013-4694\(89\)90180-6](http://dx.doi.org/10.1016/0013-4694(89)90180-6).
- Rescorla, R.A., Wagner, A.R., 1972. A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. *Classical Conditioning II: Current Research and Theory*. 2 pp. 64–99.
- Sambrook, T.D., Goslin, J., 2014. Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors. *Neuropsychologia* 61, 1–10. <http://dx.doi.org/10.1016/j.neuropsychologia.2014.06.004>.
- Sambrook, T.D., Goslin, J., 2015. A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychol. Bull.* 141 (1), 213–235. <http://dx.doi.org/10.1037/Bul0000006>.
- Sambrook, T.D., Roser, M., Goslin, J., 2012. Prospect theory does not describe the feedback-related negativity value function. *Psychophysiology* 49 (12), 1533–1544.
- San Martin, R., 2012. Event-related potential studies of outcome processing and feedback-guided learning. *Front. Hum. Neurosci.* 6 (doi: Artn 304).
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science* 275 (5306), 1593–1599. <http://dx.doi.org/10.1126/science.275.5306.1593>.
- Seymour, B., O'Doherty, J.P., Koltzenburg, M., Wiech, K., Frackowiak, R., Friston, K.J., Dolan, R., 2005. Opponent appetitive–aversive neural processes underlie predictive learning of pain relief. *Nat. Neurosci.* 8 (9), 1234–1240. <http://dx.doi.org/10.1038/Nn1527>.
- Talmi, D., Fuentesmilla, 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. <http://dx.doi.org/10.1016/j.neuroimage.2011.06.051>.
- Talmi, D., Atkinson, R., El-Deredy, W., 2013. The feedback-related negativity signals salience prediction errors, not reward prediction errors. *J. Neurosci.* 33 (19), 8264–8269. <http://dx.doi.org/10.1523/Jneurosci.5695-12.2013>.
- Walsh, M.M., Anderson, J.R., 2012. Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neurosci. Biobehav. Rev.* 36 (8), 1870–1884. <http://dx.doi.org/10.1016/j.neubiorev.2012.05.008>.
- Yu, R., Zhang, P., 2014. Neural evidence for description dependent reward processing in the framing effect. *Front. Neurosci.* 8, 56. <http://dx.doi.org/10.3389/fnins.2014.00056>.
- Zheng, Y., Li, Q., Wang, K., Wu, H.Y., Liu, X., 2015. Contextual valence modulates the neural dynamics of risk processing. *Psychophysiology* 52, 895–904.