Received: 4 April 2022 | Revised: 6 December 2022 | Accepted: 12 January 2023

DOI: 10.1111/psyp.14266

PSYCHOPHYSIOLOGY SPR WILEY

ORIGINAL ARTICLE

Valence precedes value in neural encoding of prediction error

Harry Stewardson | **Thomas D. Sambrook**

School of Psychology, University of East Anglia, Norwich, UK

Correspondence
Thomas D. Sambrook, School of Psychology, University of East Anglia, Norwich Business Park, Norwich NR4 7TJ, UK.
Email: t.sambrook@uea.ac.uk

Abstract

Event-related potentials that follow feedback in reinforcement learning tasks have been proposed to reflect neural encoding of prediction errors. Prior research has shown that in the interval of 240–340 ms multiple different prediction error encodings appear to co-occur, including a value signal carrying signed quantitative prediction error and a valence signal merely carrying sign. The effects used to identify these two encoders, respectively a sign main effect and a sign × size interaction, do not reliably discriminate them. A full discrimination is made possible by comparing tasks in which the reinforcer available on a given trial is set to be either appetitive or aversive against tasks where a trial allows the possibility of either. This study presents a meta-analysis of reinforcement learning experiments, the majority of which presented the possibility of winning or losing money. Value and valence encodings were identified by conventional difference wave methodology but additionally by an analysis of their predicted behavior using a Bayesian analysis that incorporated nulls into the evidence for each encoder. The results suggest that a valence encoding, sensitive only to the available outcomes on the trial at hand precedes a later value encoding sensitive to the outcomes available in the wider experimental context. The implications of this for modeling computational processes of reinforcement learning in humans are discussed.

KEYWORDS

adaptive scaling, Bayesian, ERP, FRN, reward prediction error, RewP

1 | INTRODUCTION

1.1 | Discriminating value and valence encoding

It is widely believed that reinforcement learning is driven by the generation of prediction errors immediately following feedback on performance. For this reason, the neural encoders of prediction errors, sources in the brain that change activation proportionally to the prediction error incurred, are an important subject of study. Neuropsychological and behavioral studies of humans and other animals suggest that prediction errors may take a number of forms. One is a *value* prediction error, a quantitative, signed term designating the degree to which an outcome is better or worse than expected. This term underlies most simple model-free

reinforcement learning models and has been shown in single cell studies (Schultz et al., 1997), fMRI studies (O'Doherty et al., 2003), and event-related potentials (ERPs) (Glazer et al., 2018; Holroyd & Coles, 2002; Sambrook & Goslin, 2015; San Martin, 2012; Walsh & Anderson, 2012). Some ERP studies have also found evidence for a simple dichotomous encoding of prediction error *valence*, that is, simply whether outcomes are better or worse than expected but without regard to how much better or worse (Fouragnan et al., 2017; Pedroni et al., 2011; Philiastides et al., 2010). This may occur in response to feedback which is dichotomous by the terms of the experimental design, for example, success versus failure, but this encoding may also be imposed on feedback that is continuous in the amount of reward delivered (Janssen et al., 2016). There is no information in this valence encoding that is not already held in the value encoding, however, it may serve as a quick, coarse evaluation (Philiastides et al., 2010; Tobler et al., 2005) or arise because the architecture of learning involves an initial streaming by valence (Fouragnan et al., 2017). In addition to value and valence reward prediction errors, there is good evidence, from single cell, fMRI, and ERP that prediction error size is encoded without reference to its sign (Brown & Braver, 2005; Matsumoto & Hikosaka, 2009; Talmi et al., 2013). This encoding is uninformative with regard to whether an outcome is rewarding or punishing. However, it may serve to orient an animal to events that carry *motivational salience* (Bromberg-Martin et al., 2010), either because they are unexpected or large. Alternatively, unsigned prediction errors may play an indirect role in reinforcement learning, for example, in determining learning rates (Caze & van der Meer, 2013).

The neural encodings underlying these three kinds of prediction error have frequently been investigated in experiments using a two-way design manipulating prediction error valence and size. The two main effects and interaction term in this design have been so widely used as a means of isolating activity generated by the three encoders described above that it is fair to describe them as "canonical effects". The canonical effect used to show valence is a main effect of positive versus negative prediction errors, the canonical effect for motivational salience is an effect of large versus small prediction errors, and the canonical effect for value is an interaction between prediction error sign and size. The interaction term indicates value encoding insofar as the effects on value of increasing prediction error size depend critically on the sign of the prediction error: value rises as the size of positive prediction errors increases but drops as the size of negative prediction errors increases. In 2×2 designs the value encoding can be demonstrated by a greater difference wave between the large positive and negative prediction errors than between the small positive and negative prediction errors.

ERP studies regularly find both main effects and their interaction at frontocentral electrodes in the feedback-locked waveform. However, not all effects are always found, and effect amplitudes and timings are highly inconsistent across studies. This inconsistency in waveforms is compounded by the choice of interval for statistical reporting. Neural activity in feedback-locked reinforcement learning is typically attributed to a single component, variously described as the feedback-related negativity (Miltner et al., 1997) or the reward positivity (RewP) (Proudfit, 2015) lying between P2 and P3 components. However, a meta-analysis of this component found that the interval in which it was operationalized varied widely between 100 and 600 ms postfeedback (Sambrook & Goslin, 2015). If all three prediction error encodings are present but at slightly different latencies then a literature attempting to resolve each of these encoders into a single scalp component, but which is inconsistent in when that component is measured, will inevitably disagree on what that component is encoding. Drawing directly from published waveforms rather than reported statistics, Sambrook and Goslin's meta-analysis showed all three effects to be robustly present, and jointly so in the interval 240–340 ms. This interval was thus chosen as the interval of interest for the study described here.

We note that the terms feedback-related negativity and reward positivity have been used in a number of senses since the original coinage of the former by Miltner et al. (1997) Amongst the several thousands of papers on the component, the term has been used to refer to the difference wave of good and bad outcomes or to a negative-going peak seen in a simple waveforms. It has been used to refer to the measured scalp effects and to the presumed underlying neural encoder producing these. It has been used to refer to both the valence and value encoders described earlier. Its redesignation as the RewP (Proudfit, 2015) was intended to capture the belief that the underlying encoder activated in response to monetary rewards but not reward omission, producing a positive-going voltage deflection at the scalp specifically for rewards. Subsequently, however, the RewP and feedback-related negativity have been operationalized as separate components either on the trivial basis of the direction in which the good and bad waveforms are differenced or substantially on the basis that they are independent encoders responding to positive and negative feedback, respectively. Because of this ambiguity we avoid using either term except when referencing the literature, instead referring to valence, motivational salience and value *encoders*, and prediction error sign, size and sign × size *effects*.

Temporally resolving encoders within the interval of 240–340 ms would be beneficial on two grounds. First, the more precisely each encoder can be operationalized, the more accurate future research on that encoder is likely to

be. Second, the ordinal sequence of encoders can inform hypotheses about how reinforcement learning is implemented in the brain. The present article takes as its starting point the proposition that while activity in the underlying encoders may overlap in time, the extent of this is likely less than that seen in scalp-recorded ERP components used to detect those encoders due to component overlap. This is particularly likely to be the case with value and valence encoders. Attempting to isolate a valence encoder by searching for the temporospatial interval in which there is a significant difference between the waveforms associated with negative and positive reward prediction errors will produce a result contaminated by the value encoder. This is because a value signal, consisting in increasingly positive voltage from large negative reward prediction errors (RPEs) to large positive RPEs will, if subjected to such an analysis, very often return a significant effect as the monotonic value signal is artificially forced into a step function around the point of zero prediction error. That is to say, valence and value show collinearity. Equating the valence *encoder* to the prediction error sign main *effect* will thus overestimate the strength of this encoder, overstate the interval in which it is active and misrepresent the latency of its peak amplitude. More generally, operationalizing the valence encoder simply in terms of the prediction error sign main effect makes it difficult to distinguish it from value encoding.

The present study improves the resolution of the encoders by capitalizing on a property of the valence encoder that is likely not shown by the value encoder, its *context dependence* (Palminteri & Lebreton, 2021). A number of studies of the RewP (e.g., Holroyd et al., 2004; Kujawa et al., 2013) have suggested that the value of an outcome is framed relative to the alternatives available. Thus, the same monetary outcome may be encoded as a positive or negative prediction error depending on what the alternative was for that trial. A similar effect has been shown at the single cell level by Tobler et al. (2005) in a study which showed that the greater of two rewards produced the same increase in activity and the lesser of the two produced the same decrease in activity, regardless of absolute magnitude. Similar effects have been shown in fMRI (Bunzeck et al., 2010; Nieuwenhuis et al., 2005). The valence encoder thus appears to be dependent on the trial context. This kind of context dependence is a consequence of a ubiquitous wider process of adaptive scaling of neurons' predictions of value, based on the distribution of possible outcomes necessary, a process necessary if a set of neurons is to be capable of encoding expected value and prediction error over variably profitable circumstances. Nevertheless, this scaling appears not to be dictated entirely by the options available on a trial. As noted earlier, the difference wave for large positive and negative prediction errors is greater than that for small positive and negative prediction errors, and importantly this effect is shown even when trial feedback is dichotomous (Sambrook & Goslin, 2015). This implies a value encoder that registers the value of outcomes not only with respect to that trial's alternatives but possibilities on other trials, that is, is *context-free*. Furthermore, this effect appears to be unaffected by whether the range of magnitudes is cued in advance of feedback (Sambrook & Goslin, 2015). A parsimonious interpretation of these effects is that in the interval in which the RewP is typically measured there exists a context-dependent valence encoder in which outcomes are compared to those available on that trial and a context-free value encoder in which outcomes are compared to those available in the wider experiment.

The present study aimed to resolve these two encoders by a manipulation of alternatives available on a trial. To the best of our knowledge this manipulation has not been carried out by any individual experiment. Instead, it has been incidentally provided in the literature insofar as studies using the standard factorial manipulation of prediction error sign and size nevertheless differ with regard to whether they have used *mixed domain gambles* or *single domain gambles*. Here, domain refers to whether outcomes are monetary gains or losses, or in the case of primary reinforcers, appetitive or aversive stimuli, for example, pleasant versus nasty tastes. In mixed domain gamble designs the size of the monetary outcome is fixed on a given trial and participants receive feedback on its domain, that is, whether they have incurred a gain or a loss. In single domain gamble designs, the domain is known in advance and feedback reveals whether the outcome is large versus small (or delivered vs. omitted). Importantly, in mixed domain gambles, losses are always negative prediction errors and gains are always positive prediction errors, but in single domain gambles small gains are negative prediction errors and small losses are positive prediction errors. The literature thus provides a natural three-way mixed factorial design (domain × size × gamble) distributed across studies that is suitable for meta-analysis. This design is capable of discriminating the proposed context-free value and context-dependent valence encoders in a way that individual studies are not.

Figure 1 shows the effects on voltage that are predicted by each of these encoders. Note that the independent variable of prediction error sign has been replaced by domain in the design. This is because prediction error sign associated with an outcome is dependent on gamble design, as we note above, and cannot therefore serve as an orthogonal independent variable. Domain remains an orthogonal variable however, since gains are preferred to losses in all cases and can thereby serve to identify a valence encoding. Figure 1a,b shows the two encodings of principal interest in the study. Importantly, it can be seen in Figure 1a that

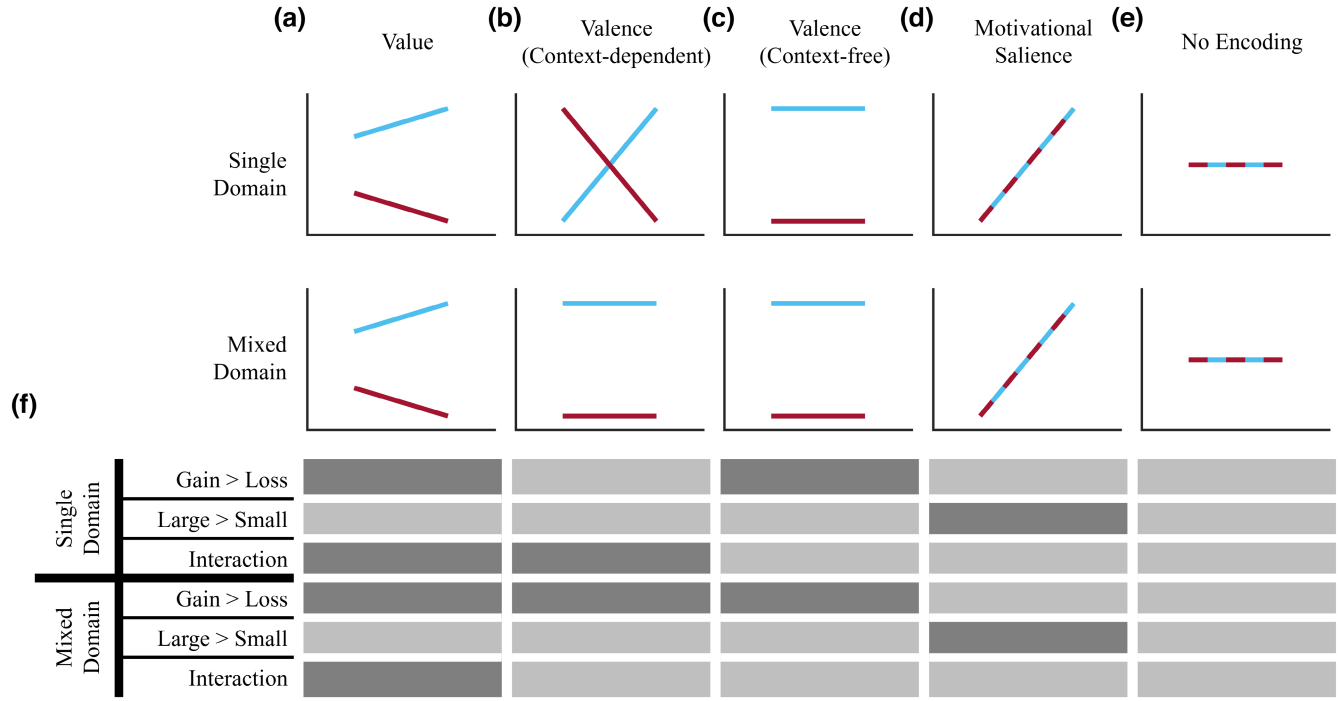

FIGURE 1 Columns show predicted effects on voltage (*y* axis) of five encoders (a–e) and how these depend on prediction error size (*x* axis, small on the left, large on the right), outcome domain (blue = gain, red = loss), and gamble design (rows 1 and 2). (f) Shows predicted effects in simple contrasts, with dark shading indicating an effect is predicted, light shading indicating the null is predicted. The interaction contrast takes the form (large gain – large loss) > (small gain – small loss).

the domain × size interaction generated by the value encoder is unaffected by gamble design, but, in Figure 1b, the context-dependent valence encoder generates an interaction only in single domain gambles. A three-way domain × size × gamble interaction is thus uniquely indicative of this context-dependent valence encoder. The remaining plots in Figure 1 show other encoders that are likely to overlap value and valence encoders. A hypothetical context-free valence encoder is modeled in Figure 1c, although there is no evidence in the literature for such an encoder at present. Its converse, a context-dependent value encoder is not shown since this gives the same predictions as the context-dependent valence encoder under the experimental design. This is because feedback is dichotomous at the level of the trial and so the continuous value signal and the discrete valence signal cannot be distinguished. In Figure 1d, a motivational salience encoder is shown, for which the predicted effects are again the same whether it is context-dependent or context-free. Figure 1e models the absence of any prediction error encoding since this serves as a null in the analysis to come. The predicted effects for specific contrasts of interest used in the subsequent analysis are shown in Figure 1f.

Although the full waveform was analyzed in the forthcoming meta-analysis, the interval 240–340 ms was designated an interval of interest based on evidence that multiple prediction error encodings are carried in this time (Sambrook & Goslin, 2015). Analysis followed two strands. In the first strand, *t* tests were applied at each sample of the great grand average waveform (see Methods) to establish the presence or absence of four effects. Three of these were the canonical effects of domain, size, and domain × size. The fourth was the novel domain × size × gamble effect discussed above. In the second analysis strand, rather than pinning the evidence for each encoder to a single canonical effect, the full range of predicted effects shown in Figure 1f was examined, including predictions of nulls. A Bayesian analysis was used to combine these separate analyses. This was partly because a Bayesian analysis can provide a meaningful quantification of the probability that a null is true and partly because Bayesian tests return simple probabilities rather than *p*-values and these probabilities can be easily multiplied to test conjunction hypotheses. As with the first analysis strand, mass univariate testing (Bayesian *t* tests) was applied at each sample of the waveform. Additionally, Bayesian *t* tests were applied to factors derived from a temporal principal components analysis (tPCA).

1.2 | Gain sensitivity of the valence encoder

The present study additionally investigated the possibility that the valence encoder is specifically sensitive to

outcomes in the gain domain. In principle, a valence encoder tasked with encoding the sign of prediction errors should be indifferent as to the domain in which prediction errors occur. Omitted or smaller than expected aversive events (e.g., monetary losses) should reinforce behavior as effectively as obtained or larger than expected appetitive events (e.g., monetary gains). That is to say, prediction error valence (better or worse than expected) is orthogonal to outcome domain (appetitive goods, aversive goods) and a single valence encoding mechanism should produce adaptive behavior in both domains. In the ERP context, this implies that a valence encoder operationalized as difference between better and worse outcomes should be of the same amplitude whether this is built from a (no loss – loss) waveform or a (gain – no gain) waveform. However, in the wider context of behavior it is nevertheless important that an animal considers domain in addition to prediction error valence in its assessment of the adaptedness of its behavior. An avoided punishment and obtained reward may constitute equal sized positive reward prediction errors, but the latter is clearly preferred when estimating the value of the environment at hand (Boureau & Dayan, 2011). For adaptive reasons, the valence encoder may thus not necessarily observe the formal orthogonality of domain and prediction error sign and this may be reflected in its associated ERP. Indeed, a number of studies have shown greater sensitivity to prediction error sign in the gain domain (Kreussel et al., 2012; Kujawa et al., 2013; Mushtaq et al., 2013; Sambrook et al., 2012). However, an alternative possibility is that this effect is an artifact of component overlap of a valence encoding and a motivational salience encoding. In the interval associated with the RewP motivational salience appears to be encoded by a relative positivity (Sambrook & Goslin, 2015; Talmi et al., 2013). This would serve to cancel out the negativity associated with large losses relative to small ones, producing a small or absent RewP in the loss domain. Conversely it would increase the positivity of large gains, increasing the RewP in the gain domain (Stewardson & Sambrook, 2021). This would produce a misleading gain sensitivity in the RewP.

The present study finds evidence for this gain sensitivity and uses moderators to ascertain whether it is likely to reflect gain sensitivity in the underlying valence encoder or is an artifact of component overlap with a motivational salience encoder. The first moderator is whether gain and loss trials are grouped together in separate blocks or are interleaved. If gain sensitivity is shown in the underlying valence encoder then this sensitivity would be expected to be greater when the two domains are interleaved within a block insofar as information on the domain is presented anew on each trial. In contrast, if the apparent gain sensitivity is an artifact of component overlap, this manipulation should have no effect. This was the finding of Yu and Zhang (2014) in an ERP study on framing effects.

The second moderator is whether feedback cues have an arbitrary or nonarbitrary relationship with the reinforcer they denote. Nonarbitrary cues are similar to the reward or punishment ultimately incurred by the participant, such as displaying "+20¢" on a screen to denote the delivery of a 20¢ reward or displaying an image of a lightning bolt to denote the delivery of an electric shock. Arbitrary cues are arbitrarily mapped to rewards and punishments, such as using one fractal to denote reward and another to denote punishment. The capacity for a cue to invoke the value of the reinforcer (either primary or monetary) appears to be greater when the relationship is nonarbitrary. For example in the "money illusion" participants are heavily influenced by the face value of money presented to them at the point of feedback rather than the real value it denotes, for example, if a 5c win is denoted by a picture of a $5 bill (Huang & Yu, 2019). In the animal context a number of experiments have shown that conditioned stimuli that share sensory properties with primary reinforcers elicit more vigorous responses (Domjan, 2005). Event potentials respond to the calorific content of foods denoted by realistic images and this occurs even when this property is not part of the instrumental learning task at hand (Toepel et al., 2009) and midbrain dopaminergic neuron alerting responses are triggered by stimuli that resemble motivationally salient cues (Bromberg-Martin et al., 2010). We would thus expect nonarbitrary cues to more strongly engage a motivational salience encoder than arbitrary ones. If the gain sensitivity of the valence encoder increases for nonarbitrary cues this would suggest the effect is merely an artifact of component overlap with a motivational salience encoder.

2 | METHOD

2.1 | Inclusion and exclusion criteria

Inclusion and exclusion criteria were as follows. Studies had to manipulate outcome domain (appetitive or aversive) and outcome size (large vs. small, or delivery vs. omission) in a 2 × 2 design with the frequency of occurrence of the four outcomes equal. Domain could be manipulated either by monetary gain or loss or by primary reinforcers (e.g., taste, shock). If more than two levels of outcome size were available, the more extreme values were used. Designs were only included if feedback was dichotomous at the level of the trial. Designs had to conform to mixed gamble or single domain gambles as described in the Introduction. While studies were included in which participants were passive with respect

to outcomes, if a study included an active condition this was used in preference, as prediction error encoding is thought to be stronger in such cases (Stewardson & Sambrook, 2021; Yeung et al., 2005). In order to maintain homogeneity in the meta-analysis studies were restricted to those on young healthy adults. Studies researching other populations were used if a control group of young healthy adults was included. Studies were included only if a grand average waveform for each of the four conditions was provided at a frontocentral electrode (Fz, FCz, Cz, or a frontocentral pool) running from −100 to 500 ms.

2.2 | Search strategies

English language journals and books were searched using the following databases: PsycInfo, PsycArticles, ERIC, PubMed, and Web of Science. The search was conducted in Endnote and the results were compiled in Zotero. The search string used was "feedback related negativity" OR "feedback negativity" OR "FRN" OR "reward positivity" OR "RewP" OR "feedback error-related negativity" OR "feedback error related negativity" OR "feedback correct-related negativity" OR "feedback correct related negativity", and covered journal titles, abstracts, and keywords. This yielded 1244 studies. Studies were then checked for their fit with the inclusion and exclusion criteria and the final sample consisted of 36 studies (20 mixed domain gambles, 16 single domain gambles). Of the single domain gamble studies (which were the portion of the data set used for moderator analysis of gain sensitivity of the RewP), nine were blocked versus seven interleaved, and nine used nonarbitrary cues versus seven using arbitrary cues. Studies are listed in Table 1.

2.3 | Coding procedures

A great grand averaging approach (Sambrook & Goslin, 2015) was used to retrieve data from published articles. Screenshots of published waveforms were digitized with Plot Digitizer software (http://sourceforge.net/projects/plotdigitizer) by using a mouse to manually lay points along the waveform at approximately 5-ms intervals. A purpose-written program available at Sambrook and Goslin (2015) was used to linearly interpolate data at 1-ms intervals in order to provide a consistent set of samples for each waveform in the following mass univariate analyses. Waveforms were replotted from these extracted data and inspected to prevent the introduction of gross errors. The process described above thus retrieved the grand average data for each condition in each experiment, along with a degree of measurement error, estimated to be under 0.1 uv (Sambrook & Goslin, 2015). Waveforms were then differenced to generate the required contrasts shown in Figure 1f and used to test for significance of effects over studies much as single experiments test for significance of effects over participants.

2.4 | Statistical procedures

2.4.1 | Identification of encoders by canonical effects using conventional statistics

Two complementary analysis strategies were used. The first strategy used mass univariate frequentist statistics to generate the time course of four effects of interest: domain, size, domain × size, and domain × size × gamble. Each study provided a simple effect size for each of these effects and significance testing of this effect size across studies constituted our meta-analytic method. Meta-analysis typically weights effect sizes by their sample size or standard deviation. Because the great grand averaging technique directly digitizes the full grand average waveforms rather than relying on statistics reported for a specific interval, standard deviations were unknown at the majority of time points. Weighting was therefore applied using study sample size and for the purpose of testing this necessitated the use of weighted t-tests. The mean effect size over studies was calculated from individual study mean differences whose sample size was used as a weight. To calculate the weighted standard deviation, weights based on study size were applied to deviations of each study mean from the grand mean (Sambrook & Goslin, 2015).

In order to render it suitable for performing *t* tests, the domain × size × gamble design was broken down into simple pairwise contrasts. One-sample weighted *t* tests were used to test the effect of domain (gain – loss), size (large – small), and domain × size ((large gain – small gain) – (large loss – small loss)). A weighted independent samples *t* test was used to test the effect of domain × size × gamble by comparing the domain × size effect in single and mixed domain gambles. Weighted Student *t* tests were performed using the R package 'weights' (Pasek, 2021). Since no weighted Welsh *t* test was available and sample sizes were moderately unequal in the independent samples *t* test (20 vs. 16), a Levene's test was used to assess equality of variance and was found to be non-significant ($p > .05$) throughout the waveform. As a supplementary analysis, the effects of interest were also calculated using unweighted *t* tests and directly from an omnibus 2 × 2 × 2 mixed factorial ANOVA (unweighted) since this would typically be used in a three-way design if weighting were not needed.

TABLE 1 Studies used in the meta-analysis.

Experiment	*N*	Site	Fig.	Gamble	Blocking	Cue type
Aarts and Pourtois (2012)	60	FCz	3h,l	Single	Blocked	Non-arbitrary
Bell et al. (2016)	36	Pool	5	Mixed	Interleaved	Arbitrary
Broyd et al. (2012)	38	Pool	3	Single	Interleaved	Arbitrary
Chen et al. (2017)						
Friend condition	14	Fz	2a	Mixed	Interleaved	Non-arbitrary
Stranger condition	14	Fz	2a	Mixed	Interleaved	Non-arbitrary
Chen et al. (2018)	20	FCz	2	Single	Blocked	Non-arbitrary
Ernst and Steinhauser (2018)	34	FCz	5a	Mixed	Blocked	Arbitrary
Forder and Dyson (2016)	33	Pool	2a	Mixed	Blocked	Arbitrary
Glienke et al. (2015)	20	Fz	4	Mixed	Interleaved	Arbitrary
HajiHosseini et al. (2012)	26	Fz	2a,c	Mixed	Interleaved	Non-arbitrary
Heydari and Holroyd (2016)	30	FCz	4a,b	Single	Blocked	Arbitrary
Hird et al. (2018)	20	Pool	2	Single	Blocked	Arbitrary
Holroyd et al. (2006)	23	Fz	2	Mixed	Interleaved	Non-arbitrary
Kamarajan et al. (2009)	50	FCz	4	Mixed	Interleaved	Non-arbitrary
Kujawa et al. (2013)	22	Pool	1	Single	Interleaved	Non-arbitrary
Long et al. (2018)	26	FCz	3a	Mixed	Interleaved	Arbitrary
Luo and Qu (2013)						
Low magnitude	18	FCz	3a	Mixed	Blocked	Non-arbitrary
High magnitude	18	FCz	3b	Mixed	Blocked	Non-arbitrary
Mei et al. (2018)	32	Pool	5	Single	Blocked	Non-arbitrary
Mei et al. (2018)	57	FCz	1	Single	Interleaved	Arbitrary
Parvaz et al. (2015)	25	FCz	2a	Mixed	Interleaved	Non-arbitrary
Peterburs et al. (2013)	28	FCz	2	Mixed	Interleaved	Non-arbitrary
Pfabigan et al. (2011)	20	FCz	1	Mixed	Interleaved	Arbitrary
Pfabigan et al. (2015)	31	Fz	2	Single	Interleaved	Non-arbitrary
Sambrook et al. (2012)	66	FCz	5	Single	Interleaved	Arbitrary
Santesso et al. (2012)	29	Fz	2	Single	Interleaved	Arbitrary
Sato et al. (2005)	18	Fz	1	Mixed	Interleaved	Arbitrary
Soder and Potts (2018)	50	Pool	2	Single	Blocked	Non-arbitrary
Talmi et al. (2013)	20	Pool	4	Single	Blocked	Arbitrary
Watts and Bernat (2018)	132	Pool	4	Single	Blocked	Non-arbitrary
Wei et al. (2018)	22	FCz	4	Mixed	Interleaved	Non-arbitrary
Wu and Zhou (2009)						
Low expectancy	16	FCz	1	Mixed	Interleaved	Non-arbitrary
High expectancy	16	FCz	1	Mixed	Interleaved	Non-arbitrary
Yu and Zhou (2006a)	20	Fz	1	Single	Interleaved	Non-arbitrary
Yu and Zhou (2006b)	20	Fz	1	Mixed	Interleaved	Non-arbitrary
Zheng et al. (2017)	36	FCz	6	Single	Blocked	Non-arbitrary

2.4.2 | Identification of encoders using Bayesian *t* tests

The second analysis strategy used Bayesian one-sample *t* tests (Rouder et al., 2009), again applied at each time point, but this time applied to the difference waves corresponding to the six contrasts shown in Figure 1f. These contrasts comprise the full set of contrasts that discriminate the encoders. One-sample Bayesian *t* tests were used to assess at each time point the extent to which the evidence favored the prediction (be that effect or null) shown in each cell of Figure 1f. Where an effect was predicted, the *t* test calculated the evidence for this versus the combined alternatives, that is, a null or opposite effect. Where a null was predicted the t test established the evidence for this against an effect in either direction. *T* tests were

performed using the ttestBF function in the R package 'BayesFactor' (Morey & Rouder, 2015). A noninformative Jeffreys prior was placed on the variance of the normal population, and a Cauchy prior placed on the standardized effect size, scaled to $2/\sqrt{2}$. Alternative scalings of 0.5 and 1 were also run to check the robustness of scaling and returned very similar results. The Bayes Factor produced by each test showed the ratio of the evidence for the prediction shown in a given cell in Figure 1f to the evidence for the two alternatives. The Bayes Factor was converted to a simple probability using a ratio to probability conversion, that is, probability = Bayes Factor/(Bayes Factor + 1). After all cells in Figure 1f had been so populated with probabilities, the model evidence for each of the five encoders was calculated by the column wise product of the six observed probabilities. Finally, these compound probabilities were rescaled so that column products summed to one, with this achieved by simply dividing each compound probability by the sum of the five compound probabilities. The resulting scaled value thus represented the *relative probability* of each encoder compared to the others. Unlike analysis strand one, Bayesian *t* tests were not weighted by study sample size due to the absence of any agreed method of applying these weights. However, a comparison of weighted and unweighted t-test output for analysis strand one (see Results), suggested that for the current data set the effects of weighting were small in the interval of interest in any case.

Temporal principal components analysis (tPCA) was performed on the six difference waves described in Figure 1f using identical means to those used by Sambrook and Goslin (2016) and following published guidelines (Dien, 2010; Dien et al., 2005, 2007). The ERP PCA Toolkit version 2.89 was used (Dien, 2010). This used each time point as a variable and each of the six difference wave contrasts in Figure 1f as observations. This yielded 108 observations (domain, size, and interaction waves from 20 mixed gamble designs and from 16 single domain gamble designs). A covariance association matrix was used. Factors were retained if they explained more variance than a factor extracted from a null data set, that is, they passed a parallel test (Horn, 1965) and were subjected to the Promax rotation algorithm. Factors oscillating around the baseline were discarded since these were unlikely to underly ERP components shown in the undecomposed data. Factors accounting for less than 3% of the data were discarded. For the purposes of visualization, factors were reconstructed into voltage denominated waveforms using the product of the factor pattern matrix and the standard deviations (Dien et al., 2003). These could then be interpreted in the same manner as the original waveforms before they underwent decomposition by tPCA. For the identification of encoders, the same process of Bayesian *t* tests as described above was applied but instead of sequentially applying testing at each point on the waveform testing was applied on the single factor score supplied by the tPCA for each study × contrast observation.

2.5 | Estimation of small sample and publication bias

The phenomena under study in the present study are unlikely to have been subject to publication bias. The RewP is a robustly elicited component and the primary studies used in this meta-analysis studied treatment effects rather than the existence of the component itself. None were directly concerned with the relative timings of encoders. Additionally, the great grand average technique enjoys considerable protection from publication bias because it begins analysis anew with the full waveform rather than starting with statistics based on a time-specific operationalization of the component that led to publication. The method's post hoc repurposing of data can thus be expected to greatly reduce the risk of publication bias. Publication bias was assessed by correlating sample size and effect size at each point on the great grand average waveform. No significant correlation was found in the interval of interest for any of the three within-subjects canonical effects (see Figure S1).

2.6 | Gain sensitivity of the valence encoder

For the analysis of gain sensitivity of the valence encoding, the standard RewP operationalization was used, that is, better outcome – worse outcome. Since the aim was to investigate how this varied between gain and loss domains, only single domain gamble studies could be used. Gain sensitivity was operationalized as the difference of the RewP in the two domains, that is, (large gain – small gain) – (small loss – large loss). The effect of two moderators, blocking and cue type was examined with weighted independent samples t tests. Levene's tests performed on these moderators showed the assumption of equality of variance to be met in the interval of interest.

3 | RESULTS

3.1 | Identification of encoders by canonical effects using conventional statistics

Simple waveforms for the eight cells in the design are given in Figure 2. Figure 3 shows *t* values for effects of interest. The size effect, indicating motivational salience

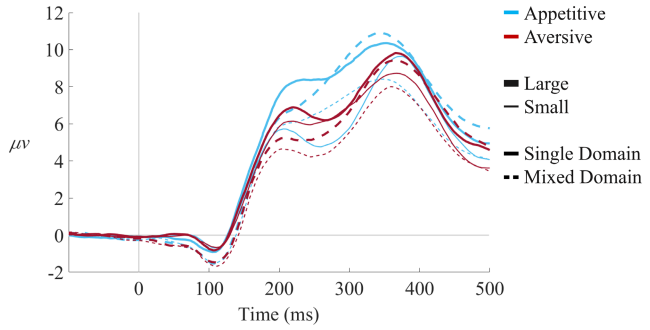

FIGURE 2 Simple great grand average waveforms for the 2×2×2 design.

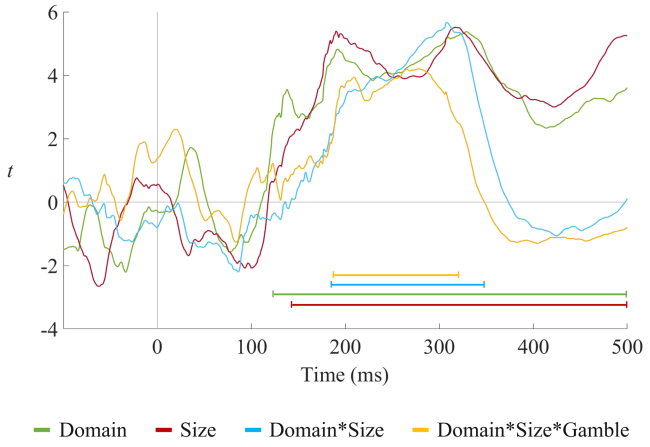

FIGURE 3 Weighted *t*-values for canonical effects with intervals of significance shown. Critical *t* value ($p < .025$, two tailed) is $t(35) = 2.03$ for all effects other than the three way interaction for which $t(19.02) = 2.09$.

encoding is significant from 141 ms to interval end at 500 ms. Domain and domain × size effects, indicative of both value and valence encoders (see Figure 1), are significant from 122 ms to interval end and 183–350 ms, respectively, peaking at 329 ms (simple effect size 2.32 uv, 95% CI [1.47, 3.17]) and 317 ms (simple effect size 1.92 uv, 95%CI [1.22, 2.63]), respectively. The three-way domain × size × gamble effect, uniquely indicative of the valence encoder, is significant from 188 to 327, peaking at 278 ms, with simple effect size 3.17 uv, 95%CI [1.69, 4.65]. This interaction respected the form predicted for a context-dependent valence encoder (see Figure 1b) with a stronger domain × sign interaction in single domain gambles than mixed domain gambles (see Figure S2). For all effects described above, unweighted *t* tests produced similar effects in the interval of interest (see Figure S3) as did an unweighted factorial ANOVA (see Figure S4). As a further supplementary test, electrode site (FCz vs. Fz) was introduced as an additional factor in the design and the interaction of this factor with the others inspected. This produced modest three-way interactions of site with form × size and with domain × size with these effects being greater at FCz (see Figure S5).

3.2 | Bayesian approach to encoder identification

Figure 4a shows the relative probability of each encoder based on Bayesian *t* tests. Prior to 132 ms, evidence favors no prediction error encoding. Motivational salience is indicated from 132–189 ms (peak 161 ms, relative probability = .93), context-dependent valence from 190–276 ms (peaks 205 and 250 ms, relative probability = .75 and .71), value from 278–342 ms (peak 328 ms, relative probability = .93) and motivational salience from 343 ms to interval end at 500 ms (relative probability > .90 from 350 ms onwards). Removing motivational salience from the model had only small effects: no encoding until 160 ms, valence from 161–277 ms, value from 278–366 ms, no encoding from 367–487 ms, and context-free valence encoding from 488 to interval end (see Figure S6).

tPCA produced three temporal factors. These are shown in Figure 4b. Factor 1 accounted for 39.64% of the variance, peaked at 384 ms, and showed very strong evidence (relative probability = .99) of being a motivational salience encoder. Factor 2 accounted for 39.57% of the variance, peaked at 250 ms, and showed moderate evidence of being a valence encoder (relative probability = .74). Factor 3 accounted for 11.00% of the variance, peaked at 302 ms, and showed moderate evidence of being a value encoder (relative probability = .73). The close agreement in timing with effects seen in Figure 4a suggests those effects are based on discrete encoders identified by the tPCA.

3.3 | Gain sensitivity

When valence encoding, using the RewP, was investigated separately in the gain and loss domains, strong gain sensitivity was found running from 177–365 ms, with no RewP observable in the loss domain (see Figure S7). However, experiments in which domain was interleaved produced slightly reduced rather than increased gain sensitivity and there was no significant effect of blocking on gain sensitivity at any point in the waveform (see Figure S8). These results do not support the idea of an intrinsically gain sensitive valence encoder. In contrast, the effect of cue type (nonarbitrary vs. arbitrary) on the effect of outcome size was shown to be significant in the interval 294–344 ms ($p < .025$ critical $t = 2.20$, d.f. = 11.06): see Figure S9. These results thus support the existence of an overlapping motivational salience encoding giving the false appearance of gain sensitive valence encoder.

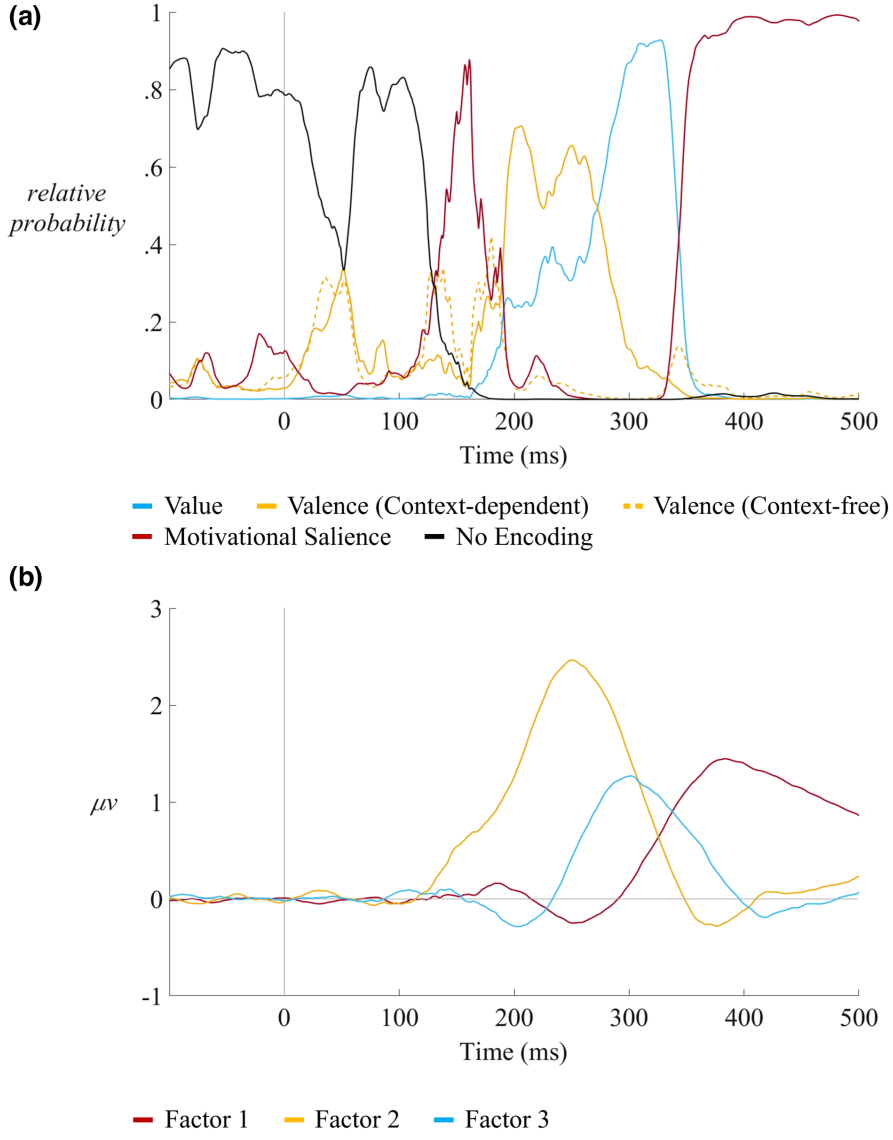

FIGURE 4 (a) Relative probability of each encoder based on Bayesian *t* tests performed on difference waves described in Figure 1f. (b) Principal components derived from difference waves described in Figure 1f.

4 | DISCUSSION

The present study attempted to resolve the timing of prediction error encoders suspected to be present in the feedback-locked frontocentral waveform and used analyses designed to avoid favoring one encoder over another. It used a meta-analytic technique that preserved information in the full waveform provided by individual studies and so removed the danger of encoders showing inflated presence due to the original researchers' focus of interest. Because the discrimination of value and valence encoders was not the focus of any of these papers, the findings are unlikely to be subject to publication bias.

In keeping with previous work (Foti et al., 2011; Sambrook & Goslin, 2016; Zheng et al., 2015) the present study found an interval of reward prediction error activity lying between two intervals of unsigned prediction error activity. Importantly, it resolved reward prediction error activity in the interval 240–340 ms into an earlier valence encoding and a later value encoding. The three-way interaction of domain, size, and gamble, uniquely diagnostic of the valence encoder, suggested the valence encoder peaks at 278 ms, a finding corroborated by tPCA. This is followed by a value encoder characterized by a domain × size interaction that occurs regardless of gamble design, that is, which shows no three-way interaction.

The present study adds to a number of studies that have attempted to isolate a categorical valence encoding from a later quantitative encoding of prediction error. An additional finding in the present study is that the apparently stronger encoding of valence in the gain domain is likely a consequence of component overlap with a motivational

salience component, suggesting there is no need to posit separate valence encoders for gain and loss domains.

Because value encoding, a representation of signed prediction error, necessarily carries within it an encoding of valence, the presence of a separate, earlier, valence-only encoding suggests this may serve to initiate separate learning processes in response to positive and negative reward prediction errors. Separate learning by prediction error sign is not predicted in traditional models of reinforcement learning (O'Doherty et al., 2007; Sutton & Barto, 1998). Furthermore, while learning mechanisms underlying approach and avoidance behavior may benefit from functional separation, the adaptive consequences of approach or avoidance are based on the valence of the prospective outcome, what we term its domain, rather the valence of recent prediction errors. Nevertheless there is behavioral evidence that people's reinforcement learning differs in response to positive and negative prediction errors (Frank et al., 2005) and that there is a genetic basis for this (Frank et al., 2004; Frank & Hutchison, 2009). Differences in learning rates (Gershman, 2015; Palminteri et al., 2017) have been proposed to underlie these effects, and theoretical models (e.g., Caze & van der Meer, 2013) have suggested an adaptive basis for separate learning rates for positive and negative prediction errors. The present findings are consistent with an initial registration of valence followed by that of signed prediction error to which different learning rates can be applied depending on the concurrent valence signal.

An alternative possibility is that there is a more marked functional separation of positive and negative prediction errors by anatomical structure or neurotransmitter (e.g., Daw et al., 2002) and that the learning processes associated with one or other sign of prediction error are simply less detectable at the scalp, for example, due to dipole orientation. Under this account the valence signal shown in the present study and elsewhere may not reflect a learning signal at all, but rather result from earlier general processes associated with motivational events, common to both better and worse than expected outcomes but detected at the scalp for one or other prediction error sign. Which prediction error sign this might be cannot be reliably identified by an analysis of simple effects in a 2×2 domain×size design due to the likely interference of components associated with motivational salience or unsigned prediction error (see Sambrook & Goslin, 2016, Figure 1). However, a number of studies (Foti et al., 2011; Sambrook & Goslin, 2016; Zheng et al., 2015) have found that scalp activity in the interval associated with the RewP correlates with parametrically varied positive but not negative prediction error.

A methodological feature of this study was the characterization of the waveform in terms of the relative evidence for each of a set of encoders of interest. The strength of this approach is that all expected effects and all expected nulls predicted by an encoder are tested for and contribute to the overall evidence. Bayesian rather than frequentist statistics were used. This was partly since these are capable of quantifying the evidence for the null but partly because they generate raw probabilities rather than *p*-values which can then be assimilated into an overall evidence term by simply multiplying them together. Since the resulting products are very small, interpretation requires them to be scaled relative to each other, producing *relative* evidence. Such relative evidence has some undesirable consequences. The method tacitly assumes only one encoder is active at a given time. If two encoders are simultaneously active, the evidence assigned to one will be down-weighted to the extent that evidence for another is present. Thus, an encoder that is growing in strength will nevertheless appear to be waning if another encoder is growing in strength faster. The method may thus be prone to misattributing the true peak of an encoder based on a seemingly unwarranted comparison to another encoder's activity. As noted earlier however, ignoring the fact of component overlap and independently assigning peaks to encoders based on canonical effects is equally prone to mislocating peaks when encoders share those canonical effects. Because both methods are compromised by component overlap but are compromised in different ways they are best used in concert. This is exemplified by the comparison of Figures 3 and 4a. Figure 3, which implements the canonical effects method, shows the domain×size effect to be largely nested within the domain effect. This might be taken to imply a shorter value encoding within a longer valence encoding, and this nesting was also shown in an earlier meta-analysis (Sambrook & Goslin, 2015) by similar means. However, since the domain main effect is produced in all cases by both encoders, but the domain×size interaction is not always present (see Figure 1a,b), a shorter, nested interval would necessarily be expected for the interaction and this nesting is thus an artifact of the method. In Figure 4a, the estimated latency for the peak of the value encoder lies outside that for the valence encoding. This also reflects an artifact of the method, namely that the two encodings cannot be coincident. In short, the canonical effects approach will tend to underestimate the gap between two encoders' activity and the relative probability approach will tend to overestimate it. There is evidence of this in the present data. The relative probability method places valence and value peaks at 206 ms and 328 ms while canonical effects places them much closer at 329 and 317 ms. Additionally, in the case of the valence encoder, the present study provides an alternative canonical effect in the form of the three-way interaction which should be unaffected by component

overlap. This positions the valence effect at an intermediate 278 ms, supporting the predicted effects of component overlap described above. tPCA, also potentially capable of resolving component overlap, positions the valence encoding at 245 ms and the value encoding at 302 ms, thus also lying between the estimates made by the methods prone to component overlap.

It should also be noted that while the relative probability method is prone to forcing the estimated latencies of encoders apart it will do this in a direction that should respect their true underlying ordinal sequencing. With respect to the current study, while the true latency of valence and value encoders may be unresolved, there is clear evidence that valence encoding precedes value encoding. This informs our understanding of how these encoders interact. For example, valence encoding may serve as a preliminary "streaming" of feedback as has been suggested by Fouragnan et al. (2017), dispensing with the need for later value encoding. Additionally, this ordering suggests that valence encoding is not achieved by a dichotomization of an earlier value encoding.

Some limitations of the study must be noted. One is the restriction of analysis to frontocentral electrodes and the absence of data from parietal sites. Reward prediction errors appear to produce a separate parietal component in the temporospatial interval associated with the P3 (Stewardson & Sambrook, 2020) and the effects of this encoder are likely felt at the frontocentral site. This will result in some mischaracterization of encoders which might have been resolvable with spatial information and a temporospatial principal components analysis. Nevertheless, insofar as much research published on the RewP is based on analysis only of a single frontocentral electrode, this meta-analysis provides appropriate intervals for operationalizing components restricted to that site. Another potential limitation lies in the absence of any time-frequency decomposition. For example, HajiHosseini and Holroyd (2015) found a valence encoding in the beta frequency band in the later interval of 350–500 ms.

The meta-analysis was restricted to studies in which feedback was dichotomous at the level of the trial. Given the relative scarcity of studies that present three or more discrete outcomes, or provide feedback on a continuous scale, this was necessary to produce a sufficiently homogenous data set. Nevertheless, studies with dichotomized feedback allow for multiple interpretations. While the evidence presented here is consistent with the existence of a valence encoding, there are other possibilities. One is that feedback is binarized but not around the value of zero prediction error. Hajcak et al. (2006), for example, proposed that in a four-outcome task, the RewP classifies all but the best outcome as equally bad, and a number of authors have suggested that neutral outcomes are categorized with bad outcomes (Holroyd et al., 2006; Toyomaki & Murohashi, 2005). Another possibility is that rather than valence encoding, the component observed here reflects error monitoring (Gehring & Willoughby, 2002) or goal congruency (Fromer et al., 2019). In support for the present interpretation however, Janssen et al. (2016) found strong evidence for valence encoding even when feedback was continuous. Nevertheless, demonstrating the same pattern of effects shown here using non dichotomized feedback would strongly support the case we have made.

5 | CONCLUSIONS

Separating neural encoding of prediction error valence from prediction error value is possible in experimental designs that manipulate prediction error sign and size. However, a simple analysis (termed here a canonical effect) of whether outcomes are better or worse than expected does not isolate valence encoding because value encoding is also sensitive to this outcome. In contrast, a Bayesian analysis of specific contrasts in the design, including those that predict a null can separate valence and value encodings. Additionally, gamble design, a property typically held constant within a given experiment becomes a variable over studies allowing meta-analysis to isolate prediction error valence encoding by means of a three-way interaction of domain, size, and gamble. The present study employs all these methods as well as temporal principal components analysis to separate valence and value encodings showing the former to occur somewhat earlier. Further analyses suggest the valence encoder is not preferentially sensitive to the valence of outcomes in the gain rather than loss domain implying it may serve a general-purpose role in reinforcement learning.

AUTHOR CONTRIBUTIONS

Harry Stewardson: Conceptualization; formal analysis; investigation; writing – review and editing. **Thomas D. Sambrook:** Conceptualization; formal analysis; supervision; writing – original draft.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Thomas D. Sambrook https://orcid.org/0000-0001-5260-797X

REFERENCES

Aarts, K., & Pourtois, G. (2012). Anxiety disrupts the evaluative component of performance monitoring: An ERP study.

Neuropsychologia, *50*(7), 1286–1296. https://doi.org/10.1016/j.neuropsychologia.2012.02.012

Bell, R., Sasse, J., Moller, M., Czernochowski, D., Mayr, S., & Buchner, A. (2016). Event-related potentials in response to cheating and cooperation in a social dilemma game. *Psychophysiology*, *53*(2), 216–228. https://doi.org/10.1111/psyp.12561

Boureau, Y. L., & Dayan, P. (2011). Opponency revisited: Competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology*, *36*(1), 74–97. https://doi.org/10.1038/npp.2010.151

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron*, *68*(5), 815–834. https://doi.org/10.1016/j.neuron.2010.11.022

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*(5712), 1118–1121. https://doi.org/10.1126/science.1105783

Broyd, S. J., Richards, H. J., Helps, S. K., Chronaki, G., Bamford, S., & Sonuga-Barke, E. J. S. (2012). An electrophysiological monetary incentive delay (e-MID) task: A way to decompose the different components of neural response to positive and negative monetary reinforcement. *Journal of Neuroscience Methods*, *209*(1), 40–49. https://doi.org/10.1016/j.jneumeth.2012.05.015

Bunzeck, N., Dayan, P., Dolan, R. J., & Duzel, E. (2010). A common mechanism for adaptive scaling of reward and novelty. *Human Brain Mapping*, *31*(9), 1380–1394. https://doi.org/10.1002/Hbm.20939

Caze, R. D., & van der Meer, M. A. A. (2013). Adaptive properties of differential learning rates for positive and negative outcomes. *Biological Cybernetics*, *107*(6), 711–719. https://doi.org/10.1007/s00422-013-0571-5

Chen, Y. M., Xu, J., Zhou, L., & Zheng, Y. (2018). The time course of incentive processing in anticipatory and consummatory anhedonia. *Journal of Affective Disorders*, *238*, 442–450. https://doi.org/10.1016/j.jad.2018.05.053

Chen, Y. Z., Lu, J. M., Wang, Y. W., Feng, Z. Q., & Yuan, B. (2017). Social distance influences the outcome evaluation of cooperation and conflict: Evidence from event-related potentials. *Neuroscience Letters*, *647*, 78–84. https://doi.org/10.1016/j.neulet.2017.03.018

Daw, N. D., Kakade, S., & Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks*, *15*(4–6), 603–616. https://doi.org/10.1016/S0893-6080(02)00052-7

Dien, J. (2010). The ERP PCA toolkit: An open source program for advanced statistical analysis of event-related potential data. *Journal of Neuroscience Methods*, *187*(1), 138–145. https://doi.org/10.1016/j.jneumeth.2009.12.009

Dien, J., Beal, D. J., & Berg, P. (2005). Optimizing principal components analysis of event-related potentials: Matrix type, factor loading weighting, extraction, and rotations. *Clinical Neurophysiology*, *116*(8), 1808–1825. https://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. *Human Brain Mapping*, *28*(8), 742–763. https://doi.org/10.1002/Hbm.20304

Dien, J., Spencer, K. M., & Donchin, E. (2003). Localization of the event-related potential novelty response as defined by principal components analysis. *Cognitive Brain Research*, *17*(3), 637–650. https://doi.org/10.1016/S0926-6410(03)00188-5

Domjan, M. (2005). Pavlovian conditioning: A functional perspective. *Annual Review of Psychology*, *56*, 179–206. https://doi.org/10.1146/annurev.psych.55.090902.141409

Ernst, B., & Steinhauser, M. (2018). Effects of feedback reliability on feedback-related brain activity: A feedback valuation account. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(3), 596–608. https://doi.org/10.3758/s13415-018-0591-7

Forder, L., & Dyson, B. J. (2016). Behavioural and neural modulation of win-stay but not lose-shift strategies as a function of outcome value in rock, paper, scissors. *Scientific Reports*, *6*, 33809. https://doi.org/10.1038/srep33809

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*(12), 2207–2216. https://doi.org/10.1002/Hbm.21182

Fouragnan, E., Queirazza, F., Retzler, C., Mullinger, K. J., & Philiastides, M. G. (2017). Spatiotemporal neural characterization of prediction error valence and surprise during reward learning in humans. *Scientific Reports*, *7*, 4762. https://doi.org/10.1038/s41598-017-04507-w

Frank, M. J., & Hutchison, K. (2009). Genetic contributions to avoidance-based decisions: Striatal D2 receptor polymorphisms. *Neuroscience*, *164*(1), 131–140. https://doi.org/10.1016/j.neuroscience.2009.04.048

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science*, *306*(5703), 1940–1943. https://doi.org/10.1126/science.1102941

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*(4), 495–501. https://doi.org/10.1016/j.neuron.2005.06.020

Fromer, R., Wolf, C. K. D., & Shenhav, A. (2019). Goal congruency dominates reward value in accounting for behavioral and neural correlates of value-based decision-making. *Nature Communications*, *10*, 4926. https://doi.org/10.1038/s41467-019-12931-x

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. https://doi.org/10.1126/science.1066893

Gershman, S. J. (2015). Do learning rates adapt to the distribution of rewards? *Psychonomic Bulletin & Review*, *22*(5), 1320–1327. https://doi.org/10.3758/s13423-014-0790-3

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

Glienke, K., Wolf, O. T., & Bellebaum, C. (2015). The impact of stress on feedback and error processing during behavioral adaptation. *Neuropsychologia*, *71*, 181–190. https://doi.org/10.1016/j.neuropsychologia.2015.04.004

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, *71*(2), 148–154. https://doi.org/10.1016/j.biopsycho.2005.04.001

HajiHosseini, A., & Holroyd, C. B. (2015). Sensitivity of frontal beta oscillations to reward valence but not probability.

Neuroscience Letters, *602*, 99–103. https://doi.org/10.1016/j.neulet.2015.06.054

HajiHosseini, A., Rodriguez-Fornells, A., & Marco-Pallares, J. (2012). The role of beta-gamma oscillations in unexpected rewards processing. *NeuroImage*, *60*(3), 1678–1685. https://doi.org/10.1016/j.neuroimage.2012.01.125

Heydari, S., & Holroyd, C. B. (2016). Reward positivity: Reward prediction error or salience prediction error? *Psychophysiology*, *53*(8), 1185–1192. https://doi.org/10.1111/psyp.12673

Hird, E. J., El-Deredy, W., Jones, A., & Talmi, D. (2018). Temporal dissociation of salience and prediction error responses to appetitive and aversive taste. *Psychophysiology*, *55*(2). https://doi.org/10.1111/psyp.12976.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037//0033-295x.109.4.679

Holroyd, C. B., Hajcak, G., & Larsen, J. T. (2006). The good, the bad and the neutral: Electrophysiological responses to feedback stimuli. *Brain Research*, *1105*, 93–101. https://doi.org/10.1016/j.brainres.2005.12.015

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*(2), 245–253. https://doi.org/10.1111/j.1469-8986.2004.00152.x

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185.

Huang, Y., & Yu, R. J. (2019). Common and distinct neural substrates of the money illusion in win and loss domains. *NeuroImage*, *184*, 109–118. https://doi.org/10.1016/j.neuroimage.2018.09.032

Janssen, D. J. C., Poljac, E., & Bekkering, H. (2016). Binary sensitivity of theta activity for gain and loss when monitoring parametric prediction errors. *Social Cognitive and Affective Neuroscience*, *11*(8), 1280–1289. https://doi.org/10.1093/scan/nsw033

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

Kreussel, L., Hewig, J., Kretschmer, N., Hecht, H., Coles, M. G. H., & Miltner, W. H. R. (2012). The influence of the magnitude, probability, and valence of potential wins and losses on the amplitude of the feedback negativity. *Psychophysiology*, *49*(2), 207–219. https://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. https://doi.org/10.1111/psyp.12002

Long, C. Q., Sun, Q., Jia, S. W., Li, P., & Chen, A. T. (2018). Give me a chance! Sense of opportunity inequality affects brain responses to outcome evaluation in a social competitive context: An event-related potential study. *Frontiers in Human Neuroscience*, *12*, 135. https://doi.org/10.3389/fnhum.2018.00135

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

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*(7248), 837–U834. https://doi.org/10.1038/nature08028

Mei, S. T., Yi, W., Zhou, S. Y., Liu, X., & Zheng, Y. (2018). Contextual valence modulates the effect of choice on incentive processing. *Social Cognitive and Affective Neuroscience*, *13*(12), 1249–1258. https://doi.org/10.1093/scan/nsy098

Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes factors for common designs. R package version 0.9.12-4.2. https://CRAN.R-project.org/package=BayesFactor

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*(6), e66350. https://doi.org/10.1371/journal.pone.0066350

Nieuwenhuis, S., Slagter, H. A., von Geusau, N. J. A., Heslenfeld, D. J., & Holroyd, C. B. (2005). Knowing good from bad: Differential activation of human cortical areas by positive and negative outcomes. *European Journal of Neuroscience*, *21*(11), 3161–3168. https://doi.org/10.1111/j.1460-9568.2005.04152.x

O'Doherty, J. P., Dayan, P., Friston, K. J., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329–337. https://doi.org/10.1016/S0896-6273(03)00169-7

O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Reward and Decision Making in Corticobasal Ganglia Networks*, *1104*, 35–53. https://doi.org/10.1196/annals.1390.022

Palminteri, S., & Lebreton, M. (2021). Context-dependent outcome encoding in human reinforcement learning. *Current Opinion in Behavioral Sciences*, *41*, 144–151. https://doi.org/10.1016/j.cobeha.2021.06.006

Palminteri, S., Lefebvre, G., Kilford, E. J., & Blakemore, S. J. (2017). Confirmation bias in human reinforcement learning: Evidence from counterfactual feedback processing. *PLoS Computational Biology*, *13*(8), e1005684. https://doi.org/10.1371/journal.pcbi.1005684

Parvaz, M. A., Konova, A. B., Proudfit, G. H., Dunning, J. P., Malaker, P., Moeller, S. J., Maloney, T., Alia-Klein, N., & Goldstein, R. Z. (2015). Impaired neural response to negative prediction errors in cocaine addiction. *Journal of Neuroscience*, *35*(5), 1872–1879. https://doi.org/10.1523/JNEUROSCI.2777-14.2015

Pasek, J. (2021). Weights: Weighting and weighted statistics. R package version 1.0.2. https://CRAN.R-project.org/package=weights

Pedroni, A., Langer, N., Koenig, T., Allemand, M., & Jancke, L. (2011). Electroencephalographic topography measures of experienced utility. *Journal of Neuroscience*, *31*(29), 10474–10480. https://doi.org/10.1523/Jneurosci.5488-10.2011

Peterburs, J., Suchan, B., & Bellebaum, C. (2013). You do the math: Coding of bets and outcomes in a gambling task in the feedback-related negativity and P300 in healthy adults. *PLoS One*, *8*(11), e81262. https://doi.org/10.1371/journal.pone.0081262

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*(5), 656–664. https://doi.org/10.1111/j.1469-8986.2010.01136.x

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

Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzer, P., & Heekeren, H. R. (2010). Temporal dynamics of prediction error processing during reward-based decision making. *NeuroImage*, *53*(1), 221–232. https://doi.org/10.1016/j.neuroimage.2010.05.052

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology*, *52*(4), 449–459. https://doi.org/10.1111/psyp.12370

Rouder, J. N., Speckman, P. L., Sun, D. C., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/Pbr.16.2.225

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. https://doi.org/10.1037/bul0000006

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

Sambrook, T. D., Roser, M., & Goslin, J. (2012). Prospect theory does not describe the feedback-related negativity value function. *Psychophysiology*, *49*(12), 1533–1544. https://doi.org/10.1111/j.1469-8986.2012.01482.x

San Martin, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, *6*, 304. https://doi.org/10.3389/Fnhum.2012.00304

Santesso, D. L., Bogdan, R., Birk, J. L., Goetz, E. L., Holmes, A. J., & Pizzagalli, D. A. (2012). Neural responses to negative feedback are related to negative emotionality in healthy adults. *Social Cognitive and Affective Neuroscience*, *7*(7), 794–803. https://doi.org/10.1093/scan/nsr054

Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & Kuboki, T. L. (2005). Effects of value and reward magnitude on feedback negativity and P300. *Neuroreport*, *16*(4), 407–411. https://doi.org/10.1097/00001756-200503150-00020

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Soder, H. E., & Potts, G. F. (2018). Medial frontal cortex response to unexpected motivationally salient outcomes. *International Journal of Psychophysiology*, *132*, 268–276. https://doi.org/10.1016/j.ijpsycho.2017.11.003

Stewardson, H. J., & Sambrook, T. D. (2020). Evidence for parietal reward prediction errors using great grand average meta-analysis. *International Journal of Psychophysiology*, *152*, 81–86. https://doi.org/10.1016/j.ijpsycho.2020.03.002

Stewardson, H. J., & Sambrook, T. D. (2021). Reward, salience, and Agency in Event-Related Potentials for appetitive and aversive contexts. *Cerebral Cortex*, *31*(11), 5006–5014. https://doi.org/10.1093/cercor/bhab137

Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction. In *Introduction to reinforcement learning*. MIT Press.

Talmi, D., Atkinson, R., & El-Deredy, W. (2013). The feedback-related negativity signals salience prediction errors, not reward prediction errors. *Journal of Neuroscience*, *33*(19), 8264–8269. https://doi.org/10.1523/Jneurosci.5695-12.2013

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, *307*(5715), 1642–1645. https://doi.org/10.1126/science.1105370

Toepel, U., Knebel, J. F., Hudry, J., le Coutre, J., & Murray, M. M. (2009). The brain tracks the energetic value in food images. *NeuroImage*, *44*(3), 967–974. https://doi.org/10.1016/j.neuroimage.2008.10.005

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

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*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Watts, A. T. M., & Bernat, E. M. (2018). Effects of reward context on feedback processing as indexed by time-frequency analysis. *Psychophysiology*, *55*(9), e13195. https://doi.org/10.1111/psyp.13195

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

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

Yeung, N., 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*(5), 535–544. https://doi.org/10.1093/cercor/bhh153

Yu, R., & Zhang, P. (2014). Neural evidence for description dependent reward processing in the framing effect. *Frontiers in Neuroscience*, *8*, 56. https://doi.org/10.3389/fnins.2014.00056

Yu, R. J., & Zhou, X. L. (2006a). Brain potentials associated with outcome expectation and outcome evaluation. *Neuroreport*, *17*(15), 1649–1653. https://doi.org/10.1097/01.wnr.0000236866.39328.1d

Yu, R. J., & Zhou, X. L. (2006b). Brain responses to outcomes of one's own and other's performance in a gambling task. *Neuroreport*, *17*(16), 1747–1751. https://doi.org/10.1097/01.wnr.0000239960.98813.50

Zheng, Y., Li, Q., Wang, K., Wu, H. Y., & Liu, X. (2015). Contextual valence modulates the neural dynamics of risk processing. *Psychophysiology*, *52*(7), 895–904. https://doi.org/10.1111/psyp.12415

Zheng, Y., Li, Q., Zhang, Y. Y., Li, Q., Shen, H. J., Gao, Q. H., & Zhou, S. Y. (2017). Reward processing in gain versus loss context: An ERP study. *Psychophysiology*, *54*(7), 1040–1053. https://doi.org/10.1111/psyp.12855

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Pearson correlation of sample size and raw effect size, i.e., voltage of difference waves. Raw effect sizes were calculated as follows: domain = gain – loss; size = large – small; domain × size = (large gain – large loss) – (small gain – small loss). Raw effect sizes were used since standardised effect sizes were not extractable from grand average waveforms in the original studies. Critical r (one-tailed) = .28, $N = 36$.

Figure S2. Domain × size interaction in single domain gambles and mixed gambles. Interaction formulated as (large gain – large loss) – (small gain – small loss).

Figure S3 Unweighted t-values for canonical effects (counterpart to Figure 3).

Figure S4. F values for canonical effects taken from an omnibus three-way factorial ANOVA (counterpart of t values in Figure S3). Critical value of $F_{1,34} = 4.12$ ($\alpha = .05$, one-tailed).

Figure S5. F values for canonical effects taken from an omnibus four-way factorial ANOVA (Figure S4 plus site factor). Only significant effects involving the site term are shown. Critical value of $F_{1,25} = 4.24$ ($\alpha = .05$, one-tailed).

Figure S6. Relative probability of each encoder with motivational salience excluded.

Figure S7. Gain-specific valence encoding in the reward positivity as shown by increased amplitude of the reward positivity (better outcome - worse outcome) in gain domain than loss domain. Lines show averages weighted by study sample size, shading indicates two weighted standard errors of the mean. Only single domain gamble experiments used ($N = 16$).

Figure S8. Gain-specific valence encoding in the reward positivity operationalised as the amplitude difference of the reward for gain and loss domains, i.e. (better gain outcome – worse gain outcome) – (better loss outcome – worse loss outcome). This is shown separately for experiments where domain is blocked ($N = 9$) and unblocked ($N = 7$). Lines show averages weighted by study sample size, shading indicates two weighted standard errors of the mean.

Figure S9. Size effect (large – small), indexing motivational salience, in experiments using concrete ($N = 9$) and abstract ($N = 7$) stimuli. Lines show averages weighted by study sample size, shading indicates two weighted standard errors of the mean.

Data S1. Supplementary data

How to cite this article: Stewardson, H., & Sambrook, T D. (2023). Valence precedes value in neural encoding of prediction error. *Psychophysiology*, *00*, e14266. https://doi.org/10.1111/psyp.14266