



PEARL

**Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors.**

Sambrook, Thomas D.; Goslin, Jeremy

**Published in:**  
Neuropsychologia

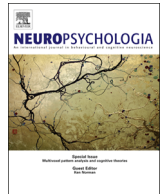
**DOI:**  
[10.1016/j.neuropsychologia.2014.06.004](https://doi.org/10.1016/j.neuropsychologia.2014.06.004)

**Publication date:**  
2014

**Link:**  
[Link to publication in PEARL](#)

**Citation for published version (APA):**  
Sambrook, T. D., & Goslin, J. (2014). Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors. *Neuropsychologia*, *61*(0), 1-10.  
<https://doi.org/10.1016/j.neuropsychologia.2014.06.004>

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Wherever possible please cite 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.



# Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors



Thomas D. Sambrook\*, Jeremy Goslin

School of Psychology, Cognition Institute, Plymouth University, Plymouth PL4 8AA, United Kingdom

## ARTICLE INFO

### Article history:

Received 18 January 2014

Received in revised form

5 June 2014

Accepted 6 June 2014

Available online 16 June 2014

### Keywords:

Feedback related negativity (FRN)

Event-related potential (ERP)

Reward prediction error (RPE)

Unsigned prediction error

Saliency

Dopamine

## ABSTRACT

Reinforcement learning models make use of reward prediction errors (RPEs), the difference between an expected and obtained reward. There is evidence that the brain computes RPEs, but an outstanding question is whether positive RPEs (“better than expected”) and negative RPEs (“worse than expected”) are represented in a single integrated system. An electrophysiological component, feedback related negativity, has been claimed to encode an RPE but its relative sensitivity to the utility of positive and negative RPEs remains unclear. This study explored the question by varying the utility of positive and negative RPEs in a design that controlled for other closely related properties of feedback and could distinguish utility from salience. It revealed a mediofrontal sensitivity to utility, for positive RPEs at 275–310 ms and for negative RPEs at 310–390 ms. These effects were preceded and succeeded by a response consistent with an unsigned prediction error, or “salience” coding.

© 2014 Elsevier Ltd. All rights reserved.

## 1. Introduction

The function of a brain is to monitor its owner's environment, responding to events to increase the chance of survival and reproductive success. Brains receive a stream of sensory information of near infinite detail and the brain is a costly organ to run. Given that most environments show a degree of stability, an efficient neural response to this stream of information is to form expectations based on reliable environmental cues and to respond only to deviations from those. Such deviations are known as prediction errors. These have been argued to provide a common basis for computation in perceptual, attentional, cognitive, and motivational processes (den Ouden, Kok, & de Lange, 2012). In particular, deviations of reward from an expected quantity, that is *reward prediction errors* (RPEs), have been shown by formal models (Sutton & Barto, 1998) to be important terms in reinforcement learning, and there is strong evidence that RPEs are coded in the primate midbrain (Schultz, 2010).

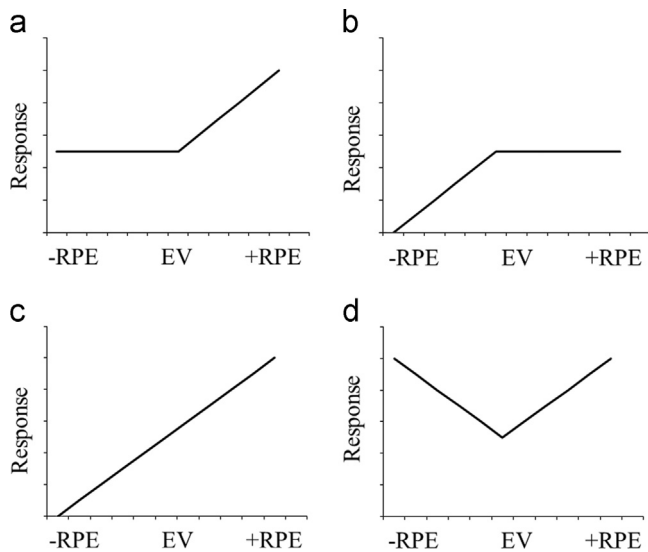
An event related potential (ERP) component known as feedback related negativity (FRN) occurring at mediofrontal sites at 200–350 ms has been proposed to encode an RPE carried from the midbrain to the anterior cingulate cortex (San Martin, 2012; Walsh & Anderson, 2012; Yu & Zhang, 2014). The FRN is so named because it exhibits a relative negativity for worse than expected

outcomes. However, this does not necessarily mean exclusive, or indeed any, sensitivity to negative reward prediction errors (–RPEs). The methodology in which the FRN emerged was based simply on comparing –RPE outcomes to +RPE outcomes; as such the negativity observed is merely relative and might equally have its basis in a positive voltage shift for +RPE outcomes. In fact, competing claims have been made in this regard, with some arguing that the FRN is preferentially sensitive to +RPEs (Cohen, Elger, & Ranganath, 2007; Eppinger, Mock, & Kray, 2009; San Martin, Manes, Hurtado, Isla, & Ibanez, 2010) and others arguing greater sensitivity to –RPEs (Bellebaum & Daum, 2008; Bellebaum, Polezzi, & Daum, 2010; Pfabigan, Alexopoulos, Bauer, & Sailer, 2011). In Fig. 1a and b we schematically represent these two possible response functions. Fig. 1c shows the response function of a component that codes both +RPEs and –RPEs (“integrated coding”), and Fig. 1d a response function to unsigned prediction errors (UPEs), that is to the absolute size of the prediction error irrespective of its valence. While this last response function has been plotted against RPE utility like the others, this merely represents how it would *behave* in an experiment studying RPEs, the component is not coding RPE utility at all but the quite different properties of UPE size. Such a component might serve a general function of registering motivational salience (Bromberg-Martin, Matsumoto, & Hikosaka, 2010).

A challenge for FRN research is that the post-feedback wave-form may comprise a number of different components with different response functions that at least partially overlap. One danger is that this overlap, rather than merely obscuring the

\* Corresponding author. Tel.: +44 1752 584826; fax: +44 1752 584808.

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



**Fig. 1.** A schematic representation of possible response functions to RPE utility. “Response” is generic and could refer to increase in single cell firing rate, BOLD activation, or amplitude increase for either a positive or negative going ERP component. In the specific case of the FRN the y axis corresponds to positivity of voltage. (a) coding of +RPE utility only, (b) coding of –RPE utility only, (c) “integrated” coding of the utility of all RPEs, and (d) coding of UPE size, or “salience”. EV = expected value, i.e. an RPE utility of zero.

individual components, has the capacity to synthesise entirely artefactual response functions. For example if a UPE size response function (Fig. 1d) overlies an integrated RPE response function (Fig. 1c) the sum effect in the EEG will be a spurious response function corresponding to Fig. 1a, a +RPE encoder. Where temporal and spatial overlap of such components is perfect, this problem is insoluble for the ERP methodology. However, even when two components might in principle be temporally dissociable, in practice this distinction may fail to be made since components are generally quantified over a relatively wide interval of 100 ms or so, a problem we address in the present experiment.

The example of a spurious +RPE encoder given above is topical. The recent trend in FRN research, following Holroyd, Pakzad-Vaezi, and Krigolson (2008), has been to claim that the FRN is solely responsive to +RPEs, despite the clear loss of adaptive value this would hold relative to an integrated RPE coder. Tellingly, there has also been recent growth in the number of papers claiming that the FRN does not code RPE utility at all and that is in fact just a UPE size response (Hauser et al., 2014; Oliveira, McDonald, & Goodman, 2007; Talmi, Atkinson, & El-Derey, 2013). If the post-feedback waveform comprises an early UPE size response closely followed by an integrated RPE utility response then an interval of measurement that catches part of both these responses will generate the spurious +RPE encoding response function as described.

The approach we took in the present study was therefore to quantify components according to the activity present in actual data using a “bottom-up” strategy, rather than imposing a possibly misguided interval of measurement (on which there is little agreement in the literature). We also depart from the tendency in FRN studies to represent RPEs as categorical levels such as “large good outcomes” vs. “small good outcomes” so that they can fit into a factorial design. This limitation belies the essentially continuous nature of RPEs. In our experiment we manipulated RPEs as a continuous independent variable, analysing the effect of RPE utility on voltage by correlating the two variables. Where the correlation between these two values was found to be significant, it could be assumed that voltage was influenced by RPE utility,

thus indicating the presence of an RPE encoding component. This relatively novel technique has been used by Hauk, Davis, Ford, Pulvermuller, and Marslen-Wilson (2006). The appropriateness of correlation coefficients for answering our research question can be readily appreciated by a glance in Fig. 1. The response functions there are schematic; however with the addition of a scatterplot of data points, each subplot might as easily represent two lines of best fit (one for +RPEs and one for –RPEs) set end to end, with the closeness of scatter on each line corresponding to two correlation coefficients for the responsiveness of voltage to +RPEs and –RPEs respectively.

Practically, identification of RPE encoding components in the post-feedback waveform, and elucidation of their response functions were achieved in two stages, as follows. Separate correlations of voltage with +RPE utility and –RPE utility were calculated at each time point on the post-feedback waveform. Separating the two kinds of RPE at the correlation stage was essential: if the correlation were calculated over the full range of RPEs then a significant positive value (for example) of Pearson’s  $r$  would leave us uninformed as to whether we were observing a response function corresponding to Fig. 1a, b or c. Having generated running correlation coefficients over the post-feedback waveform, these were clustered into discrete intervals which were tested for significance using the procedure of Maris and Oostenveld (2007). This showed us the intervals where –RPE and +RPE encoding was occurring, but not the overall response functions in these intervals. This was achieved in the second stage by considering +RPE and –RPE encoding together and comparing the joint activity to the response functions in Fig. 1. For example an interval of the waveform where there was a strong correlation between +RPE utility and voltage could be described as a +RPE encoder only if there was no correlation between –RPE utility and voltage in that interval. If there was in fact a same-signed correlation between –RPE utility and voltage then this pattern of results suggested an integrated encoder. If there was an oppositely-signed correlation between –RPE utility and voltage in this interval then this suggested a UPE size encoding.

These novel measures were designed to deal with the likely presence of multiple components in the post-feedback waveform, and maximise the possibility that they might be separated out to their true response functions. An essential requirement of this endeavour was that the experimental design should avoid confounds seen in the existing literature that may have served to distort the apparent sensitivity of the FRN to +RPEs and –RPEs, thus masking the FRN’s true response function. One of these is the *domain of the outcome*, that is, whether the outcome constitutes an absolute loss or gain. There is substantial evidence from behavioural economics (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) that outcome domain is highly salient for humans. In contrast, RPEs merely describe the difference between an outcome’s utility and its prior expected utility. The valence of the RPE (+/–) is thus formally orthogonal to whether the outcome constitutes an absolute gain or loss, that is whether the outcome is a value greater or less than zero. A –RPE of  $-5\text{c}$  might, for example, arise from making a *gain* of  $20\text{c}$  when a gain of  $25\text{c}$  was the expectation for that trial. Experiments on the FRN which use “mixed gambles”, in which –RPEs are always losses and +RPEs are always gains confound outcome domain and RPE valence. For example, such a study might vary the utility of –RPEs and +RPEs by manipulating the magnitude of outcomes as follows:  $+25\text{c}$ ,  $+5\text{c}$ ,  $-5\text{c}$ ,  $-25\text{c}$  (in order of descending RPE utility). Alternatively, the study might also achieve the same end by giving rewards and punishments with different prior likelihoods as follows:  $+25\text{c}$  when reward was unlikely,  $+25\text{c}$  when reward was likely,  $-25\text{c}$  when punishment was likely,  $-25\text{c}$  when punishment was unlikely. In either of these two experiments, an

apparent differential sensitivity to the utility of +RPEs and –RPEs might simply arise from differential sensitivity to either gain or loss outcomes generally. Indeed, in directly addressing sensitivity of the FRN in the two domains, Kreussel et al. (2012), Kujawa, Smith, Luhmann, and Hajcak (2013), Mushtaq, Stoet, Bland, and Schaefer (2013), Sambrook, Roser, and Goslin (2012) and Yu and Zhang (2014) have all shown a relative insensitivity to the utility of RPEs expressed in the loss domain, with Holroyd, Larsen, and Cohen (2004) finding non-significant results in the same direction. This calls into question the validity of claims that the FRN is insensitive to the utility of –RPEs given that much of this evidence is based on mixed gambles. In the present experiment, expectations were manipulated in such a way that –RPEs and +RPEs were both expressed as losses and gains half the time, recovering the orthogonal nature of the relationship of outcome domain and RPE valence.

A second important confound concerns whether outcomes are deliveries or omissions. When FRN experiments do not use mixed gambles they near-ubiquitously offer gambles in the gain domain only and modulate RPE utility by whether a reward has been obtained or not. Thus a likelihood-modulated FRN experiment might offer outcomes as follows: 25¢ when reward was unlikely, 25¢ when reward was likely, nothing when non-reward was likely, nothing when non-reward was unlikely; alternatively, a magnitude modulated FRN experiment would offer 25¢, 5¢, non-reward (where the alternative was 5¢), non-reward (where the alternative was 25¢). However, it has been argued that non-rewards are less salient than rewards (Esber & Haselgrove, 2011). If this is so then an apparent lack of sensitivity to the utility of –RPEs may occur simply because they have been expressed as non-rewards. In the present experiment there were no non-rewards: all outcomes constituted deliveries of some numerical quantity, with –RPEs being worse than expected quantities.

A third possible source of confounds surrounds the use of reward likelihood (rather than reward magnitude) to manipulate RPE utility. Although this has historically been the preferred means of manipulating RPE utility in FRN experiments, unexpected events are known to have very strong effects in the time course of the FRN (Folstein & Van Petten, 2008) and while unexpectedness may play a formal role in dictating the utility of RPEs, it is likely to bring with it substantial non-specific, alerting responses. For this reason, the present experiment manipulated RPEs using outcome magnitude rather than likelihood. Using reward magnitudes of equal frequency preserves the formal manipulation of RPE utility, but removes non-specific surprise effects.

A related, though more insidious confound, may exist in the form of perceptual mismatch between an expected, or hoped for stimulus and the actual feedback. Jia et al. (2007) have shown that a negativity is elicited in the FRN interval when a stimulus differs from one predicted by a participant, regardless of whether match or mismatch is the winning criterion, while Donkers, Nieuwenhuis, and van Boxtel (2005) showed that when a stimulus breaks a pattern with an ongoing sequence this leads to a net negativity even when this mismatch denotes a positive outcome. Based on informal questioning of participants in previous experiments we find many claim to hold an internal representation of the “winning stimulus” just prior to feedback. If this is typical, then a negative peak attributed to the FRN may in fact partly reflect a perceptual mismatch phenomenon. Because prediction errors were a continuous variable in the present experiment we expected this to undermine any habitual representation of a discrete winning stimulus and so prevent the opportunity for a mismatch component to be introduced for –RPEs specifically.

Our aim in this experiment was to isolate neural activity associated with RPEs. Yeung, Holroyd, and Cohen (2005) have

observed that reinforcement learning experiments, in which participants believe they are performing a feedback-guided learnable task, typically confound RPEs with error signals, that is, with non-economic judgements of whether an error was committed. To avoid this confound, participants in the present experiment were explicitly told that outcomes were unrelated to their key presses. Yeung et al. showed that while the FRN is reduced under such conditions it is still present, and a number of other studies have been able to demonstrate an FRN in conditions where participants were aware that they could exert no control over an outcome (Donkers et al., 2005; Donkers & van Boxtel, 2005; Holroyd, Krigolson, & Lee, 2011; Marco-Pallares, Kramer, Strehl, Schroder, & Munte, 2010; Potts, Martin, Burton, & Montague, 2006). We nevertheless used a task that superficially resembled a reinforcement learning task, in which participants selected from a choice of icons, to facilitate comparison with the literature.

In summary, our motivation in the present experiment was an attempt to separate components coding for + or – RPEs by using a data-driven methodology that removed the need for fixed-interval quantification, and the incumbent risk of combining components within a single analysis window. The rationale for the design was that it should avoid known confounds and that it should be able to describe the relationship between RPE utility and voltage independently for +RPEs and –RPEs and map these relationships onto the canonical response function templates shown in Fig. 1: describing the data as correlation coefficients served this purpose well. Based on the extensive FRN literature, we hypothesised that RPE utility encoding would be present (i.e. not just a UPE size response), but we had no prior hypothesis regarding which of the forms 1a–c this would take, nor when it would occur.

## 2. Methods

### 2.1. Participants

Sixty two undergraduates (9 left-handed, 22 male) participated for course credit and the opportunity to win a small sum of money. Data from 7 participants were rejected (five for equipment failure, two for eye blink artefacts on over 50% of trials).

### 2.2. Task rationale

Prediction errors were manipulated using reward magnitude. Participants undertook gambles in separate gain and loss domain blocks, hoping in gain domain blocks to win money, and in loss domain blocks to avoid loss of money. 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.

On each trial, participants received feedback on the outcome in the form of a number in a 60 point range (23–82) denoting the points won (gain domain trials) or lost (loss domain trials). In each domain there were thus a range of 30 +RPE and 30 –RPE outcomes, independent of whether points were actually won or lost. The principal variable of interest was RPE utility, that is, a signed value corresponding to the difference between the actual points gained or lost on each trial and the average and therefore expected value (52.5 points). Participants were not explicitly informed that 52.5 was the expected value, in order to avoid the possibility that they might impose a categorical good vs. bad discrimination at this point. It was anticipated that exposure to the outcomes would fairly quickly allow the extraction of expected value by any neural component devoted to this process, with the uniform distribution of the 60 outcomes in each block helping in this regard.

### 2.3. Task procedure

The experimental task was presented using E-Prime software and is summarised in Fig. 2. Participants were shown a graphic depicting four symbols and selected one using a keypad. A fixation cross appeared (600–700 ms duration), followed by the points won or lost on that trial (700 ms duration) and then a blank screen (800 ms duration). Participants performed the task in 32 blocks of 60 trials each. In every block each of the numbers 23–82 appeared once in a random order.



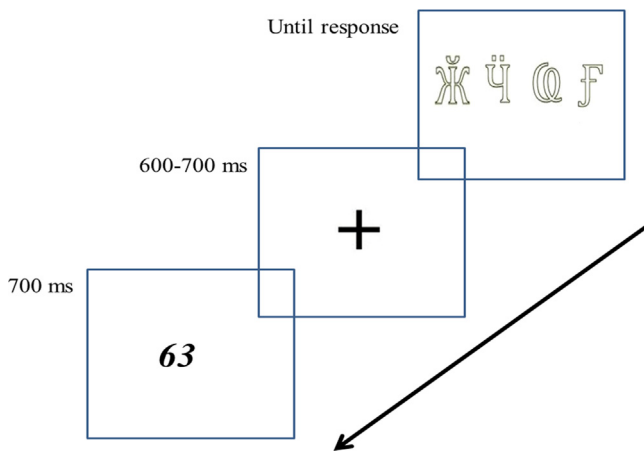


Fig. 2. Summary of a single trial of the task procedure.

For half the participants the first 16 blocks were gain domain blocks, and the second 16 were loss domain blocks, for the other half this pattern was reversed. Participants were verbally informed at the change point. At the end of each block participants were shown a number that was ostensibly, though was not in fact, the sum of all the numbers that had been displayed in that block. If this fictional number was greater than 3000, participants either won £2 (gain domain block), or lost £2 (loss domain block); if the number was below 3000 there was no effect. In fact the value of this fictional number was predetermined to be over 3000 half the time in each domain for a net gain of £.00 across the experiment for this aspect of the task. Participants were told (truthfully) that their key-presses were unrelated to the outcome of the trial, and were told (untruthfully) that each trial was determined by a random number generator set to return an average points value of 3000 across each block (the value was in fact fixed at 3150, the sum of the values 23–82). It was explained to participants (again untruthfully) that each run of the random number generator was independent, and so it was unknown how much money they might make or lose on this aspect of the task: this was to encourage interest in the feedback.

In order to motivate participants to pay attention, 10% of trials (192 trials) were followed by a probe in which participants had 1 s to make a key press to indicate what the number they had just seen was vs. a number one point higher or lower. Participants lost £.01 with each failure. Since this could be done simply by stimulus matching against the just departed stimulus, i.e. without processing magnitude, a further probe was included. At the end of each block, before being told the points total for that block, participants estimated whether the points total had exceeded or fallen short of 3000 and were awarded £.20 each time this matched the reported sum. Since the reported sum was, unbeknownst to the participants, fictitious and randomly determined, participants could only perform at chance on this question (50% average success rate): it was included merely to motivate attention to the points awarded on each trial of the block that preceded the question. Participants played 1920 trials in total, with a 30 s break between each of the 32 blocks. Total earnings for the experiment averaged £2.91 per person (approximately \$4.50).

At the conclusion of the experiment a check on participant preferences for outcomes was made. Participants were asked to rate how happy they were to see the outcomes 28, 38, 48, 58, 68 and 78 in each of the two domains by using a mouse to place a mark on a continuous rating scale labelled at one end as “very unhappy” and at the other end “very happy”. This process was repeated two further times and an average was taken.

#### 2.4. EEG recording

EEG data were collected from 11 actively amplified Ag/AgCl electrodes (actiCAP, Brain Products, Gilching, Germany) mounted on an elastic cap. The electrodes were Fz, FCz, Cz, CPz, Pz, F3, F4, FC3, FC4, FP1, and FP2. 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 electrodes FP1 and FP2 and 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 $\Omega$ . EEGs were amplified using a BrainAmp amplifier (Brain Products), continuously sampled at 500 Hz, and filtered offline with a band-pass filter from 2 to 30 Hz designed to remove P3 effects. ERPs were computed by averaging artifact-free EEGs (~86% = 1650 trials). EEGs were rejected if eye movement electrodes showed a voltage change exceeding 75  $\mu$ V/200 ms or if any midline site showed either a voltage change exceeding 20  $\mu$ V/ms or exceeded a value of  $\pm$ 100  $\mu$ V relative to baseline. EEGs were time-locked to 200 ms before the onset of the feedback to 700 ms afterward, and then were baseline-corrected using the interval –100 to 0 ms.

#### 2.5. EEG analysis

While feedback consisted of a single number, this could correspond to quite different quantities depending on the economic terms by which it was evaluated. These are laid out in Table 1. In order to analyse RPE utility effects, stimulus values were recorded to the RPE utilities they represented as indicated in the table. This resulted in a range of 60 utility values running from –29.5 to +29.5 that was independent of whether the outcome constituted an actual loss or gain. Initial factorial analyses of the FRN were performed by creating an average waveform for bad outcomes from all –RPEs and good outcomes from +RPEs. The bulk of the analysis was performed using correlations however, and these were performed separately for –RPEs and +RPEs and individually for each participant. A Pearson correlation coefficient between voltage and RPE utility was calculated at each time point. Data points in this correlation corresponded to individual trials. Since each participant saw an average of 1650 trials, of which half were +RPEs and half were –RPEs, correlations were obtained from an average of 825 sample points. While the RPE utility was fixed for a given trial, because the voltage varied over time, the correlation coefficient on each trial therefore also varied by time. The correlation coefficients were plotted against time to produce a figure which was analogous to conventional ERP plots but which showed the *strength of RPE encoding*, derived from the full range of utilities experienced, rather than *actual voltage* for a given bracket of RPE utility (e.g. high) as would be used in a factorial design. The interpretation of such a figure is straightforward: where the waveform is at baseline there is no effect of the variable (RPE utility) on voltage, where there are deviations from baseline this indicates a relationship, suggesting that RPEs are being coded by voltage. In this respect, such a correlational waveform can be interpreted as though it was in fact a traditional difference wave. Points on this waveform showing significant deviations could then be found by conducting a one sample *t*-test on the values of *r* (relative to an expected value of 0 under the null hypothesis) at each time point over the 55 participants. To facilitate subsequent Monte Carlo simulations the one sample *t*-test was implemented as a paired samples *t*-test comparing a column of 55 observed values of *r* against a column of 55 expected values set to zero.

The multiple comparisons resulting from the analysis of the whole waveform were addressed using a method based on the widely used cluster randomisation procedure of Maris and Oostenveld (2007). This procedure allows analysis of the entire ERP waveform without incurring the excess conservatism of a strict Bonferroni correction for each time point analysed. It achieves this by recognising that because voltages are strongly correlated at adjacent time points the effective number of comparisons being made is much lower than the number of sample points in the waveform.

In the first step of the procedure a one sample *t*-test on the values of *r* was performed at each time point and in each electrode channel in the manner described above, and used to identify significant ( $p < .05$ ) *t*-values. Because of the gradual growth and decay of the correlation coefficients both over time and space, these significant *t*-values also occurred in clusters of time points and electrode sites. Clusters were identified by finding significant *t*-values that were contiguous in time or space (adjacent time points for the same electrode or electrodes within 4 mm of each other and at the same time point). Only clusters containing eight or more samples (i.e. 16 ms) were considered for analysis. For each such cluster, a cluster-level *t*-value was calculated as the sum of all single sample *t*-values within the cluster. Analysis thereafter was based on these clusters and their associated

Table 1  
Properties of feedback under different economic dimensions.

Domain	Feedback			
	Stimulus	RPE valence	RPE utility	UPE size
Gain	23	–RPE	–29.5	29.5
	24	–RPE	–28.5	28.5
	52	–RPE	–.5	.5
	53	+RPE	.5	.5
	81	+RPE	28.5	28.5
	82	+RPE	29.5	29.5
Loss	23	+RPE	29.5	29.5
	24	+RPE	28.5	28.5
	52	+RPE	.5	.5
	53	–RPE	–.5	.5
	81	–RPE	–28.5	28.5
	82	–RPE	–29.5	29.5

cluster level  $t$ -value, rather than the individual (and highly non-independent)  $t$ -values derived from time  $\times$  electrode points, reducing the multiple comparisons to a manageable number.

Since cluster level  $t$ -values could not be tested for significance against a standard  $t$  distribution, in step two of the procedure, the significance of each cluster was calculated by comparing its cluster-level  $t$ -value to a Monte Carlo generated distribution of cluster level  $t$ -values in the interval occupied by the cluster. This was generated from 50,000 datasets that corresponded to the null hypothesis. To create such a dataset, for each time point within the interval under consideration, the 55 observed  $r$  values were paired with 55 expected values of 0, but this time observed and expected values were switched for a randomly determined number of subjects prior to running the  $t$ -test, ensuring a  $t$ -statistic based on data corresponding to the null hypothesis. The process was repeated for all time points in the interval under simulation and the  $t$  values were summed to create a cluster level  $t$  statistic under the null hypothesis. Having generated 50,000 such cluster level  $t$  statistics, the  $p$  value of the actual observed cluster found in step 1 was calculated as the proportion of the randomisation null distribution that exceeded its cluster-level  $t$  statistic. A Bonferroni correction was then made such that alpha was set to .025 divided by the number of clusters found to be significant at step one. The process was then repeated for the next cluster identified at step one.

### 3. Results

#### 3.1. Behavioural data

Participants were found to answer the probe question correctly on 73.87% of probed trials, with no significant difference in accuracy between loss domain blocks and gain domain blocks (paired samples  $t$  test:  $t_{54} < 1$ ). Preference data collected at the end of the experiment showed a very high correlation between RPE utility and rating ( $r = .91$ ,  $p < .001$ ,  $N = 660$ ). The relationship between RPE utility and rating was the same for both gain and loss domains as shown by very similar average beta values for participant-wise regression of rating against RPE utility (mean  $\beta$  for gain = .93, mean  $\beta$  for loss = .94) with a paired samples  $t$ -test on the beta values showing no significant difference ( $t_{54} < 1$ ). These results indicate that participants attended to feedback and affectively responded to it an appropriate way in both gain and loss domains.

In order to prevent the experiment capturing reinforcement learning effects, participants were told that there was no link between their behaviour and the outcome of the trial. We checked to see if their behaviour was consistent with this belief. Although the outcome of trials was indeed random, this led to some icons being more profitable on a given block than others and it is possible that participants modified their choices on this basis. On a participant/block basis, and after removing cases where icons were never or always chosen by a participant, there was indeed a correlation between the profitability of icons and the frequency with which they were chosen ( $r = .04$ ,  $p = .001$ ,  $N = 6880$ ). We also examined whether the RPE utility on a trial affected the likelihood of switching to a different icon on the following trial and whether this depended on the RPE valence. While intuitively the dependent variable for this analysis would appear to be the decision to switch, and the independent variables to be the RPE utility and valence on the present trial (since these cause the switch), this would result in a binary dependent variable, to be avoided when using ANOVA. Thus, as ANOVA is indifferent to the temporal relationship of the variables we allocated switching behaviour on the next trial (switch vs. no switch) and RPE valence (+ vs. -) as independent variables, and RPE utility on the current trial as the dependent variable. The results showed that RPE utility was indeed significantly lower on trials which were followed by a switch of icons ( $F_{1,105594} = 7.74$ ,  $p < .001$ ,  $\eta^2 < .001$ ). This utility difference was greater for -RPE trials (.29 points) than for +RPE trials (.03 points), producing a significant interaction term ( $F_{1,105594} = 4.97$ ,  $p = .026$ ,  $\eta^2 < .001$ ). Thus despite our instructions, to some degree the participants appeared to regard the

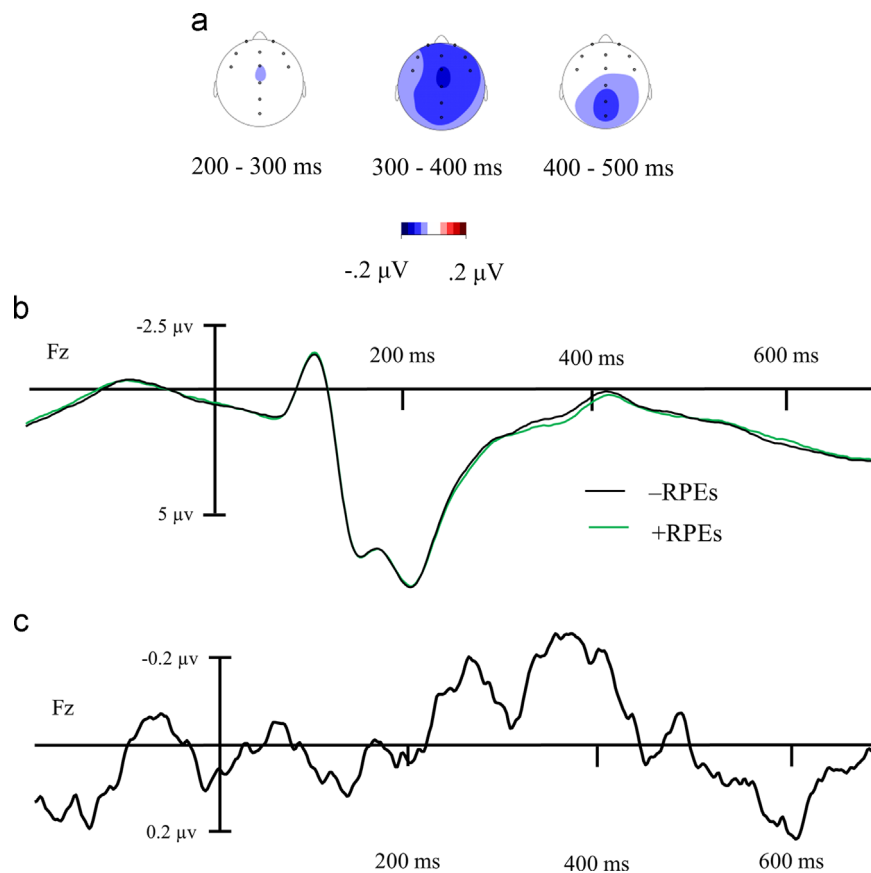
experiment as a reinforcement learning task, and furthermore showed different learning effects for rewards and punishment. These effects, while significant, were of negligible size however, as indicated by the small  $\eta^2$  values. Moreover, the speed with which icons were chosen following the presentation of the icon choice array (interquartile range: 261–592 ms) suggests that for the most part icons were regarded as unimportant, that button presses were simply used to elicit outcomes, and that only feedback was processed.

#### 3.2. Electrophysiological data

While the experiment was designed with a parametric analysis of the effect of RPE utility on voltage in mind, we first present a standard factorial treatment, since this is conventional in the literature. The scalp topography of valence effects, captured by a difference wave of -RPEs and +RPEs, is shown in Fig. 3. In keeping with the literature, this shows an early frontocentral negativity, shifting parietally at greater latencies. Simple waveforms (with no high pass filter) for +RPEs and -RPEs are provided, along with the associated difference wave at Fz. Because there is no accepted convention for how the FRN is quantified in the literature, here we quantified it in three different ways, a peak to peak measure, a mean voltage measure and a peak of difference wave measure. The peak to peak measure was comprised of the difference between the most positive peak in the interval 100–300 ms and the most negative peak in the interval 200–400 ms. A two (valence: good vs. bad)  $\times$  five (electrode: Fz, FCz, Cz, CPz, Pz) ANOVA revealed a significant valence effect ( $F_{1,54} = 8.97$ ,  $p = .004$ ), a significant electrode effect ( $F_{4,216} = 47.52$ ,  $p < .001$ ) but no interaction ( $F_{4,216} = 2.15$ ,  $p = .14$ ). The same ANOVA performed on mean voltage in the interval 200–400 ms revealed an effect close to significance for valence ( $F_{1,54} = 3.64$ ,  $p = .062$ ), and electrode ( $F_{4,216} = 2.88$ ,  $p = .086$ ) but no interaction ( $F < 1$ ). Using a peak of difference wave measure, one sample  $t$ -tests revealed that all electrodes showed a significant negative peak in the interval 200–400 ms. There was no interaction with electrode however, as shown by a non-significant ( $F < 1$ ) effect of electrode on this difference wave peak. The factorial analyses above suggest a mediofrontal response to valence, although this was not significantly greater at frontocentral sites. As anticipated however, the participants' passive stance, and the absence of categorical winning and losing stimuli resulted in weak effects, and so we now turn to the more powerful correlational analyses.

Fig. 4a shows the grand average correlation coefficients of voltage and utility at three representative midline sites. In this figure, deviations from baseline indicate points in time at which voltage amplitude appears to code RPE utility. The approximate threshold for significance ( $p < .05$ ) is shown; however this suffers from the multiple comparisons problem leading to an inflated possibility of Type I error. To correct for this, the Maris and Oostenveld cluster randomisation procedure described above was used, applied in the interval 100–700 ms and at Fz, FCz, Cz, CPz and Pz, and the surviving clusters of RPE-related activity in the waveform are shown in Fig. 4b. These clusters are considered individually below. To aid visualisation of how Fig. 4 would correspond to a standard time  $\times$  voltage plot readers may refer to Supplementary Fig. 1.

Analysis of the +RPEs revealed three clusters of activity that were significant at the .008 threshold set by Bonferroni correction: cluster 1<sup>+</sup> occurring at all midline sites from ~140 to 180 ms ( $p = .000007$ ), cluster 2<sup>+</sup>, for which the largest temporal response was at Fz from 204 to 312 ms ( $p = .0003$ ) and cluster 3<sup>+</sup> for which the largest temporal response was at Pz from 418 to 600 ms ( $p = .000002$ ).



**Fig. 3.** Effects of valence, that is, the average of all +RPEs contrasted with all –RPEs. (a) Scalp topography of difference, (b) grand averages at Fz, and (c) difference wave at Fz.

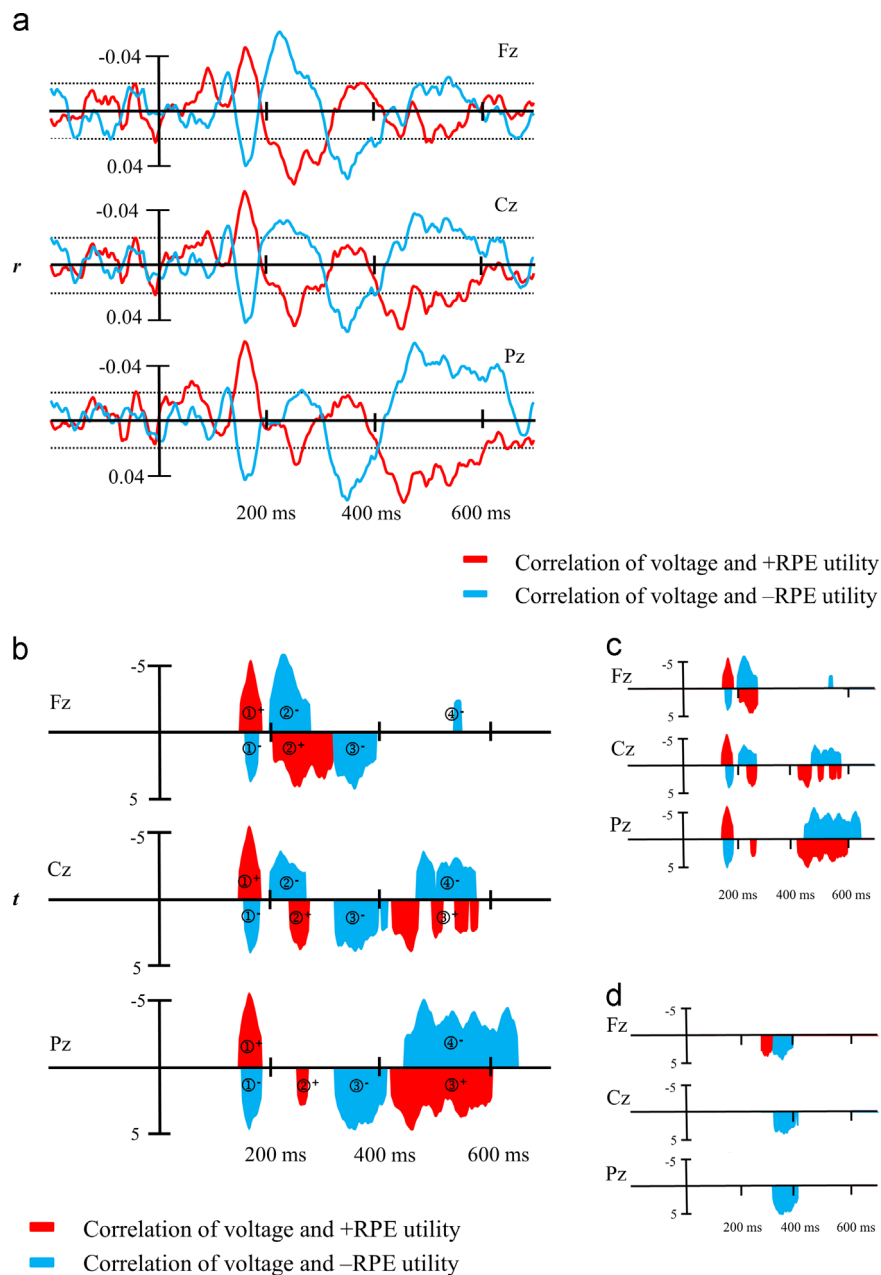
Analysis of the –RPEs revealed four clusters of activity that were significant at the .004 threshold set by Bonferroni correction: cluster 1<sup>–</sup> occurring at all midline sites from ~145 to 180 ms ( $p=.0001$ ), cluster 2<sup>–</sup>, for which the largest temporal response was at Fz from 198 to 272 ms ( $p=.00002$ ), cluster 3<sup>–</sup>, occurring at all midline sites from ~310 to 390 ms ( $p<.000002$ ) and cluster 4<sup>–</sup>, for which the largest temporal response was at Pz from 442 to 648 ms ( $p=.000006$ ).

### 3.3. Discrimination of salience and utility effects

An important objective of the experiment was to distinguish intervals in the post-feedback waveform which appeared to code for the utility of +RPEs and –RPEs from those merely coding for the size of UPEs. Since UPE size and RPE utility are perfectly positively correlated for +RPEs (the bigger the prediction error the higher the utility) and perfectly negatively correlated for –RPEs (the bigger the prediction error, the lower the utility), intervals in which a UPE size response occurs should show a response to RPE utility that is oppositely signed for +RPEs and –RPEs. Fig. 4b allows identification of these. From ~145 to 180 ms, across the midline broadly, a UPE size response appears to occur, as indicated by correlations between voltage and utility which are oppositely signed for +RPEs and –RPEs. The negative sign of the correlation for +RPEs indicates that in the original waveforms increased utility for good outcomes was associated with increased negativity of voltage. In contrast, the positive sign of the correlation for –RPEs indicates that in the case of bad outcomes *decreased utility* (i.e. large –RPEs) was associated with negativity of voltage. Thus large prediction errors, regardless of their sign, induced a voltage-negativity in this interval. A second UPE size response was seen from ~200 to 270 ms, strongest at Fz.

The reversal of the correlation signs relative to the preceding UPE size response indicates that here large prediction errors, regardless of their sign, were associated with positivity of voltage. Note that responsiveness to +RPEs alone persisted beyond this interval up to 312 ms, an effect we consider below. From ~310 to 390 ms across all midline sites there was a significant correlation between voltage and the utility of –RPEs only. The positive sign of the correlation indicates that positivity of voltage was associated with increased utility (i.e. small –RPEs showed more positive voltages than large –RPEs). Finally, from ~400 to 640 ms, and most pronounced at Pz, another UPE size response was seen, with positivity of voltage associated with large prediction errors. Fig. 4c shows the incidence of UPE size responses alone, with utility effects removed.

A considerable advantage of the cluster randomisation technique is that it does not require any assumptions about the timing of components, instead locating *all* intervals of significant activity while nevertheless avoiding increased Type 1 error. Notwithstanding this absence of *a priori* stipulations about where effects should be measured, it is still appropriate to interpret its results in the light of theoretical expectations. One such expectation is that a response to RPE utility will occur in the interval 200–350 ms. In keeping with this expectation, Fig. 4b suggests an encoding of +RPEs from 272 to 310 ms at Fz. However, the cluster randomisation technique has combined this effect with the earlier UPE size coding. This means that a statistical demonstration of coding of +RPE utility is wanting, since while the transition from UPE size to +RPE utility coding at 272 ms is visually compelling, the significance of the +RPE coding depends to an unknown degree on conglomeration with the earlier UPE size effect. To establish the reality of the +RPE encoding in the 272–310 ms interval, the cluster randomisation procedure was run only in the interval 272–



**Fig. 4.** (a) Grand average of Pearson correlation coefficients of voltage and RPE utility. Dotted lines show approximate threshold for significance of a one sample  $t$ -test of correlation coefficients ( $N=55$ , expected value of  $r$  under the null hypothesis is 0). (b) Intervals of significant correlation between voltage and RPE utility after correction for multiple comparisons.  $t$ -values were obtained from a one sample  $t$ -test of correlation coefficients ( $N=55$ , expected value of  $r$  under the null hypothesis is 0) at each time point and formed into clusters of significance for which a cluster-level  $t$  value was calculated as the sum of individual  $t$ -values. This was compared to a Monte Carlo simulated distribution generated under the null hypothesis ( $r=0$ ) to establish significance. Non-significant ( $p > .05$ ) values of  $t$  have been set to zero. See text for further details. (c) Intervals coding for unsigned prediction error size, or salience. (d) Intervals coding for RPE utility.

700 ms. The cluster remained significant at the .0125 threshold set by Bonferroni correction:  $p = .002$ ). Fig. 4d shows the incidence of utility responses alone, with UPE size effects removed.

The use of Pearson correlation coefficients to examine RPE encoding presumes that the relationship between voltage and utility is linear, and this assumption was made to simplify the analysis and its exposition. Having identified intervals of encoding of +RPEs from 272 to 310 ms and -RPEs from 310 to 390 ms under this linear assumption, we examined the nature of the relationship by plotting grand average voltages at each of the 60 levels of utility. Fig. 5a reveals a linear relationship between voltage and utility for both +RPEs and -RPEs. Fig. 5b provides the corresponding plot for the interval 200–270 ms. For comparative purposes, the  $x$  axis has been left as RPE utility; however the

V-shaped function clearly suggests that it is of UPE size that is being coded, and with respect to this variable, the function is once again linear.

#### 4. Discussion

Effective reinforcement learning should be sensitive to both punishment and reward and so should make use of the quantitative information held in the utility of both +RPE and -RPEs. The present experiment investigated whether such sensitivity to both +RPEs and -RPEs was shown in mediofrontal ERPs. To do this, we initially identified intervals of the waveform responsive simply to UPE size, since this might have complicated previous interpretations



of RPE utility effects. While early and late intervals suggested coding for UPE size, the study suggested that in the interval  $\sim 270$ – $390$  ms RPE utility, not UPE size was coded. In this interval, the mediofrontal waveform was responsive to the utility of both +RPEs and –RPEs. Importantly, increased utility in both kinds of RPE was associated with increased positivity of voltage. This is important because integrated coding of +RPEs and –RPEs as depicted in Fig. 1a requires the relationship between utility and voltage to be same-signed for both RPE valences in order to establish a common currency of utility for reward and punishment.

It is worth noting here that the value of the correlation sign is not meaningful in itself and the implications would be unaltered if, in the interval 270–390 ms it was negative for both +RPEs and –RPEs. To draw an analogy with more familiar examples, while the sign of traditional components such as the N2 and P3 is clearly pre-eminent for those components' identification, it does not itself convey any information regarding the component's function. Furthermore, the fact that for both +RPEs and –RPEs the sign of the relationship between utility and voltage alternates across the waveform, as shown in Fig. 4, is not remarkable and can be assumed to indicate sequential and independent components responding to the same properties of feedback but with opposite polarity. A well-known existing example of this is the N2–P3 complex in which unexpected events produce an ERP showing an accentuated negative peak followed by an accentuated positive peak (Folstein & Van Petten, 2008). Indeed, an analogous effect can be seen in Supplementary Fig. 1, where large UPEs are associated with an accentuated negative peak at  $\sim 175$  ms and then an accentuated positive peak at  $\sim 210$  ms and this is reflected in Fig. 4 by a correlation sign switching at  $\sim 200$  ms.

Returning to the observed utility effects, the correlation of utility and voltage for +RPEs at 272–310 ms is notable not for the actual sign of the correlation but simply for the absence of an oppositely signed correlation for –RPEs in the same interval; it is this absence that allows us to discount the possibility that this is a UPE size response, leaving an RPE utility response as the most probable alternative. The same inference applies to the correlation

of utility and voltage for –RPEs at 310–390 ms, and collectively this forms the basis for our primary conclusion that both +RPE utility and –RPE utility are coded in the mediofrontal feedback locked waveform.

As noted above, the relevance of the correlation sign lies in the fact that increased utility was associated with increased positivity of voltage for both kinds of RPE. This suggests the possibility of a single RPE processing system capable of assigning a utility value to all outcomes, both good and bad, in a manner directly comparable to that used by formal reinforcement models. Moreover, the consecutive nature of the RPE utility signal, with –RPE coding following +RPE coding, would appear to be consistent with one account of midbrain RPE generation. It has been proposed that the midbrain dopamine neurons that code +RPEs with phasic increases are limited in their ability to code –RPEs with phasic decreases because of their already low tonic firing rate, and may therefore code the utility of –RPEs with the duration of firing decrease (Bayer, Lau, & Glimcher, 2007; Mileykovskiy & Morales, 2011). This would result in a delay in the transmission of –RPEs to the neural generator of the FRN consistent with the effect shown here. Note however that the scalp topography of the response to +RPE and –RPE utility differed, with +RPEs most pronounced frontally, but –RPEs showing a broader distribution. While it is possible that there is a single frontal source for all RPE-related activity and then an additional source for –RPEs specifically, it is also possible that there is a distinct generator for each kind of RPE. While functionally there is no need for –RPEs and +RPEs to be processed in the same neural structure in order to effect reinforcement learning this does of course undermine the case for a single unitary RPE encoder.

The current study found the strength of coding of –RPE utility to be stronger than coding of +RPE utility. However, a recent review of the FRN by Walsh and Anderson (2012) found coding of +RPE utility to be more commonly observed than –RPE utility coding. One source of this asymmetry may well be the interval of measurement of the FRN, which shows wide variability across the literature, but often does not extend beyond 350 ms or even 300 ms, and thus could miss later –RPE effects. A second source concerns interference from the P3. While this study used a 2 Hz filter in an attempt to mitigate P3 effects, many FRN studies do not. Since the P3 is typically more positive for unexpected and high magnitude outcomes regardless of valence (San Martin, 2012), a late +RPE signal would be augmented by this UPE effect while a late –RPE signal would be diminished. This constitutes a specific example of the distorting effects that components encoding UPE size can have on interpreting utility effects in components that they overlie.

Two other experiments have investigated the ERP response to parametrically manipulated +RPEs and –RPEs. Talmi, Fuentemilla, Litvak, Duzel, and Dolan (2012) explicitly looked for an integrated RPE utility coding across +RPEs and –RPEs and reported failure in this regard, instead finding a UPE size response. This is likely a consequence of the interval chosen however, 200–300 ms, where this study also found a strong UPE size signal. Using concurrent MEG across a wider time range however, the authors found an integrated utility signal at 320 ms, close to the 310 ms point in the present experiment where +RPE utility coding switches to –RPE utility coding.

Pedroni, Langer, Koenig, Allemand, and Jancke (2011) also investigated +RPEs and –RPEs and found consecutive responses at similar latencies to this study: +RPE utility coding at 290–310 ms and –RPE utility coding at 360–380 ms. An important difference however was that the study found the correlation between voltage and –RPE utility to be negative, not positive. A key difference between the study and that of ours and Talmi et al.'s is that in Pedroni et al.'s, –RPEs were achieved by omission of

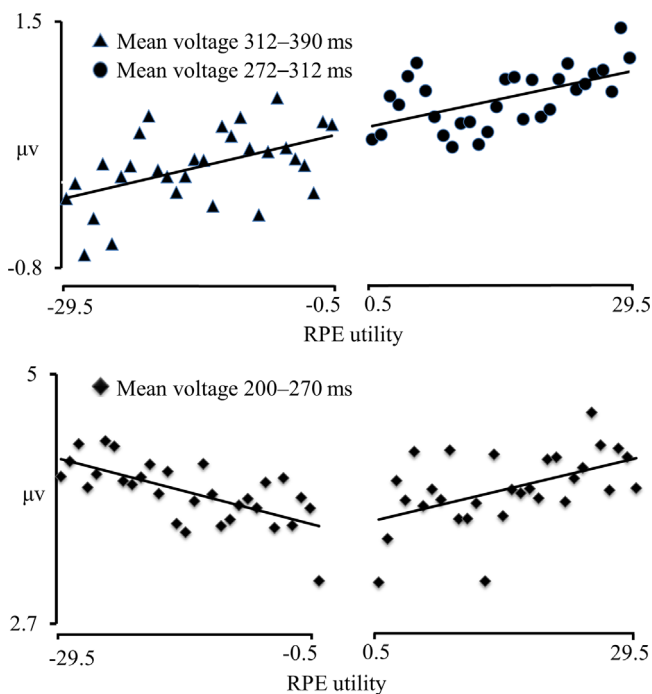


Fig. 5. Grand average of voltages at each level of RPE utility at Fz for (a) +RPEs in the interval 272–312 ms and –RPEs in the interval 312–390 ms, and (b) all prediction errors in the interval 200–270 ms.

reward. Esber and Haselgrove's (2011) claim of an inherent lowered saliency of non-rewards has already been noted. Furthermore, single cell work has suggested the possibility of separate systems for coding reward delivery and reward omission, regardless of the utility that either eventuality comprises. In particular, Joshua, Adler, Mitelman, Vaadia, and Bergman (2008) have found midbrain dopamine neurons that increased their response to aversive outcomes but decreased it to omitted rewards, despite both these outcomes constituting  $-RPEs$ . Thus the manner in which  $-RPEs$  were generated may explain the reversal of the sign observed in Pedroni et al.'s study with respect to that of ours and Talmi et al.'s.

The utility effects at 275–390 ms occurred between what appeared to be strong responses to UPE size occurring in the intervals  $\sim 145$ – $180$  ms,  $\sim 200$ – $270$  ms and  $\sim 400$ – $640$  ms. These intervals were indicated by significant oppositely signed correlations of RPE utility with voltage, represented by the inverted waveforms of  $+RPEs$  and  $-RPEs$  in Fig. 4. The UPE size effect at 200–270 ms is consistent with other demonstrations of strong UPE size effects in this interval (Hauser et al., 2014; Talmi et al., 2013) and fMRI work showing such a signal (Metereau & Dreher, 2013; Rutledge, Dean, Caplin, & Glimcher, 2010). The sustained strong UPE size effect at Pz running from 400 to 640 ms corresponds spatiotemporally to the slow wave component, shown by Foti, Weinberg, Dien, and Hajcak (2011) to be responsive to UPE size rather than RPE utility.

The demonstrated UPE size effects, especially in the 200–270 ms interval, appear to vindicate the concern that FRN measurements are susceptible to UPE size contamination. The direction of the UPE size effect, greater positivity for larger prediction errors, can be expected to increase voltage differences for  $+RPEs$  of different size and decrease them for  $-RPEs$  of different size, leading to an apparent preferential sensitivity of the FRN to  $+RPEs$ . However, this UPE size effect is also of interest in its own right, as the same effects have been seen in single cell recordings (Matsumoto, Matsumoto, Abe, & Tanaka, 2007) and in fMRI meta-analyses (Bartra, McGuire, & Kable, 2013), described as indicating motivational salience (Bromberg-Martin et al., 2010). As large UPEs were as common as small ones in this study, the “saliency” indexed by this UPE size response does not reflect novelty of the stimulus itself, but rather the notability of the appearance of a value that is an outlier with respect to a current estimate of a distribution. Such comparisons, and the updated expectations they produce, are central to Bayesian models of optimal foraging. They might, furthermore, benefit from independence of RPE circuitry: an animal that had eaten to satiation, for example, would still do well to note an unusually large source of food.

As we note earlier, an apparently unambiguous response function can still be an artefact synthesised from two components with quite different response functions where overlap is complete. It is possible that an apparent UPE size response is in fact the aggregate effect of separate populations of neurons coding  $+RPE$  and  $-RPE$  utility. In this scheme, one population codes  $+RPE$  utility as in Fig. 1a, and another codes  $-RPE$  utility as in Fig. 1b, but with a reflected response function, that is with *increases* of activation for large  $-RPEs$  rather than decreases. Such response functions have been found in single cells (Asaad & Eskandar, 2011; Matsumoto et al., 2007) often in close proximity in the primate ventromedial prefrontal cortex and anterior cingulate cortex (Kennerley, Dahmubed, Lara, & Wallis, 2009; Monosov & Hikosaka, 2012; Quilodran, Rothe, & Procyk, 2008). This makes their separation by EEG and fMRI impractical, leaving the question unresolved at present of whether an apparent UPE size signal might in reality reflect an aggregate of two utility signals occurring earlier than the FRN, and which might in fact be involved in its

generation. This interpretation is less parsimonious on face value, partly because it invokes two underlying mechanisms rather than one, but also because neural discrimination between large  $-RPEs$  and large  $+RPEs$  then requires a following neural integrator receiving excitation from one source and inhibition from the other in order to generate the required integrated utility function in Fig. 1c. None of these possibilities are neurally implausible however, and the application of techniques other than EEG will be required to resolve these two interpretations of the UPE size signal that are both equally consistent with the data.

Finally, it is worth discussing the consequences of using a correlational waveform as the basis for analysis rather than a standard voltage-based waveform. It should be noted first that the correlation coefficients are small simply because they are based upon single trial ERP voltages, which are inherently noisy. In interpreting these correlations, the  $p$  value is more relevant than the  $r$  value because the latter is greatly affected by the amount of averaging prior to running the correlation. For example, take a representative correlation,  $r = .04$ , found on the  $+RPE$  correlational waveform of Subject #6 at 306 ms. This value of  $r$  is based on  $\sim 825$  individual data points of voltage plotted against 30 levels of RPE, which reveals a great deal of scatter between the points. By averaging voltage at each RPE level we are left with 30 data points, a much reduced scatter, and now a calculated  $r = .21$ . Reducing this to three bins of low, mid and high RPE gives a reported  $r = .52$ . This final level of averaging better approximates to the sort of comparison and effect size that is seen in standard factorial designs.

The question of peaks should also be addressed. FRN studies often, but by no means always, produce a waveform with an N2 peak superimposed on an ongoing positivity running from  $\sim 100$  ms to  $\sim 400$  ms. Opinion varies as to the value of using this peak to identify the FRN: while some studies use its amplitude, or its amplitude relative to an adjacent positive peak as a measure of the FRN, others ignore the N2 peak, instead using a simple mean amplitude measure across a set interval, or a measure based on a difference wave of good and bad outcome waveforms (either peak or mean amplitude). It is our opinion that the N2 peak of a single (i.e. undifferentiated) waveform is not a reliable guide to either the amplitude or latency of the FRN component. The theoretical basis for this position has been clearly stated by Luck (2005), which is that peaks are not equivalent to components because each peak in a single waveform represents the summed effect of many components. A further important methodological objection to use of the N2 peak is that it may be absent, particularly for large  $+RPEs$ , and setting peak values to zero in such cases (Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003) produces a floor effect that distorts the comparison of  $+RPEs$  and  $-RPEs$ . As argued by Luck (2005), components are generally better described by an *experimental effect*. This has traditionally been achieved by differencing pairs of waves drawn from different levels of an independent variable. While correlational waveforms are simpler and more powerful in cases where the independent variable is continuous, they achieve exactly the same end and should produce qualitatively similar results. In comparison, single waveforms, taken in isolation, can produce quite different and possibly misleading phenomena. As a case in point, single waveforms for the present experiment (see Supplementary Fig. 1) show a negative deflection at  $\sim 175$  ms. However, this is incidental to determining the latency of the FRN, since there is no effect of valence at this point and sensitivity to valence is by definition a property of the FRN component. In contrast, a UPE size effect does occur close to this negative peak, though of course that need not be so, a peak might occur with equal amplitude for large and small UPEs.

To conclude, this experiment investigated whether a medio-frontal ERP response existed to both  $+RPE$  and  $-RPE$  utility. The experimental controlled confounds that have not previously been

controlled. Rather than using a standard factorial design, RPE size was manipulated parametrically, using a powerful correlational analysis. A cluster randomisation based correction for multiple comparisons allowed us to investigate the responsiveness of the waveform beyond the window typically used to assess the FRN without risk of false positives. Our results suggest that medio-frontal ERPs are responsive to the utility of both  $-$ RPEs and  $+RPEs$  and to UPE size (saliency).

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.neuropsychologia.2014.06.004>.

## References

- Asaad, W. F., & Eskandar, E. N. (2011). Encoding of both positive and negative reward prediction errors by neurons of the primate lateral prefrontal cortex and caudate nucleus. *Journal of Neuroscience*, *31*, 17772–17787.
- 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.
- Bayer, H. M., Lau, B., & Glimcher, P. W. (2007). Statistics of midbrain dopamine neuron spike trains in the awake primate. *Journal of Neurophysiology*, *98*, 1428–1439.
- Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience*, *27*, 1823–1835.
- Bellebaum, C., Poleszki, D., & Daum, I. (2010). It is less than you expected: the feedback-related negativity reflects violations of reward magnitude expectations. *Neuropsychologia*, *48*, 3343–3350.
- Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*, *68*, 815–834.
- Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage*, *35*, 968–978.
- den Ouden, H. E. M., Kok, P., & de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, *3*, 1–12.
- Donkers, F. C. L., Nieuwenhuis, S., & van Boxtel, G. J. M. (2005). Medial frontal negativities in the absence of responding. *Cognitive Brain Research*, *25*, 777–787.
- Donkers, F. C. L., & van Boxtel, G. J. M. (2005). Medial frontal negativities to averted gains and losses in the slot-machine task—a further investigation. *Journal of Psychophysiology*, *19*, 256–262.
- Eppinger, B., Mock, B., & Kray, J. (2009). Developmental differences in learning and error processing: evidence from ERPs. *Psychophysiology*, *46*, 1043–1053.
- Esber, G. R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus saliency: a model of attention in associative learning. *Proceedings of the Royal Society B—Biological Sciences*, *278*, 2553–2561.
- Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, *45*, 152–170.
- 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. *Human Brain Mapping*, *32*, 2207–2216.
- Hauk, O., Davis, M. H., Ford, M., Pulvermuller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage*, *30*, 1383–1400.
- Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., et al. (2014). The feedback-related negativity (FRN) revisited: new insights into the localization, meaning and network organization. *Neuroimage*, *84*, 159–168.
- Holroyd, C. B., Krigolson, O. E., & Lee, S. (2011). Reward positivity elicited by predictive cues. *Neuroreport*, *22*, 249–252.
- Holroyd, C. B., Larsen, J. T., & Cohen, J. D. (2004). Context dependence of the event-related brain potential associated with reward and punishment. *Psychophysiology*, *41*, 245–253.
- Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*, *14*, 2481–2484.
- Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*, 688–697.
- Jia, S. W., Li, H., Luo, Y. J., Chen, A. T., Wang, B. X., & Zhou, X. L. (2007). Detecting perceptual conflict by the feedback-related negativity in brain potentials. *Neuroreport*, *18*, 1385–1388.
- Joshua, M., Adler, A., Mitelman, R., Vaadia, E., & Bergman, H. (2008). Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. *Journal of Neuroscience*, *28*, 11673–11684.
- Kahneman, D., & Tversky, A. (1979). Prospect theory—analysis of decision under risk. *Econometrica*, *47*, 263–291.
- Kennerley, S. W., Dahmubed, A. F., Lara, A. H., & Wallis, J. D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables. *Journal of Cognitive Neuroscience*, *21*, 1162–1178.
- 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*, 207–219.
- 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*, 134–138.
- Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
- Marco-Pallares, J., Kramer, U. M., Strehl, S., Schroder, A., & Munte, T. F. (2010). When decisions of others matter to me: an electrophysiological analysis. *BMC Neuroscience*, *11*, 1–8.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*, 177–190.
- Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, *10*, 647–656.
- Metereau, E., & Dreher, J. C. (2013). Cerebral correlates of salient prediction error for different rewards and punishments. *Cerebral Cortex*, *23*, 477–487.
- Mileykovskiy, B., & Morales, M. (2011). Duration of inhibition of ventral tegmental area dopamine neurons encodes a level of conditioned fear. *Journal of Neuroscience*, *31*, 7471–7476.
- Monosov, I. E., & Hikosaka, O. (2012). Regionally distinct processing of rewards and punishments by the primate ventromedial prefrontal cortex. *Journal of Neuroscience*, *32*, 10318–10330.
- Mushtaq, F., Stoet, G., Bland, A. R., & Schaefer, A. (2013). Relative changes from prior reward contingencies can constrain brain correlates of outcome monitoring. *Plos One*, *8*, e66350.
- Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *Journal of Cognitive Neuroscience*, *19*, 1994–2004.
- Pedroni, A., Langer, N., Koenig, T., Allemand, M., & Jancke, L. (2011). Electroencephalographic topography measures of experienced utility. *Journal of Neuroscience*, *31*, 10474–10480.
- Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, *48*, 656–664.
- Potts, G. F., Martin, L. E., Burton, P., & Montague, P. R. (2006). When things are better or worse than expected: the medial frontal cortex and the allocation of processing resources. *Journal of Cognitive Neuroscience*, *18*, 1112–1119.
- Quilodran, R., Rothe, M., & Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, *57*, 314–325.
- Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the reward prediction error hypothesis with an axiomatic model. *Journal of Neuroscience*, *30*, 13525–13536.
- Sambrook, T. D., Roser, M., & Goslin, J. (2012). Prospect theory does not describe the feedback-related negativity value function. *Psychophysiology*, *49*, 1533–1544.
- San Martin, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, *6*.
- San Martin, R., Manes, F., Hurtado, E., Isla, P., & Ibanez, A. (2010). Size and probability of rewards modulate the feedback error-related negativity associated with wins but not losses in a monetarily rewarded gambling task. *Neuroimage*, *51*, 1194–1204.
- Schultz, W. (2010). Dopamine signals for reward value and risk: Basic and recent data. *Behavioral and Brain Functions*, *6*(24), 1–9.
- Sutton, R., & Barto, A. G. (1998). *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press.
- Talmi, D., Atkinson, R., & El-Dereby, W. (2013). The feedback-related negativity signals saliency prediction errors, not reward prediction errors. *Journal of Neuroscience*, *33*, 8264–8269.
- 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*, 635–645.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect-theory—cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.
- Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, *36*, 1870–1884.
- Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex*, *15*, 535–544.
- Yu, R., & Zhang, P. (2014). Neural evidence for description dependent reward processing in the framing effect. *Frontiers in Neuroscience*, *8*, 56.