

## RESEARCH ARTICLE

# Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis

Elsa Fouragnan<sup>1,2</sup>  | Chris Retzler<sup>1,3</sup> | Marios G. Philiastides<sup>1</sup> 

<sup>1</sup>Institute of Neuroscience & Psychology, University of Glasgow, Glasgow, United Kingdom

<sup>2</sup>Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom

<sup>3</sup>Department of Behavioural & Social Sciences, University of Huddersfield, Huddersfield, United Kingdom

## Correspondence

Marios G. Philiastides, Institute of Neuroscience and Psychology, University of Glasgow, 58 Hillhead Street, G12 8QB, Glasgow, Lanarkshire, United Kingdom. Email: marios.philiastides@glasgow.ac.uk or

Elsa Fouragnan, Department of Experimental Psychology, University of Oxford, Tinsley Building, Mansfield Road, Oxford, OX1 3SR, UK. Email: elsa.fouragnan@psy.ox.ac.uk

## Funding information

Biotechnology and Biological Sciences Research Council (BBSRC), Grant/Award Number: BB/J015393/1–2 (M.G.P.); Economic and Social Research Council (ESRC), Grant/Award Number: ES/L012995/1 (M.G.P.)

## Abstract

Learning occurs when an outcome differs from expectations, generating a reward prediction error signal (RPE). The RPE signal has been hypothesized to simultaneously embody the valence of an outcome (better or worse than expected) and its surprise (how far from expectations). Nonetheless, growing evidence suggests that separate representations of the two RPE components exist in the human brain. Meta-analyses provide an opportunity to test this hypothesis and directly probe the extent to which the valence and surprise of the error signal are encoded in *separate* or *overlapping* networks. We carried out several meta-analyses on a large set of fMRI studies investigating the neural basis of RPE, locked at decision outcome. We identified two valence learning systems by pooling studies searching for differential neural activity in response to *categorical* positive-versus-negative outcomes. The first valence network (negative > positive) involved areas regulating alertness and switching behaviours such as the midcingulate cortex, the thalamus and the dorsolateral prefrontal cortex whereas the second valence network (positive > negative) encompassed regions of the human reward circuitry such as the ventral striatum and the ventromedial prefrontal cortex. We also found evidence of a largely distinct surprise-encoding network including the anterior cingulate cortex, anterior insula and dorsal striatum. Together with recent animal and electrophysiological evidence this meta-analysis points to a sequential and distributed encoding of different components of the RPE signal, with potentially distinct functional roles.

## KEYWORDS

decision making, learning, meta-analysis, prediction error, reward

## 1 | INTRODUCTION

Effective decision-making depends upon accurate outcome representations associated with potential choices. These representations can be defined through reinforcement learning (RL) (Rescorla & Wagner, 1972; Sutton, 1998), a modelling framework that uses the reward prediction error (RPE), the difference between actual and expected outcomes, as a learning signal to update future outcome expectations. In this framework, RPE is a signed quantity and learning is driven by two separate components of the RPE signal: its *valence* (i.e., the sign of the RPE, representing whether an outcome is better [+] or worse [–] than expected) and its *surprise* (i.e., the modulus of the RPE, representing the degree [high or low] of deviation from expectations). Whereas the valence informs an agent whether to reinforce or extinguish a certain behaviour (Fouragnan, Retzler, Mullinger, & Philiastides, 2015;

Fouragnan, Queirazza, Retzler, Mullinger, & Philiastides, 2017; Frank, Seeberger, & O'reilly, 2004), the surprise component determines the extent to which the strength of association between outcome and expectations needs to be adjusted (Collins & Frank, 2016; Niv et al., 2015; den Ouden, Kok, & de Lange, 2012).

This modelling framework has received considerable attention in neuroscience since the early 90s when animal neurophysiological studies identified dopaminergic neurons in the midbrain, in particular in the ventral tegmental area (VTA), the substantia nigra pars compacta (SNc) and reticulata (SNr), whose tonic response profile appears to simultaneously capture both components of the RPE signal outlined above (Montague, Dayan, & Sejnowski, 1996; Schultz, Apicella, & Ljungberg, 1993; Schultz, Dayan, & Montague, 1997). Specifically, these neurons show anticipatory increase and suppression of their tonic activity in response to positive and negative RPE, respectively. While the anticipatory increase is

proportional to the magnitude of positive RPE, the magnitude of negative RPE is encoded by the duration of the basal tonic suppression.

This discovery was a breakthrough in the field of learning and decision making and has continued to be influential in the field over the past two and a half decades (see Schultz, 2016a; Schultz, 2016b for a review). As a result, this neurophysiological work has strongly motivated human functional magnetic resonance imaging (fMRI) research to identify the corresponding macroscopic blood-oxygen-level-dependent (BOLD) pattern of the signed RPE. This pattern of activity was expected to be such that the strength of the BOLD would proceed from high positive RPEs > low positive RPEs > low negative RPEs > high negative RPEs. More specifically, studies have employed a model-based fMRI approach, whereby different types of reinforcement-learning models are first fitted to subjects' behaviour to yield parametric predictors for signed RPE against which fMRI data are subsequently regressed (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Fouragnan et al., 2013; Gläscher, Daw, Dayan, & O'Doherty, 2010; O'Doherty et al., 2004; O'Doherty, Hampton, & Kim, 2007; Queirazza, Fouragnan, Steele, Cavanagh, & Philiastides, 2017).

These fMRI studies have employed different algorithms to derive the signed RPE, ranging from the simple formulation of the temporal difference learning algorithm to incorporating action learning, notably using the Q-learning and SARSA ('state, action, reward, state and action') algorithms (Schonberg et al., 2010; Seymour, Daw, Dayan, Singer, & Dolan, 2007; Tanaka et al., 2006). According to qualitative reviews of this previous findings (O'Doherty et al., 2007) as well as quantitative, coordinate-based meta-analyses of these studies, the regions correlating with the different formulations of signed RPE have been found to be predominantly subcortical, including the striatum and amygdala, with some cortical regions, such as the ventromedial prefrontal cortex and the cingulate cortex also reported (Bartra, McGuire, & Kable, 2013; Garrison, Erdeniz, & Done, 2013; Liu, Hairston, Schrier, & Fan, 2011). Additionally, substantial effort has been undertaken to identify how different types of outcomes (primary rewards such as food, or secondary rewards such as monetary outcomes) can modulate the signed RPE in the same regions and the extent to which it can be considered a domain-general, common currency signal (Sescousse, Caldú, Segura, & Dreher, 2013).

While using trial-by-trial estimates of signed RPE from reinforcement-learning models has provided an enormously productive framework for understanding learning and decision-making, a growing number of studies have also discussed the complementary role of surprise, namely the unsigned RPE, which can also be estimated at the single-trial level. These include, but are not limited to, the use of trial-by-trial estimates of the modulus of RPE or Bayesian surprise according to Bayesian learning theory (Hayden, Heilbronner, Pearson, & Platt, 2011; Iglesias et al., 2013). Additionally, human electroencephalography (EEG) studies, attempting to offer a temporal account of the cortical dynamics associated with RPE processing, did not find a systematic monotonic response profile consistent with a single RPE representation but instead offered evidence suggestive of separate representations for valence and surprise at the macroscopic level of responses recorded on the scalp. Specifically, multiple recent EEG studies combining

model-based RPE estimates with single-trial analysis of the EEG revealed an early outcome stage reflecting a purely categorical valence signal and a later processing stage reflecting separate representations for valence and surprise (Fouragnan et al., 2015, 2017; Philiastides, Biele, Vavatzanidis, Kazzer, & Heekeren, 2010b). These later valence and surprise signals appeared in spatially distinct but temporally overlapping neural signatures.

These findings suggest that, in addition to the fully monotonic firing pattern of midbrain neurons, there exist individual representations for valence and surprise, potentially subserving different functional roles during reward-based learning (e.g., approach-avoidance behaviour and the speed of learning via varying degrees of attentional engagement, respectively). Here, we conducted an fMRI meta-analysis to more directly explore the possibility that there exist separate neuronal representations encoding valence and surprise promoting reward learning in humans. We discuss the findings of our work in the context of recent reports from animal neurophysiology and human neuroimaging experiments that provide evidence toward a distributed coding of the different facets of the RPE signal (Brischoux, Chakraborty, Brierley, & Ungless, 2009; Fouragnan et al., 2015, 2017; Matsumoto & Hikosaka, 2009).

## 2 | MATERIALS AND METHODS

### 2.1 | Literature search

We selected fMRI studies using the Pubmed database (<http://www.ncbi.nlm.nih.gov/pubmed>) with the following search keywords: '(fMRI OR neuroimaging) AND (prediction error OR reward OR surprise)' along with three initial filters preselecting studies in which participants were human adults of over 19 years of age and excluding reviews. This initial selection resulted in 724 candidates for inclusion to which a further 64 articles were added from existing in-house reference libraries. Note that previous meta-analyses used the terms 'prediction error' or 'reward' but we are the first to include 'surprise' in our systematic search for relevant articles (Bartra et al., 2013; Garrison et al., 2013; Sescousse et al., 2013).

Abstracts from the 788 candidate-articles identified were then evaluated for inclusion in the corpus according to the following criteria. We required studies of healthy human adults, reporting changes in BOLD as a function of three different components of RPE: the categorical valence, surprise and signed RPE, including statistical comparisons either in the form of binary contrasts or continuous parametric analyses. Because the main objective of the present meta-analysis is to examine the neural coding of RPE processing at decision outcome, we also imposed the restriction that fMRI analyses were time-locked to the presentation of outcomes (feedback). We used studies involving outcomes consisting of abstract points, monetary payoffs, consumable liquids and arousing pictures but excluded articles in which outcomes consisted of social feedback. We also required that studies used functional brain imaging and did not use pharmacological interventions and ensured that the reported coordinates were either in Montreal Neurological Institute (MNI) or Talairach space. Finally, we excluded articles

in which results were derived from region of interest (ROI) since our meta-analytic statistical methods assume that foci are randomly distributed in the whole brain under the null hypothesis. After applying these constraints our meta-analysis comprised 102 publications with a total of 2,316 participants, 144 contrasts and 991 activation foci. The number of participants per study ranged from 8 to 66 (median = 24, interquartile range [IQR] = 7).

## 2.2 | Study categorization

The goal of this meta-analysis was to separately categorize studies along the three components of RPE, locked at time of outcome, in order to: (1) identify the extent to which there exist distinct neural representations for valence and surprise and (2) identify whether the neural correlates of the signed RPE simply intersect those of valence and surprise (possibly due to colinearities across these components) or appear as unique clusters of activation reflecting the true combined influence of the two measures.

To group the relevant articles according to the three main RPE components we used the following definitions: (1) *valence* represents the sign of the RPE and as such it is positive when an outcome is better than expected and negative when worse than expected, (2) *surprise* represents the absolute degree of deviation from expectations and is treated as an unsigned quantity and (3) *signed RPE* simultaneously reflects the influence of both valence and surprise and appears as a fully signed parametric signal. According to these definitions, we identified several fMRI statistical analyses conducted in the original studies that fall under each of the three RPE components (Table 1). The main assumptions of these fMRI analyses, with regard to the BOLD signal as a function of each RPE component, are presented schematically in Figure 1.

For the valence components, the literature has looked at neural responses which vary categorically along positive-negative axes, as represented in patterns A (i) and (ii) of Figure 1. We therefore extracted activations exhibiting a relative BOLD signal increase for negative relative to positive outcomes (NEG > POS: pattern A [i]) and greater BOLD for positive relative to negative outcomes (POS > NEG: pattern A [ii]), respectively. We considered six types of fMRI statistical comparisons which reported coordinate results from either: (1) a contrast associated with negative > positive outcomes, (2) a contrast associated with negative > no outcomes, (3) a negative correlation with a trial-by-trial regressor modulated by [+1] for positive outcomes and [-1] for negative outcomes, (4) the positive correlation with the regressor described in (3), (5) a contrast associated with positive > negative outcomes and (6) a contrast associated with positive > no outcomes. We grouped results from contrasts 1–3 (i.e., NEG > POS) and contrasts 3–6 (i.e., POS > NEG) to capture regions yielding greater BOLD activity for negative relative to positive outcomes and a greater activity for positive relative to negative outcomes, respectively (Table 1).

While the fMRI literature on RPE processing has produced a large amount of theoretical and empirical evidence for the valence and the signed RPE components, comparatively little has been done to directly investigate surprise as a separate component. Fewer studies have used

fMRI regressors that were parametrically modulated by trial-to-trial changes in surprise using the unsigned RPE (Fouragnan et al., 2017; Hayden et al., 2011; Iglesias et al., 2013). These studies used the terms 'surprise', 'unsigned RPE' or outcome 'salience' to refer to the mathematical modulus of RPE from computational learning models. In addition to these articles, our literature search has revealed a number of other measures (see below), which are highly correlated with outcome surprise, as defined by learning theory. We therefore used these measures as proxies of surprise to gain insights into the spatial extent of the relevant neural responses and the degree to which they overlap with those associated with valence.

Specifically, a recent line of research has investigated the neural basis of 'Bayesian surprise' or 'volatility', computed as the direct modulus of Bayesian predictive error (Ide, Shenoy, Yu, & Li, 2013; Iglesias et al., 2013; Mathys et al., 2014; O'Reilly et al., 2013) which corresponds to the absolute difference between categorical outcomes and the probabilistic expectation of these outcomes, estimated using Bayesian inference. In the framework of Bayesian learning, the absolute Bayesian RPE plays an important role in learning from rapid changes in behavioural exploration (Courville, Daw, & Touretzky, 2006). Finally, other studies used the term 'associability' which is a parameter in the Pearce-Hall model (Hall & Pearce, 1979; Pearce & Hall, 1980) defined as the degree of divergence between an actual outcome and the original expectation (e.g., the associative strength between a choice and an outcome - we note that in the RL framework, associability can also refer to the learning rate). It is clear from these reports that there is a lack of consistent terminology to refer to unsigned RPE, which emphasizes the need for a more unified framework for studying RPE processing.

To test for consistencies in the neuronal responses across these different reports, and provide initial support for a unified representation of surprise, we grouped fMRI analyses which reported outcome-locked activations resulting from: (1) a positive correlation with a trial-by-trial regressor of the modulus (unsigned) RPE resulting from RL models across both positive and negative outcomes ('surprise' or 'unsigned RPE'), (2) a positive correlation with a trial-by-trial regressor of the unsigned RPE resulting from Bayesian modelling ('Bayesian Surprise' or 'volatility'), (3) a positive correlation with a trial-by-trial regressor of the free parameter of the Pearce-Hall model ('associability' term), (4) a contrast associated with (high positive outcomes and high negative outcomes) > (low positive outcomes and low negative outcomes OR no outcomes), (5) a positive correlation with a parametric regressor of surprising positive RPE alone and (6) a positive correlation with a parametric regressor of surprising negative RPE alone (Table 1). Figure 1 illustrates the hypothesized pattern of BOLD signal predicted by these contrasts (pattern B), exhibiting a V-shaped response profile that is maximal for both highly surprising negative and positive RPEs. Despite possible subtle differences in the definition of these measures we expected that only foci consistently correlating with deviations from reward expectations would be revealed in this analysis.

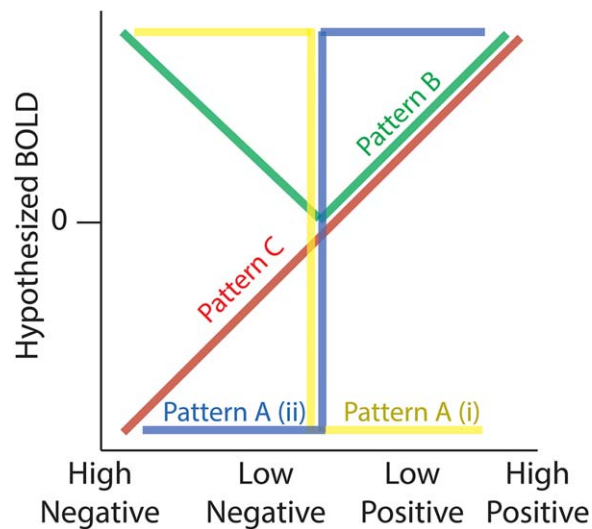
One reason the surprise component has not been looked at closely in isolation is because the literature has focused primarily on signed RPE representations instead. This approach was motivated by

**TABLE 1** Categorisation of fMRI studies into the three RPE components (valence, surprise, signed RPE) and broken down by the relevant fMRI contrast/regressor

Statistical comparisons	Number	Total	Reference
<b>Valence</b>		32	(de Bruijn, de Lange, Cramon, & Ullsperger, 2009; Daniel et al., 2011; Demos, Heatherton, & Kelley, 2012; van Duijvenvoorde et al., 2014; Elward, Vilberg, & Rugg, 2015; Ferdinand & Opitz, 2014; Fouragnan et al., 2015; Gläscher, Hampton, & O'Doherty, 2009; Haruno et al., 2004; Häusler, Oroz Artigas, Trautner, & Weber, 2016; Jocham et al., 2016; Kahnt, Heinze, Park, & Haynes, 2010; Katahira et al., 2015; Klein-Flügge, Hunt, Bach, Dolan, & Behrens, 2011; Klein-Flügge et al., 2011; Knutson, Westdorp, Kaiser, & Hommer, 2000; Knutson, Adams, Fong, & Hommer, 2001; Koch et al., 2008; Leknes, Lee, Berna, Andersson, & Tracey, 2011; Losecaat Vermeer, Boksem, & G Sanfey, 2014; Marsh et al., 2010; Mattfeld, Gluck, & Stark, 2011; Noonan, Mars, & Rushworth, 2011; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001; O'Doherty, Critchley, Deichmann, & Dolan, 2003; Rodriguez, 2009; Rolls, Grabenhorst, & Parris, 2008; Scholl et al., 2015; Seymour et al., 2007; Spicer et al., 2007; Spoomaker et al., 2011; Ullsperger & Cramon, 2003; Yacubian et al., 2006)
<b>Pattern A i (NEG&gt;POS)</b>			
Negative > Positive	19		
Negative > No outcomes	9		
Negative correlation with a regressor defining valence RPE (with a binary modulation whereby positive RPE = 1, and negative RPE = -1)	4		
<b>Valence</b>		33	(Amiez, Sallet, Procyk, & Petrides, 2012; Aron et al., 2004; Bickel, Pitcock, Yi, & Angtuaco, 2009; de Bruijn et al., 2009; Canessa et al., 2013; Daniel et al., 2011; van Duijvenvoorde et al., 2014; Elliott, Friston, & Dolan, 2000; Ernst et al., 2004; Forster & Brown, 2011; Fouragnan et al., 2015; Fujiwara, Tobler, Taira, Iijima, & Tsutsui, 2009; Häusler et al., 2016; Hester, Barre, Murphy, Silk, & Mattingley, 2008; Hester, Murphy, Brown, & Skilleter, 2010; Jocham et al., 2016; Katahira et al., 2015; Knutson et al., 2000; Knutson et al., 2001; Knutson et al., 2001; Kurniawan, Guitart-Masip, Dayan, & Dolan, 2013; Losecaat Vermeer et al., 2014; Luking, Luby, & Barch, 2014; Paschke et al., 2015; Sarinopoulos et al., 2010; Scholl et al., 2015; Schonberg et al., 2010; Seymour et al., 2007; Späti et al., 2014; Spoomaker et al., 2011; Ullsperger & Cramon, 2003)
<b>Pattern A ii (POS&gt;NEG)</b>			
Positive > Negative	18		
Positive > No outcomes	9		
Positive correlation with a regressor defining valence RPE (with a binary modulation whereby positive RPE = 1, and negative RPE = -1)	6		
<b>Surprise</b>		41	(Allen et al., 2016; Amado et al., 2016; Amiez et al., 2012; Boll, Gamer, Gluth, Finsterbusch, & Büchel, 2013; Browning, Holmes, Murphy, Goodwin, & Harmer, 2010; Chumbley et al., 2014; Daw et al., 2011; Dreher, 2013; Ferdinand & Opitz, 2014; Forster & Brown, 2011; Fouragnan et al., 2015; Fouragnan et al., 2017; Fujiwara et al., 2009; Ide et al., 2013; Iglesias et al., 2013; Jensen et al., 2007; Knutson et al., 2001; Kotz, Dengler, & Wittfoth, 2015; Leong, Radulescu, Daniel, DeWoskin, & Niv, 2017; Losecaat Vermeer et al., 2014; Manza et al., 2016; McClure, Berns, & Montague, 2003; Metereau & Dreher, 2013; Metereau & Dreher, 2015; Meyniel & Dehaene, 2017; Nieuwenhuis, Slagter, von Geusau, Heslenfeld, & Holroyd, 2005; O'Reilly et al., 2013; den Ouden et al., 2012; Poudel, Innes, & Jones, 2013; Rodriguez, 2009; Rohe, Weber, & Fließsbach, 2012; Rohe & Noppeney, 2015; Rohe & Noppeney, 2015; Rolls et al., 2008; Schwartenbeck, FitzGerald, & Dolan, 2016; Silvetti & Verguts, 2012; Tobia, Gläscher, & Sommer, 2016; Watanabe, Sakagami, & Haruno, 2013; Wunderlich et al., 2009; Wunderlich, Symmonds, Bossaerts, & Dolan, 2011; Yacubian et al., 2006; Zalla et al., 2000; Zhang, Mano, Ganesh, Robbins, & Seymour, 2016)
<b>Pattern B</b>			
Unsigned RPE ('RL surprise')	12		
Unsigned Bayesian RPE ('Volatility', 'Bayesian surprise')	13		
Positive and Negative outcomes > No or low outcomes	9		
'Associability' term of the Pearce et Hall model	2		
Parametric changes in magnitude of surprising positive RPE (unsigned)	3		
Parametric changes in magnitude of surprising negative RPE (unsigned)	2		
<b>Signed RPE</b>		38	(Abler, Walter, Erk, Kammerer, & Spitzer, 2006; Behrens et al., 2007; van den Bos, Cohen, Kahnt, & Crone, 2012; Cohen & Ranganath, 2007; Daw et al., 2011; Delgado et al., 2000; Delgado, 2007; Diederen et al., 2017; Diuk, Tsai, Wallis, Botvinick, & Niv, 2013; Dunne, D'Souza, & O'Doherty, 2016; Gläscher et al., 2010; Guo et al., 2016; Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Ide et al., 2013; Katahira et al., 2015; Leong et al., 2017; Li & Zhang, 2006; Lin, Adolphs, & Rangel, 2012; Mattfeld et al., 2011; McClure et al., 2003; Metereau & Dreher, 2013; Metereau & Dreher, 2015; O'Doherty et al., 2003; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Pessiglione et al., 2008; Ribas-Fernandes et al., 2011; Rolls et al., 2008; Schlagenhaut et al., 2013; Schonberg et al., 2010; Scimeca, Katzman, & Badre, 2016; Seymour et al., 2007; Takemura, Samejima, Vogels, Sakagami, & Okuda, 2011; Tanaka et al., 2004; Tanaka et al., 2006; Valentin & O'Doherty, 2009; Watanabe et al., 2013; Wunderlich et al., 2011)
<b>Pattern C</b>			
Signed RPE (from model-free RL models)	16		
Signed RPE (from model-based RL models)	8		
Signed Bayesian RPE	10		
High positive RPEs > low positive RPEs > low negative RPEs > high negative RPEs	4		

neurophysiology experiments showing monotonic responses as a function of both valence and surprise and by a theoretical framework suggesting that learning is driven by a single signed RPE representation. To identify the spatial extent of these representations we also looked at fMRI data reporting positive correlations with signed RPE (negative correlation were discarded). Specifically, we combined four types of

fMRI analyses, which estimated trial-by-trial signed RPE from different computational models. We used fMRI reports from (1) model-free and (2) model-based RL methods. Model-free methods include Markov Chain Monte Carlo and temporal difference methods (Samson, Frank, & Fellous, 2010; Seymour et al., 2007). Model-based methods include dynamic programming and certainty equivalent methods (Daw, Niv, &



**FIGURE 1** Hypothesized profiles for BOLD responses as function of the three RPE components. Pattern A (i and ii) describe the two categorical valence responses (yellow and blue colours indicate (i) responses being greater for negative compared with positive outcomes [NEG > POS] and (ii) responses being greater for positive compared with negative outcomes [POS > NEG]). Pattern B captures surprise effects with greater responses for higher outcome deviations from expectations, independent of the sign (valence) of the RPE. Pattern C shows a monotonically increasing response profile consistent with a signed RPE representation [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Dayan, 2005; Doya, Samejima, Katagiri, & Kawato, 2002). More on these algorithms can be found in the review by (Kaelbling, Littman, & Moore, 1996). We also included continuous parametric analyses using trial-by-trial signed RPE from (3) Bayesian RL framework described above (Iglesias et al., 2013; Mathys et al., 2014; den Ouden et al., 2012). Finally, our analysis for signed RPE also contained one type of parametric analysis that employed fixed RPE values (not estimated from RL models) ranked on a scale such that (4) high positive RPEs > low positive RPEs > low negative RPEs > high negative RPEs (Table 1). Figure 1 illustrates the hypothesized pattern of BOLD signal predicted by these contrasts (pattern C) and it is assumed to increase linearly as a function of signed RPE.

Crucially, we note that an issue requiring closer scrutiny pertains to the difficulty in disambiguating the signed RPE pattern of activity from those associated with valence and surprise. Specifically, pattern C (signed RPE) is generally highly correlated with pattern A (ii), (POS > NEG valence) and in studies in which only positive RPEs are considered, pattern C (signed RPE) and pattern B (surprise) are perfectly correlated. Nonetheless, comparing clusters of activations across the three RPE components could potentially reveal whether or not there exist unique clusters of activations associated with signed RPE.

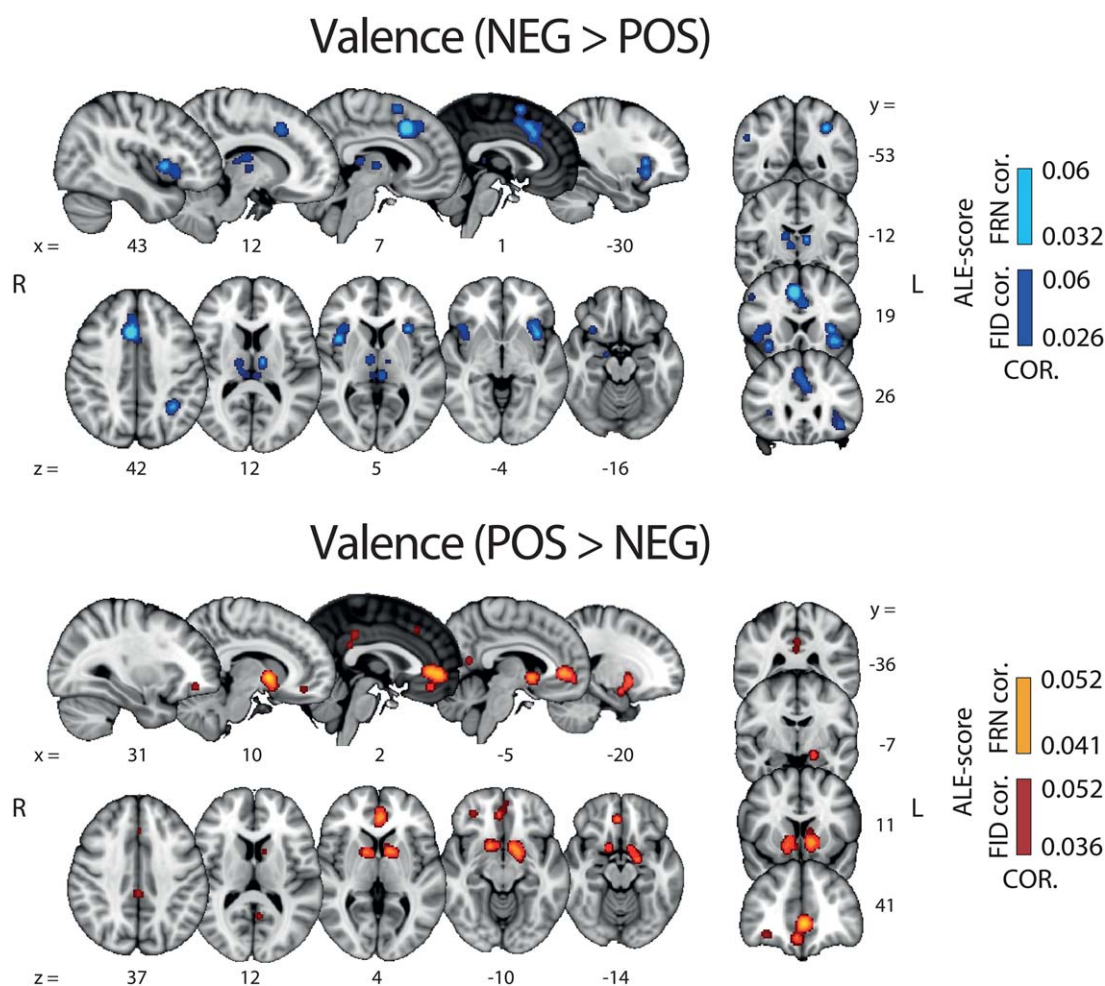
### 2.3 | Activation likelihood estimation analysis

We conducted the meta-analysis using the GingerALE software (version 2.3.6) (Eickhoff et al., 2009) that employs a revised (and rectified;

Eickhoff, Laird, Fox, Lancaster, & Fox, 2017) version of the activation likelihood estimation (ALE) algorithm (Laird et al., 2005; Turkeltaub, Eden, Jones, & Zeffiro, 2002), which identifies common areas of activation across studies. This method performs coordinate based meta-analysis which considers each reported foci as a 3D Gaussian probability distribution, centred at the coordinates provided by each study reflecting the spatial uncertainty associated with each reported set of coordinates. Note that each contrast provided to the ALE algorithm is treated as a separate experiment. The probabilities distributions are then combined to create a modelled activation map, namely an ALE map for that contrast. Studies are weighted according to the number of subjects they contain by adjusting the full width at half maximum of the Gaussian distributions. The convergence of results across the whole brain is obtained by computing the union of all resulting voxel-wise ALE scores. To distinguish meaningful convergence from random noise, statistics are computed by comparing ALE scores with an empirical null-distribution representing a random spatial association between studies. To infer true convergence, a random-effect inference is applied to capitalize on the differences between studies rather than between foci within a particular study. The null-hypothesis is modelled by randomly sampling voxels from each of the ALE maps from which the union is obtained. The ALE maps are assessed against the null distribution using a cluster level threshold of specific *p* values. Contrast analyses between categories of the entire dataset are determined by ALE subtraction method, including a correction for differences in sample size between the categories.

Here, we manually extracted all coordinates from the studies shown in Table 1 and entered them into separate files for each of the three RPE components in preparation for the ALE analyses. Any studies that provided coordinates in Talairach space were converted into MNI space by the Matlab (MathWorks, Natick, Massachusetts) function *tal2mni* in the fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). We conducted ALE analyses for each of the three components of RPE individually. Along the valence component, we looked at both patterns A (i) and A (ii) in Figure 1 (i.e., to identify activations for negative > positive RPE and vice versa, respectively). Accordingly, we ran separate ALE analyses for each of the two patterns. In addition, we performed two conjunction analyses—one between the valence and surprise components to investigate our hypothesis of largely separate neural representations and another between all three RPE components to identify regions that simultaneously encode these representations. Subsequently, we also performed all possible pairwise contrast analyses between the three patterns (A, B and C), using the individual maps associated with each pattern.

A total of 402 foci from 66 contrasts were used with 262 foci from 31 contrasts for Pattern A (i) revealing BOLD patterns greater for negative than positive outcomes and 205 from 35 contrasts for Pattern A (ii) (e.g., the opposite contrast). For the surprise (Pattern B) and signed RPE (Pattern C) analyses, we applied individual ALE analyses, with 284 foci from 40 contrasts for surprise and 240 foci from 38 contrasts for signed RPE. Overall, the number of contrasts used for each separate outcome component was large enough (>30) to allow sufficient power for the required statistical tests (Eickhoff & Etkin, 2016).



**FIGURE 2** Results of whole-brain ALE analysis along the valence component. Overlays of brain areas activated by correlations with NEG > POS (blue) and POS > NEG (orange) (Pattern A [i] and [ii], respectively; Figure 1) ( $p$  values corrected with FDR-ID [FID] and FDR-pN [FRN] < .05 and a minimum cluster volume of 50 mm<sup>3</sup>). Representative slices are shown with MNI coordinates given below each image [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Finally, we transformed the resulting ALE maps from the Colins MNI individual brain space (Colin27\_T1\_seg\_MNI) to the MNI normalized brain space (MNI ICBM152 template) by applying an affine transformation using the FSL *flirt* program (Jenkinson, Bannister, Brady, & Smith, 2002), prior to overlaying onto the canonical MNI template for visualization.

### 3 | RESULTS

All coordinates used for the following ALE analyses were collated from fMRI studies in which the components of RPE have been regressed onto BOLD activity time-locked to outcome presentation. We report ALE maps with clusters surviving the False Discovery Rate (FDR) yielding two  $p$  value thresholds. The most conservative FDR correction yields a  $p$  value with no assumptions about how the data is correlated (FRN), and the least conservative FDR correction assumes independence or positive dependence (FID) with  $p$  < .05 and a minimum volume clustering value of 50 mm<sup>3</sup>. Note that, using a cluster-level family-wise error (FWE) correction implemented with a cluster-extent threshold of

$p$  < .05 and a cluster-forming threshold of  $p$  < .001 revealed virtually identical results (compared with FRN) (Eickhoff et al., 2017) as per previous reports (Garrison, Done, & Simons, 2017). For all tables presenting ALE cluster results, the size of each cluster is provided in mm<sup>3</sup> along with the associated MNI coordinates and maximum ALE score. The ALE score indicates the relative effect size for each peak voxel within each ALE analysis.

#### 3.1 | Outcome valence

The first two ALE analyses were conducted to identify regions in which BOLD signals correlate with outcome valence. Specifically, we looked at activations that yielded greater BOLD for negative relative to positive outcomes (NEG > POS; pattern A [i] in Figure 1) and greater BOLD for negative relative to positive outcomes (POS > NEG; pattern A [ii] in Figure 1), respectively. Accordingly, we considered all fMRI studies, which assumed BOLD responses varying categorically along a positive-negative axis for outcome valence.

The findings of the two valence ALE analyses are shown in Figure 2. The resulting maps revealed a highly distributed network of brain

**TABLE 2** ALE cluster results for the valence analysis: Pattern A (i) and (ii) (FDR-ID  $p < .05$ , with a minimum volume cluster size of 50 mm<sup>3</sup>)

Region	R/L	x	y	z	Cluster size	ALE score
<b>Pattern A (i) NEG &gt; POS</b>						
Dorsomedial cingulate cortex (dMCC)	R	2	24	36	12,712	0.051
Anterior insula (aINS)	R	32	24	-2	6,120	0.062
-	L	-32	22	-4	4,880	0.056
Pallidum	R	12	8	4	3,360	0.04
-	L	-14	6	2	2,520	0.029
Middle frontal gyrus	R	38	4	32	3,152	0.029
-	R	30	10	56	488	0.021
-	L	-28	12	60	104	0.019
Inferior parietal lobule (IPL)	R	40	-48	42	2,416	0.039
-	L	-38	-48	42	2,216	0.043
Middle temporal gyrus (MTG)	R	60	-28	-6	1,192	0.031
Amygdala	R	18	-6	-12	704	0.024
Thalamus	L	-12	-12	10	624	0.025
-	L	-6	-26	8	280	0.023
Habenula	R	2	-20	-18	312	0.022
Dorsolateral prefrontal cortex (dlPFC)	L	-44	28	32	360	0.020
-	R	40	34	30	344	0.020
Fusiform area	L	-40	-62	-10	272	0.023
Precentral cortex	L	-52	0	34	256	0.021
Dorsomedial orbitofrontal cortex (dmOFC)	R	38	58	-2	192	0.020
Dorsomedial prefrontal cortex (dmPFC)	R	20	50	4	120	0.018
Superior temporal sulcus	R	58	-42	22	120	0.017
<b>Pattern A (ii) (POS &gt; NEG)</b>						
Ventral striatum (vSTR)	L	-12	8	-4	4,880	0.052
-	R	8	8	-2	2,880	0.038
Ventromedial prefrontal cortex (vmPFC)	L	-2	42	0	3,416	0.037
Posterior cingulate cortex (PCC)	L	0	-32	36	240	0.016
-	L	0	-36	26	88	0.014
Ventrolateral OFC (vIOFC)	R	32	44	-10	144	0.015
Dorsomedial prefrontal cortex (dmPFC)	L	-6	-56	14	96	0.016
Medial prefrontal cortex (mPFC)	L	-2	46	20	88	0.014

activations encompassing several cortical regions and sub-cortical structures. More precisely, NEG > POS valence clusters were found in a network encompassing the anterior and dorsal part of the mid-cingulate cortex (aMCC and dMCC) including the pre supplementary motor area (pre-SMA), the bilateral anterior and middle insular cortex (aINS, mINS), the bilateral dorsolateral prefrontal cortex (dlPFC), the bilateral thalamus, right amygdala, left inferior parietal lobule (IPL) and the habenula.

POS > NEG valence clusters were found in the bilateral ventral striatum (vSTR), the ventromedial prefrontal cortex (vmPFC), the posterior part of the cingulate cortex (PCC), as well as the ventrolateral orbitofrontal cortex (vIOFC). At a lower threshold (uncorrected  $p$  value of .001), we also found the midbrain as part of this network, encompassing the VTA, which is commonly associated with the delivery of reward (D'Ardenne, McClure, Nystrom, & Cohen, 2008). Table 2 contains the complete list of regions, coordinates and statistics of these two ALE analyses.

### 3.2 | Surprise

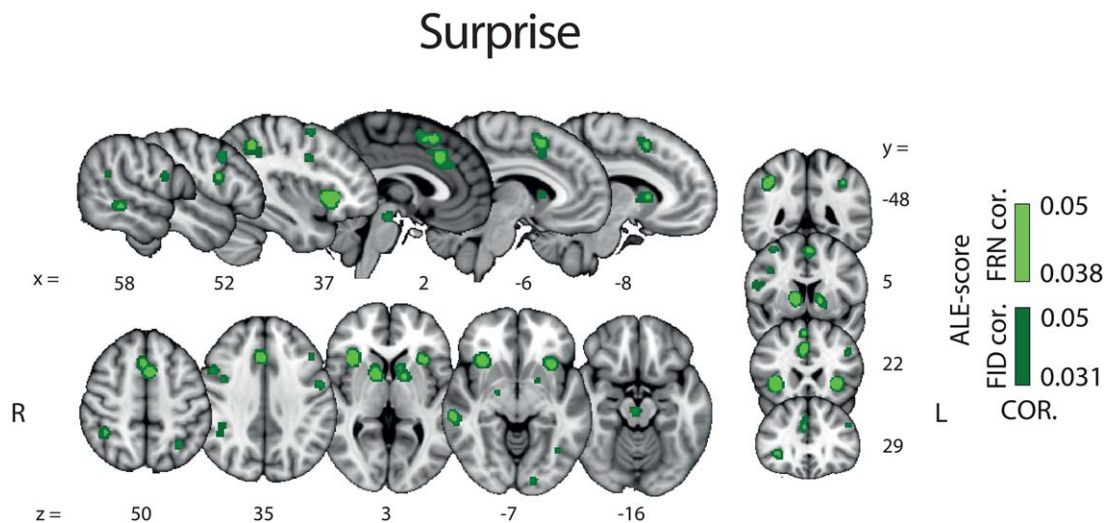
fMRI investigations of RPE have focused primarily on the valence components while neglecting potential contributions from possible separate representations along the surprise component, defined as the degree by which outcomes deviate from expectations and mathematically expressed as the modulus of RPE. A major goal of this work was to explore the possibility that there exist largely separate neuronal representations encoding surprise. To this end, we conducted a new ALE analysis in which the few empirical fMRI studies making use of the surprise from RL models were combined with other fMRI measures correlated with the surprise as defined by RL models (Table 1).

Figure 3 shows the areas in which BOLD signal correlated with surprise. We found evidence for activations in a distributed network encompassing the aMCC, dMCC, the pre-SMA the bilateral dorsal striatum (dSTR), the bilateral aINS, the MTG and the midbrain. Crucially, this activation map shows that the neural network associated with surprise is largely distinct from that of valence. This finding provides initial support for the notion that these two RPE components are encoded in separate brain areas and, as such, they might be contributing individually to promote learning. The full results of the surprise ALE analysis are also summarized in Table 3.

### 3.3 | Valence and surprise conjunction and contrast analyses

The activation maps for valence (NEG > POS and POS > NEG) and surprise ALE analyses conducted above revealed little overlap between the spatial representations of these two RPE components. To formally quantify the degree of overlap between the valence and surprise networks, we next ran a conjunction analysis between the two components. The statistical map resulting from this conjunction analysis and the two separate statistical maps of valence and surprise (as already reported in Figures 2 and 3) are overlaid in Figure 4.

Contrast analyses were conducted for each possible pairing between any dimensions of valence (POS > NEG [positive]; NEG > POS



**FIGURE 3** Results of the whole brain ALE analysis for the surprise component of RPE (pattern B, Figure 1). Overlay of brain areas activated by all analyses representing direct or indirect measures of the surprise component of RPE ( $p$  values corrected with FDR-ID [FID] and FDR-pN [FRN]  $< .05$  and a minimum cluster volume of  $50 \text{ mm}^3$ ). Representative slices are shown with MNI coordinates given below each image [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

**TABLE 3** ALE clusters results for the surprise analysis (FDR-ID  $p < .05$ , with a minimum volume cluster size of  $50 \text{ mm}^3$ )

Region	R/L	x	y	z	Cluster size	ALE score
Anterior mid-cingulate cortex (aMCC)	R	4	24	34	4072	0.029
Anterior insula (aINS)	R	32	24	-4	2496	0.050
-	L	-32	20	-4	1544	0.038
Inferior parietal lobule (IPL)	R	40	-46	42	1672	0.033
-	L	-40	-48	42	568	0.025
Dorsal striatum (dSTR)	R	12	8	4	1400	0.034
-	L	-14	10	2	1216	0.021
Middle temporal gyrus (MTG)	R	60	-28	-8	648	0.022
Lateral inferior frontal cortex	R	52	10	18	488	0.025
Lateral central frontal gyrus	L	-44	26	30	392	0.019
Precentral gyrus	R	48	12	34	360	0.019
-	L	-52	0	34	224	0.020
Midbrain	R	2	-20	-18	304	0.021
Dorsal mid-cingulate cortex (dmCC)	R	12	14	42	224	0.019
Hippocampus	R	20	-6	-10	160	0.018
Fusiform gyrus	L	-40	-60	-10	112	0.017
Mid occipital pole	L	-16	-90	-6	112	0.016
Superior temporal sulcus	R	60	-40	20	64	0.015

[negative] and POS + NEG [all valence]) and surprise. These analyses allowed us to identify the areas that were unique and specific to each individual outcome and RPE-related component. The positive valence (pattern A [ii]) minus surprise (pattern B) contrast revealed two main clusters in the vSTR and vmPFC whereas the reverse contrast revealed a network of clusters including preSMA, aINS and MTG. Contrasting negative valence (pattern A [i]) and surprise also exposed separate networks of areas for each subtraction. Specifically, this contrast revealed a network encompassing the thalamus, the habenula, the right mINS and the dmCC, whereas the reverse contrast showed clusters in the dorsal portion of the STR and the dlPFC. The statistical maps resulting from these contrast analyses are presented in Figure 5.

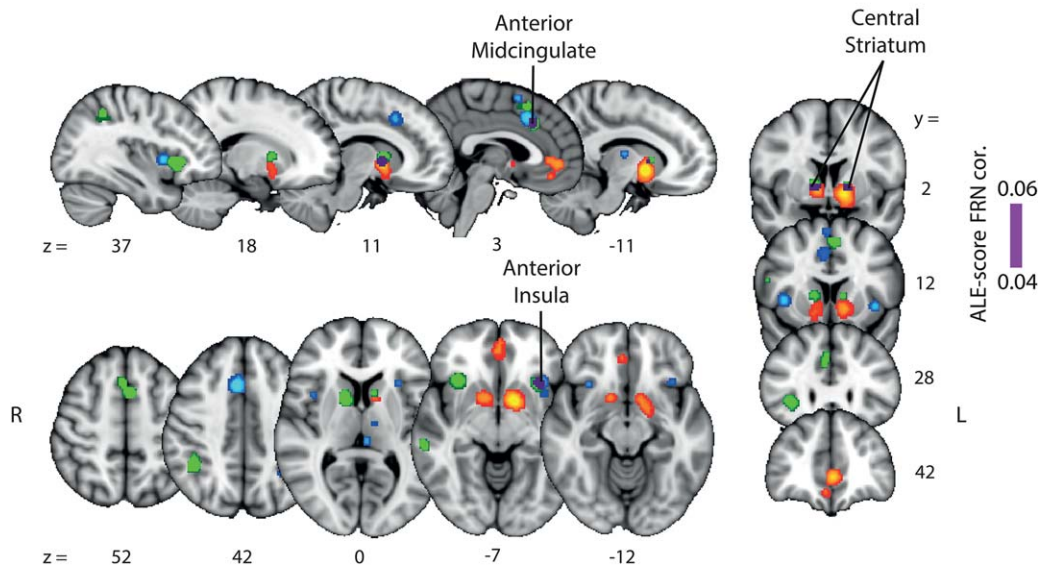
### 3.4 | Signed RPE

A major goal of this work was to investigate the spatial profile of the signed RPE component and to scrutinise more closely the extent to which it overlaps with the separate representations identified for valence (NEG > POS and POS > NEG) and surprise. The fMRI-RPE literature has focused on this component largely due to neurophysiological evidence suggesting that RPE-like learning is driven by a single, theoretically unified representation of both POS > NEG valence and surprise (Table 1).

Results from this ALE analysis revealed very few unique activations for signed RPE compared with valence and surprise. Instead, brain areas identified in this analysis overlapped mostly with areas appearing in the POS > NEG valence component and, to a lesser extent, surprise (Figure 6). Specifically, a large overlap between signed RPE and the POS > NEG valence component was found in the STR and a smaller one in the vmPFC. Similarly, areas appearing in the signed RPE analysis that overlapped with the surprise component were also found, albeit only in small clusters comprising the aMCC and dorsal STR. Taken



## Conjunction Surprise & Valence



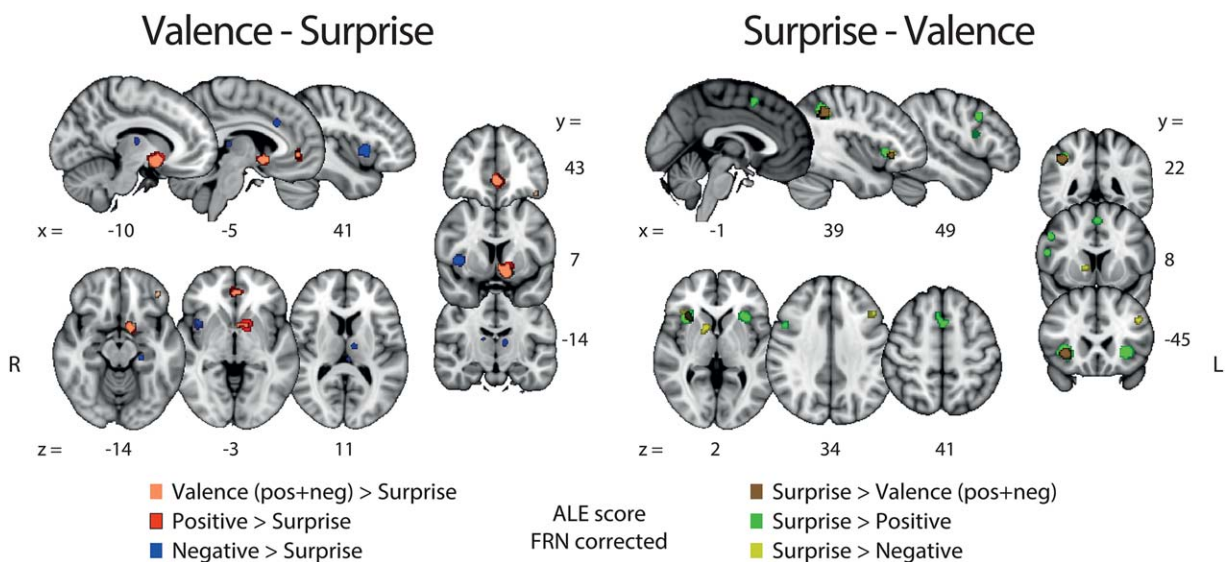
**FIGURE 4** Results of the ALE conjunction analysis between valence and surprise (purple). The regions identified earlier with separate ALE analyses along the valence (NEG > POS: blue, POS > NEG: orange) and surprise (green) components are shown for comparison purposes. The  $p$  values were corrected with FDR-pN [FRN] < .05 and a minimum cluster volume of  $50 \text{ mm}^3$  for the initial maps. Representative slices are shown with MNI coordinates given below each image [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

together, these findings emphasize the potential collinearities between the BOLD predictors used to identify neural representations associated with the three RPE components and highlight the need for developing a methodology for properly disentangling their individual contributions (Tables 4–7).

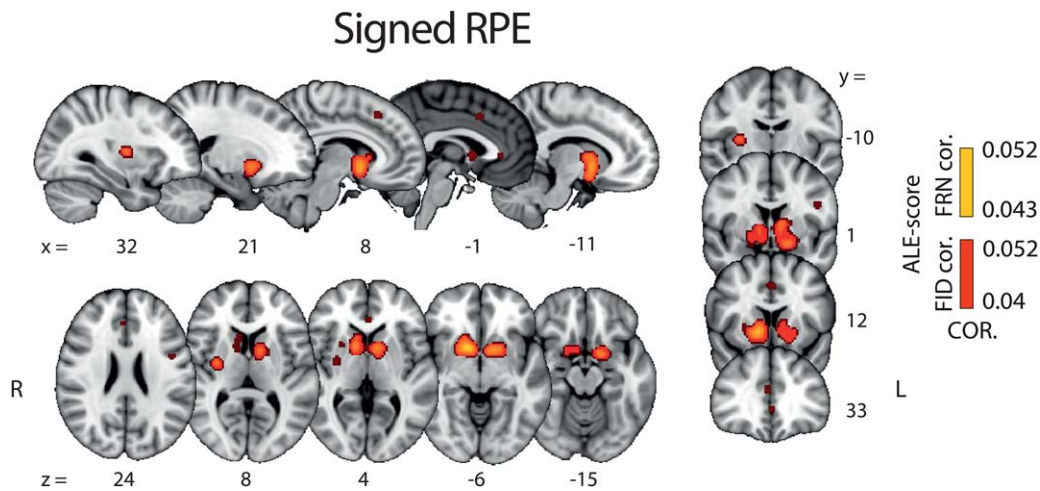
### 3.5 | Putting it all together

Subsequently, to formally test for the overlap between all three RPE components and identify potential regions integrating valence and

surprise either into a signed RPE representation or a linear superposition of the two signals (Fouragnan et al., 2017), we performed a conjunction analysis between the valence (pattern A), the surprise (pattern B) and signed RPE (pattern C) signals. We summarize our conjunction results in Figure 7, which revealed a major overlap between all activations associated with signed RPE and each of the other two RPE representations in the central part of the STR. Thus, one possibility is that the STR meets the requirement that a full monotonic representation of the error signal also simultaneously encodes valence and surprise, as per our last ALE analysis.



**FIGURE 5** Results of the ALE contrast analyses for [valence - surprise] (left panel) and [surprise - valence]. The  $p$  values were corrected with FDR-pN [FRN] < .05 and a minimum cluster volume of  $50 \text{ mm}^3$  for the initial maps. Representative slices are shown with MNI coordinates given below each image [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 6** Results of whole brain ALE analysis for signed RPE. Overlay of brain areas activated by positive correlation with signed RPE ( $p$  values corrected with FDR-ID [FID] and FDR-pN [FRN]  $< .05$  and a minimum cluster volume of  $50 \text{ mm}^3$ ). Representative slices are shown with MNI coordinates given below each image [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Another possibility is that the overlap between all outcome components outcomes in the STR is arising, at least in part, due to collinearities across the different outcome representations, particularly between the positive categorical nature of outcome valence (pattern A [ii]) and the signed RPE. To formally test this hypothesis, we performed a new series of contrast analyses between signed RPE and all dimensions of categorical valence and surprise. Particularly, we performed contrast analyses between patterns C-A(i), C-A(ii), C-A and C-B (and vice versa). The results are summarized in Figure 8. Intriguingly, we did not find any area unique to signed RPE when looking at each of the individual comparisons of signed RPE with the other three patterns. In fact, when comparing signed RPE to positive valence (pattern A [ii]), no clusters were found to be significantly different than those found with the categorical outcome valence (POS > NEG). Conversely, the STR was found for all the other signed RPE comparisons (signed RPE > negative; signed RPE > surprise). Finally, the unique network related to negative valence (pattern A [i]) was found in the dMCC, thalamus and mINS, the unique cluster related to positive valence was found in the vmPFC and the unique network related to surprise was found in the aMCC, pre-SMA and the aINS.

**TABLE 4** ALE cluster results for the conjunction analysis of valence and surprise (FDR-ID  $p < .05$ , with a minimum volume cluster size of  $50 \text{ mm}^3$ )

Region	R/L	x	y	z	Cluster size	ALE score
Striatum (STR)	R	12	6	4	1082	0.031
-	L	-12	12	4	376	0.021
Anterior insula (aINS)	L	-32	20	-6	453	0.018
Anterior mid-cingulate cortex (aMCC)	R	3	22	37	221	0.014
Inferior parietal lobule	L	40	-46	42	327	0.014

## 4 | DISCUSSION

In this fMRI meta-analysis work, we demonstrated that reward learning in humans involves separate neuronal signatures of RPE, comprising distinct representations for valence and surprise. Together with recent neurophysiological and EEG evidence (including studies using simultaneous EEG and fMRI), these findings point to a potentially sequential and distributed encoding of different RPE components with potentially functionally distinct roles.

### 4.1 | Valence networks

The ALE analyses related to valence revealed two distributed set of activations correlating with both pattern A (i) and (ii) in Figure 1. Foci for which the BOLD signal was greater for negative than positive outcomes showed significant clustering in a large network of areas including the thalamus, the aMCC and dMCC, the aINS, mINS and the dIPFC. Conversely, foci for which the BOLD signal was greater for positive than negative outcomes showed significant clusters in a separate network including vmPFC, vSTR, PCC and vIOFC. These findings clearly suggest the presence of multiple systems responding to the categorical nature of valence which supports the notion that separate valuation systems shape learning in the human brain (Fiorillo, 2013; Fouragnan et al., 2013), although their functional role remains debated. More specifically, the debate focuses on the number and exact nature of the neural systems assigning value to decision outcomes and driving behaviours that are evolutionarily appropriate in response to changes in the environment.

A first theory describes two distinct valence systems invoking two orthogonal axes of decision-making: alertness (involving the implementation of action) and learning (including the updates of value expectations for future avoidance and approach behaviours). In this framework, the first system is thought to monitor on-going activity and interrupt it when needed to trigger switching behaviours (e.g., following

**TABLE 5** ALE cluster results for the contrast analyses of valence and surprise (FDR- $p < .05$ , with a minimum volume cluster size of 50 mm<sup>3</sup>)

Region	R/L	x	y	z	Cluster size	ALE score
<b>Valence vs. surprise</b>						
Ventral striatum (vSTR)	L	-10	8	-10	1,096	3.29
Ventromedial prefrontal cortex (vmPFC)	L	-2	44	0	256	3.29
<b>Positive vs. surprise</b>						
Ventral striatum (vSTR)	L	-12	-8	-8	1,872	3.29
Ventromedial prefrontal cortex (vmPFC)	R	0	46	0	512	3.29
Ventral striatum (vSTR)	R	8	8	-6	168	3.29
<b>Negative vs. surprise</b>						
Middle insula (mINS)	R	40	10	2	544	3.29
Mid cingulate Cortex (MCC)	R	6	20	42	144	3.29
<b>Surprise vs. valence</b>						
Anterior insula (aINS)	R	32	24	-4	1,224	3.29
Anterior insula (aINS)	L	-32	20	-2	112	3.29
Ventral tegmental area (VTA)	L	-6	-16	-10	96	3.29
Ventral tegmental area (VTA)	R	2	-20	-16	72	3.29
Occipital lobe	R	24	-80	-6	72	3.29
<b>Surprise vs. positive</b>						
Anterior insula (aINS)	R	32	22	-2	1,648	3.29
Middle temporal gyrus (MTG)	R	40	-46	42	1,184	3.29
Anterior insula (aINS)	L	-32	22	-2	1,016	3.29
Inferior Frontal Gyrus	R	52	10	18	184	3.29
Supplementary motor area (SMA)	L	-2	12	52	160	3.29
<b>Surprise vs. negative</b>						
Angular gyrus	R	40	-46	40	248	3.29
Anterior insula (aINS)	R	32	28	-6	80	3.29
Dorsal striatum (dSTR)	R	12	10	2	56	3.29

negative RPEs). In contrast, the second system uses both negative and positive RPE values for decreasing or increasing internal value representations associated with decisions to ultimately drive avoidance and approach learning, respectively (Boureau & Dayan, 2011; Cools, Nakamura, & Daw, 2011; Elliot, 2006; Fiorillo, 2013; Fouragnan et al., 2015; Gray & McNaughton, 2003; Guitart-Masip et al., 2012).

**TABLE 6** ALE clusters results for the signed RPE studies (FDR-ID  $p < .05$ , with a minimum volume cluster size of 50 mm<sup>3</sup>)

Region	R/L	x	y	z	Cluster size	ALE score
Striatum (STR) (encompasses left and right hemispheres)	R	12	10	-4	10888	0.053
Putamen	R	30	-6	8	688	0.024
Anterior mid-cingulate cortex (aMCC)	R	6	26	46	160	0.018
-	L	-2	14	40	120	0.016
Anterior cingulate cortex (ACC)	R	4	36	20	112	0.017
Ventromedial prefrontal (vmPFC)	L	0	34	0	64	0.015
Lateral inferior frontal gyrus (lIFC)	L	-46	4	24	64	0.016

A second (not mutually exclusive) proposition supports the idea that there are at least two separate systems responsible for aversive and appetitive reinforcements such that punishments and rewards are encoded separately (i.e., a punishment space and a reward space; Morris, 2014). This proposition was developed on the basis of neurophysiological evidence showing that different types of neurons exhibit differential activity in response to punishing vs. non-punishing outcomes and rewarding versus non-rewarding outcomes, respectively (Fiorillo, Tobler, & Schultz, 2003; Fiorillo, 2013; Schultz, Apicella, Scarnati, & Ljungberg, 1992; Schultz, 1998). In this second theory, the punishment space is responsible for avoidance behaviours as well as avoidance learning and the reward space is responsible for approach behaviours and approach learning.

It is noteworthy that our meta-analysis on itself cannot directly distinguish between the two theories because the results do not reveal whether the relevant activations respond exclusively to either positive or negative outcomes or are modulated by both outcomes in opposite directions. This distinction is critical because the former response profile would suggest the presence of separate approach and avoidance systems that might not necessarily be linked to the learning processes as such, while the latter might point to both up- and down-regulation of activity consistent with learning and updating of reward expectations. Nonetheless, the meta-analysis results suggest that two main networks process valence. The network encompassing aINS, aMCC, thalamus and dlPFC could regulate on-going activity and alertness or

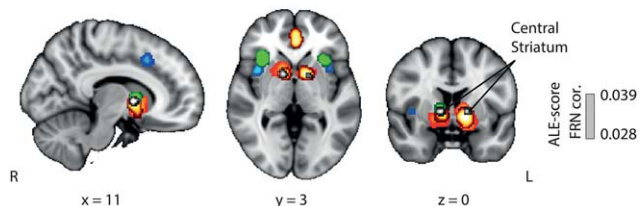
**TABLE 7** ALE cluster results for the contrast analyses of signed RPE and valence as well as signed RPE and surprise (FDR-pN  $p < .05$ , with a minimum volume cluster size of 50 mm<sup>3</sup>)

Region	R/L	x	y	z	Cluster size	ALE score
<b>Positive–signed RPE</b>						
Ventromedial prefrontal cortex (vmPFC)	R	2	44	–15	160	3.29
<b>Signed RPE–positive</b>						
No significant						
<b>Negative–signed RPE</b>						
Middle insula (mINS)	R	40	12	0	528	3.29
Dorsal middle cingulate cortex (dmCC)	R	6	22	36	208	3.29
Middle insula (mINS)	L	–38	18	–4	184	3.29
Habenula	L	–2	–26	8	168	2.58
Thalamus	R	8	–10	5	96	2.58
<b>Signed RPE–negative</b>						
Ventral striatum (vSTR)	R	10	10	–6	2208	3.29
<b>Valence–signed RPE</b>						
Ventromedial prefrontal cortex (vmPFC)	R	2	44	–12	760	3.29
Middle insula (mINS)	R	40	12	2	568	2.58
Dorsal middle cingulate cortex (dmCC)	R	6	24	38	480	2.58
<b>Signed RPE–valence</b>						
Ventral striatum (vSTR)	R	12	16	–2	184	3.29
<b>Surprise–signed RPE</b>						
Anterior insula (aINS)	L	–34	22	0	704	3.29
Anterior midcingulate cortex (amCC)	R	0	14	52	136	3.29
Pre supplementary motor area (preSMA)	R	0	14	52	136	3.29
Anterior insula (aINS)	R	38	18	–2	88	3.29
<b>Signed RPE–surprise</b>						
Ventral striatum (vSTR)	L	–10	8	–10	904	3.29
Ventral striatum (vSTR)	R	12	14	–3	192	3.29
Ventral striatum (vSTR)	R	4	6	–6	72	3.29

could represent the punishment space in accordance to the first and the second theories, respectively. Conversely, the network of regions encompassing the vmPFC, vSTR, PCC and vOFC could represent the learning system depicted in the first theory or could represent the reward space depicted in the second theory. Further research is

required to tease apart the roles of these systems, especially by investigating their precise response profiles in the appetitive (where rewarding and non-rewarding outcomes are manipulated) and in a true aversive (where punishing and non-punishing outcomes are manipulated) domains, respectively.

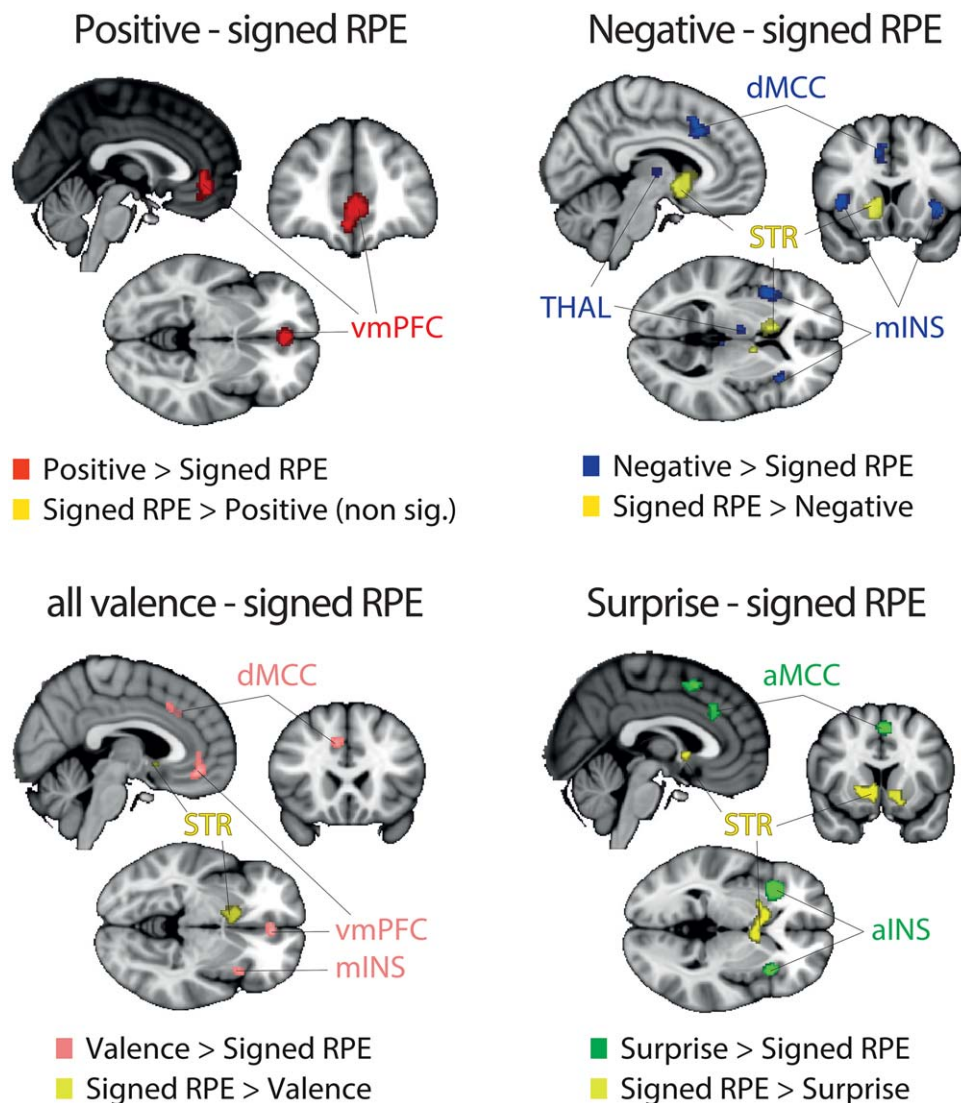
### Conjunction Surprise & Valence & Signed PE



**FIGURE 7** Results of the ALE conjunction analysis for all components of RPE. Overlay of brain areas individually activated by (1) valence (orange), (2) surprise (green), and (3) signed RPE (red), with  $p$  values corrected with FDR-pN [FRN]  $< .05$  and a minimum cluster volume of 50 mm<sup>3</sup> for the initial maps. Importantly, the overlap between the three analyses, shown in white, also corresponds to the only cluster found for the ALE conjunction analysis between valence/surprise vs. signed RPE. MNI coordinates are given below each image [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

## 4.2 | Surprise network

Emerging evidence indicates that the brain encodes the unsigned RPE signal (surprise), which alerts the organism of relative deviations from expectations, regardless of the outcome value. However, to date, only few articles have modelled surprise as such to search for independent neural representations, with the exception of recent neurophysiological developments (Brischoux et al., 2009; Matsumoto & Hikosaka, 2009), recent EEG work (Philiastides et al., 2010b; Yeung & Sanfey, 2004) and an increasing number of fMRI studies (Fouragnan et al., 2017; Gläscher et al., 2010; Li & Daw, 2011; Metereau & Dreher, 2013). Nevertheless, other fMRI studies used variables highly correlated with surprise that can be employed as proxies (Behrens, Woolrich, Walton, & Rushworth, 2007; Iglesias et al., 2013; Nassar et al., 2012; den Ouden et al., 2012; Yu & Dayan, 2005). These studies share the assumption that the corresponding BOLD response profile is maximal for high positive and



**FIGURE 8** Results of the ALE contrast analyses for [signed RPE – positive valence] (left panel), [signed RPE – negative valence] (middle panel) and [signed RPE – (positive + negative valence)] (right panel). The  $p$  values were corrected with FDR- $p$ N [FRN]  $< .05$  and a minimum cluster volume of  $50 \text{ mm}^3$  for the initial maps. Representative slices are shown with MNI coordinates given below each image [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

high negative RPE and minimal for no RPE, resembling a V-shape, as illustrated with pattern B in Figure 1. By combining these fMRI results into a single ALE-analysis, we expose for the first time the network associated with surprise while stressing the need for a common lexicon for this learning component to guide subsequent research in the field.

The surprise ALE-analysis revealed a large network including cortical and sub-cortical areas such as aMCC, bilateral aINS, dSTR and mid-brain, that differed majoritarily from those of valence processing although small overlaps were found between the two components at the junction of ventral and dorsal STR, in left aINS and aMCC. Importantly, the role of surprise is still a subject of debate. Some studies propose that this network encodes the saliency of an outcome or how much a stimulus stands out from others (Litt, Plassmann, Shiv, & Rangel, 2011; Zink, Pagnoni, Martin-Skurski, Chappelow, & Berns, 2004). As such, the surprise system could be considered as a key attentional

mechanism that enables an organism to focus its limited perceptual and cognitive resources on the most pertinent subset of the available sensory data, similarly to the attentional mechanism used to guide decisions in the case of salient stimuli (Kahnt & Tobler, 2013). Consistent with a role in attention regulation, representations of such signal have been found in lower-level visual areas (Serences, 2008), lateral intraparietal cortex (Huettel, Stowe, Gordon, Warner, & Platt, 2006; Kahnt & Tobler, 2013) and areas involved in visual and motor preparation such as the supplementary motor area (Wunderlich, Rangel, & O'Doherty, 2009) or the supplementary eye field (Middlebrooks & Sommer, 2012; So & Stuphorn, 2012).

In contrast, it has also been suggested that a surprise system can independently monitor unexpected information and act as a learning signal that allows better predictions of upcoming events, and help plan appropriate behavioural adjustments (Dayan & Balleine, 2002; Fouragnan et al., 2017; Kolling, Behrens, Mars, & Rushworth, 2012;

Wittmann et al., 2016). In particular, some studies suggest that the aINS receives information related to surprise and direct modulation from the dSTR providing crucial information for behavioural adjustment (Menon & Levitin, 2005). Along these lines, the surprise signal also captures the essence of a learning signal that the brain needs to compute to maintain a homeostatic state (Friston, Kilner, & Harrison, 2006; Friston, 2009). Practically, this means that the brain elaborates internal predictions about sensory input and updates them according to surprise, a process that can be formulated as generalized Bayesian filtering or predictive coding in the brain. Finally, still in the framework of learning, some authors argue that surprise can also be considered as a signal predicting the level of risk associated with a future decision outcome, and thus reflect a risk RPE (Fiorillo et al., 2003; Preusschoff, Quartz, & Bosaerts, 2008; Rudorf, Preusschoff, & Weber, 2012).

### 4.3 | Neuromodulatory pathways encoding multicomponent RPE signals

Supporting the idea of separate neural systems for valence and surprise, recent electrophysiological work has revealed both signals existing in neighbouring groups of neurons. The first study of this kind observed the response of dopaminergic neurons in ventral and dorsal areas of the SNc and reported two categories of dopamine neurons (Matsumoto & Hikosaka, 2009). Some dopamine neurons increase their phasic firing activity in response to valence while others responded only to the changes in unsigned RPE, regardless of the valence component. The latter population of neurons was located more dorsolaterally in the SNc, whilst the neurons encoding valence were located more ventromedially, including the VTA. Interestingly, the dorsolateral SNc projects mainly to the dorsal STR, whereas the ventral SNc and VTA project to the ventral STR, which matches the results of our last conjunction analysis (Figure 7). We found that the only region that encodes the full monotonic representation of the RPE as well as the separate valence and surprise components of RPE seems to be the central part of the STR as shown in Figure 7. This result aligns with the assumption that this region receives direct projections from the midbrain dopaminergic neurons encoding a fully monotonic signed RPE signal (Schultz et al., 1997). Additionally, the meta-analysis also revealed that both the valence (POS > NEG) and surprise networks include activity in the midbrain, confirming this hypothesis.

It is important to note that identifying neural activity associated with valence and surprise signals is challenging because in many experimental paradigms both components are highly correlated. For example, when positive RPE are manipulated in isolation, valence (POS > NEG) strongly correlates with surprise. Additionally, whether positive or negative, an unexpected outcome attracts more attention, leads to higher levels of emotional arousal and involves higher levels of motor preparation compared with no RPE (Matsumoto & Hikosaka, 2009; Maunsell, 2004; Roesch & Olson, 2004). Consequently, to disentangle these signals, one needs to design tasks in which the level of valence and surprise can independently be controlled and decoupled (Kahnt, 2017; Kahnt & Tobler, 2013) or capitalize on the variability of physiologically derived responses (i.e., endogenous variability) associated with valence

and surprise (Fouragnan et al., 2015; Fouragnan et al., 2017; Pisauro, Fouragnan, Retzler, & Philiastides, 2017).

It is important to note that since the problem of collinearity and functional specificity of some brain regions is already present in single studies, it will inevitably be carried over to studies performing conjunction meta-analyses. Virtually every experimental design engages a large number of cognitive operations and, thereby, activates functional neural networks that may be irrelevant to a particular regressor (psychological construct) of interest. For example in our study, regions related to outcome valence and surprise might share variance with outcome confidence (Gherman & Philiastides, 2015; Gherman & Philiastides, 2017; Lebreton, Abitbol, Daunizeau, & Pessiglione, 2015; Philiastides, Heekeren, & Sajda, 2014). Despite this general limitation and the difficulty of interpreting conjunction results, aggregating results across a large number of experiments allows one to expose convergence of findings across studies and increasing the generalizability of the conclusions. In particular, this meta-analysis, capitalizing on both individual maps of activations as well as contrasts between different outcome components, points to a distributed encoding of valence and surprise, with potentially distinct functional roles.

### 4.4 | Temporally specific components of RPE processing

The presence of separate RPE-related neural systems raises the question of how these systems unfold in time. Capitalizing on the high temporal resolution of EEG, three recent studies using simultaneous EEG-fMRI have started to shed light on the spatiotemporal characterisation of the RPE components. First, these studies have revealed two temporally specific EEG components discriminating between positive and negative RPEs peaking around 220 and 300 ms, respectively, largely consistent with the timing of the feedback-related negativity and feedback-related positivity ERP components (Cohen, Elger, & Ranganath, 2007; Hajcak, Moser, Holroyd, & Simons, 2006; Yeung & Sanfey, 2004). Additionally, the studies also revealed a late unsigned RPE component which overlaps temporally with the late valence signal (Philiastides et al., 2010b) but appears in a largely separate and distributed neural network (Fouragnan et al., 2017).

Based on these previous studies and the current meta-analysis, we propose that the early and late EEG valence components might reflect the separate contributions of the two networks of areas found for the ALE-valence analyses. This proposal assumes that an early network processes mainly negative RPEs in order to initiate a fast alertness response in the presence of negative outcomes. Conversely, a later network—associated with the brain's reward circuitry—is modulated by both positive and negative RPEs, consistent with a role in approach/avoidance learning and value updating (Philiastides, Biele, & Heekeren, 2010a). We also propose that the surprise network unfolds near simultaneously with the late valence component and thus influences learning through largely distinct spatial representations of the two outcomes signals, which happen to form a composite signal in overlapping areas (Fouragnan et al., 2017).

## 4.5 | Full representation of a monotonic signed RPE signal

To examine the spatial profile of a true monotonic signed RPE representation in the human brain, we pooled results from fMRI studies, which hypothesized that RPE-like learning is driven by a simultaneous representation of both categorical valence and surprise. These fMRI studies are based on the influential assumption that BOLD signal increases monotonically as a function of signed RPE, as illustrated in pattern C (Figure 1), equivalent to the teaching signal that is predicted in the Rescorla–Wagner model of RL (Rescorla & Wagner, 1972). Additionally, we combined the valence and surprise networks and subsequently compared it with the signed RPE to test the requirement that the signed RPE simultaneously encodes both components. This conjunction analysis revealed that the only brain region that seems to encode a true monotonic signal is the STR in the basal ganglia, which could explain why such a signal is not tractable with EEG recordings as highlighted earlier. This result confirms the long standing view that the BOLD activity in STR mirrors the dopaminergic signalling of the mesolimbic neurons (Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; Haber, Kunishio, Mizobuchi, & Lynd-Balta, 1995; O'Doherty et al., 2004; Pagnoni, Zink, Montague, & Berns, 2002) that fully encode the RL prediction error signal of the Rescorla–Wagner rule (Ikemoto, 2007; Schultz et al., 1992).

Nonetheless, the ALE contrast analyses between valence (the positive correlation with pattern A [ii]) and signed RPE revealed no significant activation, whereas the reverse contrast revealed a denser cluster of activity in vmPFC for valence than signed RPE. Given the evidence presented above that the signed RPE may only be encoded in the STR, we suggest that this result may arise due to collinearities between valence and signed RPE or surprise and signed RPE. More precisely, a parametric predictor for signed RPE would be positively correlated with the contrast positive > negative outcomes whereas the signed RPE and surprise would be perfectly correlated in the positive (appetitive) domain.

## 5 | CONCLUSION

In conclusion, the current meta-analysis points to a framework whereby heterogeneous signals are involved in RPE processing. The proposal of a temporally distinct and spatially distributed representation of valence and surprise is open to debate and many questions remain about how these signals interact and how they correspond to the computations made in the brain. For example, it is currently unclear whether valence and surprise encoding occur before the computation of the signed RPE, or whether these three computations are performed in parallel. Nevertheless the taxonomy proposed is conceptually useful because it breaks down the learning and valuation processes into testable components and organizes the RPE literature in terms of the computations that are potentially involved. It will require additional experiments to validate the current proposal and to better understand the complexity of RPE processing.

## ACKNOWLEDGMENTS

We thank Dr. Andrea Pisauro, Dr. Miriam Klein-Flugge and Dr. Jacqueline Scholl for helpful comments on the manuscript.

## ORCID

Elsa Fouragnan  <https://orcid.org/0000-0003-1485-0332>

Marios G. Philiastides  <http://orcid.org/0000-0002-7683-3506>

## REFERENCES

- Abler, B., Walter, H., Erk, S., Kammerer, H., & Spitzer, M. (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *NeuroImage*, 31(2), 790–795.
- Allen, M., Fardo, F., Dietz, M. J., Hillebrandt, H., Friston, K. J., Rees, G., & Roepstorff, A. (2016). Anterior insula coordinates hierarchical processing of tactile mismatch responses. *NeuroImage*, 127, 34–43.
- Amado, C., Hermann, P., Kovács, P., Grotheer, M., Vidnyánszky, Z., & Kovács, G. (2016). The contribution of surprise to the prediction based modulation of fMRI responses. *Neuropsychologia*, 84, 105–112.
- Amiez, C., Sallet, J., Procyk, E., & Petrides, M. (2012). Modulation of feedback related activity in the rostral anterior cingulate cortex during trial and error exploration. *NeuroImage*, 63(3), 1078–1090.
- Aron, A. R., Shohamy, D., Clark, J., Myers, C., Gluck, M. A., & Poldrack, R. A. (2004). Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *Journal of Neurophysiology*, 92(2), 1144–1152.
- 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.
- Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9), 1214–1221.
- Bickel, W. K., Pitcock, J. A., Yi, R., & Angtuaco, E. J. C. (2009). Congruence of BOLD response across intertemporal choice conditions: Fictive and real money gains and losses. *Journal of Neuroscience*, 29(27), 8839–8846.
- Boll, S., Gamer, M., Gluth, S., Finsterbusch, J., & Büchel, C. (2013). Separate amygdala subregions signal surprise and predictiveness during associative fear learning in humans. *The European Journal of Neuroscience*, 37(5), 758–767.
- van den Bos, W., Cohen, M. X., Kahnt, T., & Crone, E. A. (2012). Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cerebral Cortex*, 22(6), 1247–1255.
- Boureau, Y.-L., & Dayan, P. (2011). Opponency revisited: Competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 36(1), 74–97.
- Brischoux, F., Chakraborty, S., Brierley, D. I., & Ungless, M. A. (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, 106(12), 4894–4899.
- Browning, M., Holmes, E. A., Murphy, S. E., Goodwin, G. M., & Harmer, C. J. (2010). Lateral prefrontal cortex mediates the cognitive modification of attentional bias. *Biological Psychiatry*, 67(10), 919–925.
- de Bruijn, E., de Lange, F., Cramon, D., & Ullsperger, M. (2009). When errors are rewarding. *The Journal of Neuroscience*, 29(39), 12183–12186.
- Canessa, N., Crespi, C., Motterlini, M., Baud-Bovy, G., Chierchia, G., Panzaleo, G., ... Cappa, S. F. (2013). The functional and structural neural

- basis of individual differences in loss aversion. *The Journal of Neuroscience*, 33(36), 14307–14317.
- Chumbley, J. R., Burke, C. J., Stephan, K. E., Friston, K. J., Tobler, P. N., & Fehr, E. (2014). Surprise beyond prediction error. *Human Brain Mapping*, 35(9), 4805–4814.
- Cohen, M. X., & Ranganath, C. (2007). Reinforcement learning signals predict future decisions. *The Journal of Neuroscience*, 27(2), 371–378.
- Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage*, 35(2), 968–978.
- Collins, A. G. E., & Frank, M. J. (2016). Surprise! Dopamine signals mix action, value and error. *Nature Neuroscience*, 19(1), 3–5.
- Cools, R., Nakamura, K., & Daw, N. D. (2011). Serotonin and dopamine: Unifying affective, activational, and decision functions. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 36(1), 98–113.
- Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, 10(7), 294–300.
- Daniel, R., Wagner, G., Koch, K., Reichenbach, J. R., Sauer, H., & Schlösser, R. G. M. (2011). Assessing the neural basis of uncertainty in perceptual category learning through varying levels of distortion. *Journal of Cognitive Neuroscience*, 23(7), 1781–1793.
- D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, 319, 1264–1267.
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69(6), 1204–1215.
- Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8(12), 1704–1711.
- Dayan, P., & Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron*, 36(2), 285–298.
- Delgado, M. R. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Sciences*, 1104, 70–88.
- Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, 84(6), 3072–3077.
- Demos, K. E., Heatherton, T. F., & Kelley, W. M. (2012). Individual differences in nucleus accumbens activity to food and sexual images predict weight gain and sexual behavior. *The Journal of Neuroscience*, 32(16), 5549–5552.
- Diederer, K. M. J., Ziauddeen, H., Vestergaard, M. D., Spencer, T., Schultz, W., & Fletcher, P. C. (2017). Dopamine modulates adaptive prediction error coding in the human midbrain and striatum. *The Journal of Neuroscience*, 37(7), 1708–1720.
- Diuk, C., Tsai, K., Wallis, J., Botvinick, M., & Niv, Y. (2013). Hierarchical learning induces two simultaneous, but separable, prediction errors in human basal ganglia. *The Journal of Neuroscience*, 33(13), 5797–5805.
- Doya, K., Samejima, K., Katagiri, K., & Kawato, M. (2002). Multiple model-based reinforcement learning. *Neural Computation*, 14(6), 1347–1369.
- Dreher, J.-C. (2013). Neural coding of computational factors affecting decision making. *Progress in Brain Research*, 202, 289–320.
- van Duijvenvoorde, A. C. K., Op de Macks, Z. A., Overgaauw, S., Gunther Moor, B., Dahl, R. E., & Crone, E. A. (2014). A cross-sectional and longitudinal analysis of reward-related brain activation: Effects of age, pubertal stage, and reward sensitivity. *Brain and Cognition*, 89, 3–14.
- Dunne, S., D'Souza, A., & O'Doherty, J. P. (2016). The involvement of model-based but not model-free learning signals during observational reward learning in the absence of choice. *Journal of Neurophysiology*, 115(6), 3195–3203.
- Eickhoff, S. B., & Etkin, A. (2016). Going beyond finding the 'lesion': A path for maturation of neuroimaging. *The American Journal of Psychiatry*, 173(3), 302–303.
- Eickhoff, S. B., Laird, A. R., Fox, P. M., Lancaster, J. L., & Fox, P. T. (2017). Implementation errors in the GingerALE Software: Description and recommendations. *Human Brain Mapping*, 38(1), 7–11.
- Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping*, 30(9), 2907–2926.
- Elliot, A. J. (2006). The hierarchical model of approach-avoidance motivation. *Motivation and Emotion*, 30(2), 111–116.
- Elliott, R., Friston, K. J., & Dolan, R. J. (2000). Dissociable neural responses in human reward systems. *The Journal of Neuroscience*, 20(16), 6159–6165.
- Elward, R. L., Vilberg, K. L., & Rugg, M. D. (2015). Motivated memories: Effects of reward and recollection in the core recollection network and beyond. *Cerebral Cortex (New York, N.Y. : 1991)*, 25(9), 3159–3166.
- Ernst, M., Nelson, E. E., McClure, E. B., Monk, C. S., Munson, S., Eshel, N., ... Pine, D. S. (2004). Choice selection and reward anticipation: An fMRI study. *Neuropsychologia*, 42(12), 1585–1597.
- Ferdinand, N. K., & Opitz, B. (2014). Different aspects of performance feedback engage different brain areas: Disentangling valence and expectancy in feedback processing. *Scientific Reports*, 4, 5986.
- Fiorillo, C. D. (2013). Two dimensions of value: Dopamine neurons represent reward but not aversiveness. *Science*, 341(6145), 546–549.
- Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science (New York, N.Y.)*, 299(5614), 1898–1902.
- Forster, S. E., & Brown, J. W. (2011). Medial prefrontal cortex predicts and evaluates the timing of action outcomes. *NeuroImage*, 55(1), 253–265.
- Fouragnan, E., Chierchia, G., Greiner, S., Neveu, R., Avesani, P., & Coricelli, G. (2013). Reputational priors magnify striatal responses to violations of trust. *Journal of Neuroscience*, 33(8), 3602–3611.
- Fouragnan, E., Queirazza, F., Retzler, C., Mullinger, K., & Philiastides, M. (2017). Spatiotemporal neural characterization of prediction error valence and surprise during reward learning in humans. *Scientific Reports*, 7(1), 4762.
- Fouragnan, E., Retzler, C., Mullinger, K., & Philiastides, M. G. (2015). Two spatiotemporally distinct value systems shape reward-based learning in the human brain. *Nature Communications*, 6, 8107.
- Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science (New York, N.Y.)*, 306(5703), 1940–1943.
- Friston, K. (2009). The free-energy principle: A rough guide to the brain?. *Trends in Cognitive Sciences*, 13(7), 293–301.
- Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology, Paris*, 100(1–3), 70–87.
- Fujiwara, J., Tobler, P. N., Taira, M., Iijima, T., & Tsutsui, K.-I. (2009). Segregated and integrated coding of reward and punishment in the cingulate cortex. *Journal of Neurophysiology*, 101(6), 3284–3293.
- Garrison, R., Done, J., & Simons, S. (2017). Interpretation of published meta-analytical studies affected by implementation errors in the



- GingerALE software. *Neuroscience & Biobehavioral Reviews*, pii, S0149-7634(16)30764-3.
- Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, 37(7), 1297–1310.
- Gherman, S., & Philiastides, M. G. (2015). Neural representations of confidence emerge from the process of decision formation during perceptual choices. *NeuroImage*, 106, 134–143.
- Gherman, S., & Philiastides, M. G. (2017). Human VMPFC encodes early signatures of confidence in perceptual decisions. bioRxiv, 224337.
- Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66(4), 585–595.
- Gläscher, J., Hampton, A. N., & O'Doherty, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cerebral Cortex*, 19(2), 483–495.
- Gray, J. A., & McNaughton, N. (2003). *The neuropsychology of anxiety: An enquiry into the function of the septo-hippocampal system*. New York, NY: Oxford University Press.
- Guitart-Masip, M., Huys, Q. J. M., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *NeuroImage*, 62(1), 154–166.
- Guo, R., Böhmer, W., Hebart, M., Chien, S., Sommer, T., Obermayer, K., & Gläscher, J. (2016). Interaction of instrumental and goal-directed learning modulates prediction error representations in the ventral striatum. *The Journal of Neuroscience*, 36(50), 12650–12660.
- Haber, S. N., Kunishio, K., Mizobuchi, M., & Lynd-Balta, E. (1995). The orbital and medial prefrontal circuit through the primate basal ganglia. *The Journal of Neuroscience*, 15, 4851–4867.
- 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.
- Hall, G., & Pearce, J. M. (1979). Latent inhibition of a CS during CS-US pairings. *Journal of Experimental Psychology: Animal Behavior Processes*, 5(1), 31–42.
- Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The Journal of Neuroscience*, 28(22), 5623–5630.
- Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., Samejima, K., ... Kawato, M. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: A functional magnetic resonance imaging study of a stochastic decision task. *The Journal of Neuroscience*, 24(7), 1660–1665.
- Häusler, A. N., Oroz Artigas, S., Trautner, P., & Weber, B. (2016). Gain- and loss-related brain activation are associated with information search differences in risky gambles: An fMRI and eye-tracking study. *eNeuro*, 3, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5032244/>.
- Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., & Platt, M. L. (2011). Surprise signals in anterior cingulate cortex: Neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *The Journal of Neuroscience*, 31(11), 4178–4187.
- Hester, R., Barre, N., Murphy, K., Silk, T. J., & Mattingley, J. B. (2008). Human medial frontal cortex activity predicts learning from errors. *Cerebral Cortex (New York, N.Y. : 1991)*, 18(8), 1933–1940.
- Hester, R., Murphy, K., Brown, F. L., & Skilleter, A. J. (2010). Punishing an error improves learning: The influence of punishment magnitude on error-related neural activity and subsequent learning. *The Journal of Neuroscience*, 30(46), 15600–15607.
- Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., & Platt, M. L. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron*, 49(5), 765–775.
- Ide, J. S., Shenoy, P., Yu, A. J., & Li, C. R. (2013). Bayesian prediction and evaluation in the anterior cingulate cortex. *The Journal of Neuroscience*, 33(5), 2039–2047.
- Iglesias, S., Mathys, C., Brodersen, K. H., Kasper, L., Piccirelli, M., den Ouden, H. E. M., & Stephan, K. E. (2013). Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron*, 80(2), 519–530.
- Ikemoto, S. (2007). Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. *Brain Research Reviews*, 56(1), 27–78.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2), 825–841.
- Jensen, J., Smith, A. J., Willeit, M., Crawley, A. P., Mikulis, D. J., Vitcu, I., & Kapur, S. (2007). Separate brain regions code for salience vs. valence during reward prediction in humans. *Human Brain Mapping*, 28(4), 294–302.
- Jocham, G., Brodersen, K. H., Constantinescu, A. O., Kahn, M. C., Ianni, A. M., Walton, M. E., ... Behrens, T. E. J. (2016). Reward-guided learning with and without causal attribution. *Neuron*, 90(1), 177–190.
- Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. arXiv:cs/9605103. <http://arxiv.org/abs/cs/9605103>.
- Kahnt, T. (2017). A decade of decoding reward-related fMRI signals and where we go from here. *NeuroImage*, <http://www.sciencedirect.com/science/article/pii/S1053811917304688>.
- Kahnt, T., & Tobler, P. N. (2013). Salience signals in the right temporoparietal junction facilitate value-based decisions. *The Journal of Neuroscience*, 33(3), 863–869.
- Kahnt, T., Heinzle, J., Park, S. Q., & Haynes, J.-D. (2010). The neural code of reward anticipation in human orbitofrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 107(13), 6010–6015.
- Katahira, K., Matsuda, Y.-T., Fujimura, T., Ueno, K., Asamizuya, T., Suzuki, C., ... Okada, M. (2015). Neural basis of decision making guided by emotional outcomes. *Journal of Neurophysiology*, 113(9), 3056–3068.
- Klein-Flügge, M. C., Hunt, L. T., Bach, D. R., Dolan, R. J., & Behrens, T. E. J. (2011). Dissociable reward and timing signals in human midbrain and ventral striatum. *Neuron*, 72(4), 654–664.
- Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of Neuroscience*, 21, RC159.
- Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage*, 12(1), 20–27.
- Koch, K., Schachtzabel, C., Wagner, G., Reichenbach, J. R., Sauer, H., & Schlösser, R. (2008). The neural correlates of reward-related trial-and-error learning: An fMRI study with a probabilistic learning task. *Learning & Memory (Cold Spring Harbor, N.Y.)*, 15(10), 728–732.
- Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural mechanisms of foraging. *Science (New York, N.Y.)*, 336(6077), 95–98.
- Kotz, S. A., Dengler, R., & Wittfoth, M. (2015). Valence-specific conflict moderation in the dorso-medial PFC and the caudate head in emotional speech. *Social Cognitive and Affective Neuroscience*, 10(2), 165–171.
- Kurniawan, I. T., Guitart-Masip, M., Dayan, P., & Dolan, R. J. (2013). Effort and valuation in the brain: The effects of anticipation and execution. *The Journal of Neuroscience*, 33(14), 6160–6169.

- Laird, A. R., Fox, P. M., Price, C. J., Glahn, D. C., Uecker, A. M., Lancaster, J. L., ... Fox, P. T. (2005). ALE meta-analysis: Controlling the false discovery rate and performing statistical contrasts. *Human Brain Mapping, 25*(1), 155–164.
- Lebreton, M., Abitbol, R., Daunizeau, J., & Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. *Nature Neuroscience, 18*(8), 1159–1167.
- Leknes, S., Lee, M., Berna, C., Andersson, J., & Tracey, I. (2011). Relief as a reward: Hedonic and neural responses to safety from pain. *PLOS One, 6*(4), e17870.
- Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., & Niv, Y. (2017). Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron, 93*(2), 451–463.
- Li, J., & Daw, N. D. (2011). Signals in human striatum are appropriate for policy update rather than value prediction. *The Journal of Neuroscience, 31*(14), 5504–5511.
- Li, J., & Zhang, Y. (2006). Prediction error method-based second-order structural identification algorithm in stochastic state space formulation. *Earthquake Engineering & Structural Dynamics, 35*(6), 761–779.
- Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience, 7*(3), 274–281.
- Litt, A., Plassmann, H., Shiv, B., & Rangel, A. (2011). Dissociating valuation and saliency signals during decision-making. *Cerebral Cortex (N Y 1991), 21*(1), 95–102.
- Liu, X., Hairston, J., Schrier, M., & Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages: A meta-analysis of functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews, 35*(5), 1219–1236.
- Losecaat Vermeer, A., Boksem, M., & G Sanfey, A. (2014). Neural mechanisms underlying context-dependent shifts in risk preferences. *NeuroImage, 103*, 355.
- Luking, K. R., Luby, J. L., & Barch, D. M. (2014). Kids, candy, brain and behavior: Age differences in responses to candy gains and losses. *Developmental Cognitive Neuroscience, 9*, 82–92.
- Manza, P., Hu, S., Ide, J. S., Farr, O. M., Zhang, S., Leung, H.-C., & Li, C. R. (2016). The effects of methylphenidate on cerebral responses to conflict anticipation and unsigned prediction error in a stop-signal task. *Journal of Psychopharmacology, 30*(3), 283–293.
- Marsh, R., Hao, X., Xu, D., Wang, Z., Duan, Y., Liu, J., ... Peterson, B. S. (2010). A virtual reality-based fMRI study of reward-based spatial learning. *Neuropsychologia, 48*(10), 2912–2921.
- Mathys, C. D., Lomakina, E. I., Daunizeau, J., Iglesias, S., Brodersen, K. H., Friston, K. J., & Stephan, K. E. (2014). Uncertainty in perception and the Hierarchical Gaussian Filter. *Frontiers in Human Neuroscience, 8*, 825.
- Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature, 459*(7248), 837–841.
- Mattfeld, A. T., Gluck, M. A., & Stark, C. E. L. (2011). Functional specialization within the striatum along both the dorsal/ventral and anterior/posterior axes during associative learning via reward and punishment. *Learning and Memory (Cold Spring Harbor, NY), 18*(11), 703–711.
- Maunsell, J. H. R. (2004). Neuronal representations of cognitive state: Reward or attention?. *Trends in Cognitive Sciences, 8*(6), 261–265.
- McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron, 38*(2), 339–346.
- Menon, V., & Levitin, D. J. (2005). The rewards of music listening: Response and physiological connectivity of the mesolimbic system. *NeuroImage, 28*(1), 175–184.
- Metereau, E., & Dreher, J.-C. (2013). Cerebral correlates of salient prediction error for different rewards and punishments. *Cerebral Cortex, 23*(2), 477–487.
- Metereau, E., & Dreher, J.-C. (2015). The medial orbitofrontal cortex encodes a general unsigned value signal during anticipation of both appetitive and aversive events. *Cortex, 63*, 42–54.
- Meyniel, F., & Dehaene, S. (2017). Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proceedings of the National Academy of Sciences of the United States of America, 114*(19), E3859–E3868.
- Middlebrooks, P. G., & Sommer, M. A. (2012). Neuronal correlates of metacognition in primate frontal cortex. *Neuron, 75*(3), 517–530.
- Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience, 16*, 1936–1947.
- Morrens, J. (2014). Dopamine neurons coding prediction errors in reward space, but not in aversive space: A matter of location?. *Journal of Neurophysiology, 112*(5), 1021–1024.
- Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasley, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience, 15*(7), 1040–1046.
- 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. *The European Journal of Neuroscience, 21*(11), 3161–3168.
- Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *The Journal of Neuroscience, 35*(21), 8145–8157.
- Noonan, M. P., Mars, R. B., & Rushworth, M. F. S. (2011). Distinct roles of three frontal cortical areas in reward-guided behavior. *The Journal of Neuroscience, 31*(40), 14399–14412.
- O'Doherty, J., Critchley, H., Deichmann, R., & Dolan, R. J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *The Journal of Neuroscience, 23*, 7931–7939.
- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304*, 452–454.
- O'Doherty, J. P., Hampton, A., & Kim, H. (2007). model-Based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences, 1104*, 35–53.
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience, 4*, 95–102.
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Research Article. Computational Intelligence and Neuroscience, 2011*, 1.
- O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences, 110*, E3660–E3669.
- den Ouden, H. E. M., Kok, P., & de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology, 3*, 548.
- Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience, 5*(2), 97–98.
- Paschke, L. M., Walter, H., Steimke, R., Ludwig, V. U., Gaschler, R., Schubert, T., & Stelzel, C. (2015). Motivation by potential gains and losses

- affects control processes via different mechanisms in the attentional network. *NeuroImage*, 111, 549–561.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Reviews*, 87(6), 532–552.
- Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., & Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron*, 59(4), 561–567.
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, 442(7106), 1042–1045.
- Philiastides, M. G., Biele, G., & Heekeren, H. R. (2010a). A mechanistic account of value computation in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 107(20), 9430–9435.
- Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzner, P., & Heekeren, H. R. (2010b). Temporal dynamics of prediction error processing during reward-based decision making. *NeuroImage*, 53, 221–232.
- Philiastides, M. G., Heekeren, H. R., & Sajda, P. (2014). human scalp potentials reflect a mixture of decision-related signals during perceptual choices. *The Journal of Neuroscience*, 34(50), 16877–16889.
- Pisauro, A., Fouragnan, E., Retzler, C., & Philiastides, M. (2017). Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI. *Nature Communications*, 8, 15808.
- Poudel, G. R., Innes, C. R. H., & Jones, R. D. (2013). Distinct neural correlates of time-on-task and transient errors during a visuomotor tracking task after sleep restriction. *NeuroImage*, 77, 105–113.
- Preusschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *Journal of Neuroscience*, 28(11), 2745–2752.
- Queirazza, F., Fouragnan, E., Steele, J. D., Cavanagh, J., & Philiastides, M. (2017). Dorsomedial prefrontal cortex activity during reinforcement learning discriminates response to Cognitive Behavioural Therapy in depression. bioRxiv, 224410.
- Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.
- Ribas-Fernandes, J. J. F., Solway, A., Diuk, C., McGuire, J. T., Barto, A. G., Niv, Y., & Botvinick, M. M. (2011). A neural signature of hierarchical reinforcement learning. *Neuron*, 71(2), 370–379.
- Rodriguez, P. F. (2009). Stimulus-outcome learnability differentially activates anterior cingulate and hippocampus at feedback processing. *Learning and Memory (Cold Spring Harbor, NY)*, 16(5), 324–331.
- Roesch, M. R., & Olson, C. R. (2004). Neuronal activity related to reward value and motivation in primate frontal cortex. *Science (New York, N. Y.)*, 304(5668), 307–310.
- Rohe, T., & Noppeney, U. (2015). Cortical hierarchies perform bayesian causal inference in multisensory perception. *PLOS Biology*, 13(2), e1002073.
- Rohe, T., Weber, B., & Fließbach, K. (2012). Dissociation of BOLD responses to reward prediction errors and reward receipt by a model comparison. *The European Journal of Neuroscience*, 36(3), 2376–2382.
- Rolls, E. T., Grabenhorst, F., & Parris, B. A. (2008). Warm pleasant feelings in the brain. *NeuroImage*, 41(4), 1504–1513.
- Rudolf, S., Preusschoff, K., & Weber, B. (2012). Neural correlates of anticipation risk reflect risk preferences. *The Journal of Neuroscience*, 32(47), 16683–16692.
- Samson, R. D., Frank, M. J., & Fellous, J.-M. (2010). Computational models of reinforcement learning: The role of dopamine as a reward signal. *Cognitive Neurodynamics*, 4(2), 91–105.
- Sarinopoulos, I., Grupe, D. W., Mackiewicz, K. L., Herrington, J. D., Lor, M., Steege, E. E., & Nitschke, J. B. (2010). Uncertainty during anticipation modulates neural responses to aversion in human insula and amygdala. *Cerebral Cortex*, 20(4), 929–940.
- Schlagenhauf, F., Rapp, M. A., Huys, Q. J. M., Beck, A., Wüstenberg, T., Deserno, L., ... Heinz, A. (2013). Ventral striatal prediction error signaling is associated with dopamine synthesis capacity and fluid intelligence. *Human Brain Mapping*, 34(6), 1490–1499.
- Scholl, J., Kolling, N., Nelissen, N., Wittmann, M. K., Harmer, C. J., & Rushworth, M. F. S. (2015). The good, the bad, and the irrelevant: Neural mechanisms of learning real and hypothetical rewards and effort. *The Journal of Neuroscience*, 35(32), 11233–11251.
- Schonberg, T., O'Doherty, J. P., Joel, D., Inzelberg, R., Segev, Y., & Daw, N. D. (2010). Selective impairment of prediction error signaling in human dorsolateral but not ventral striatum in Parkinson's disease patients: Evidence from a model-based fMRI study. *NeuroImage*, 49(1), 772–781.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80(1), 1–27.
- Schultz, W. (2016a). Dopamine reward prediction-error signalling: A two-component response. *Nature Reviews Neuroscience*, 17(3), 183–195.
- Schultz, W. (2016b). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1), 23–32.
- Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *The Journal of Neuroscience*, 13, 900–913.
- Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *The Journal of Neuroscience*, 12(12), 4595–4610.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science (New York, N.Y.)*, 275(5306), 1593–1599.
- Schwartenbeck, P., FitzGerald, T. H. B., & Dolan, R. (2016). Neural signals encoding shifts in beliefs. *NeuroImage*, 125, 578–586.
- Scimeca, J. M., Katzman, P. L., & Badre, D. (2016). Striatal prediction errors support dynamic control of declarative memory decisions. *Nature Communications*, 7, 13061.
- Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron*, 60(6), 1169–1181.
- Sescousse, G., Caldú, X., Segura, B., & Dreher, J.-C. (2013). Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, 37(4), 681–696.
- Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *The Journal of Neuroscience*, 27(18), 4826–4831.
- Silvetti, M., & Verguts, T. (2012). Reinforcement learning, high-level cognition, and the human brain, *Neuroimaging - Cognitive and Clinical Neuroscience*, Prof. Peter Bright (Ed.), InTech, <https://doi.org/10.5772/23471>
- So, N., & Stuphorn, V. (2012). Supplementary eye field encodes reward prediction error. *The Journal of Neuroscience*, 32(9), 2950–2963.
- Späti, J., Chumbley, J., Brakowski, J., Dörig, N., Grosse Holtforth, M., Seifritz, E., & Spinelli, S. (2014). Functional lateralization of the anterior insula during feedback processing. *Human Brain Mapping*, 35(9), 4428–4439.

- Spicer, J., Galvan, A., Hare, T. A., Voss, H., Glover, G., & Casey, B. (2007). Sensitivity of the nucleus accumbens to violations in expectation of reward. *NeuroImage*, *34*(1), 455–461.
- Spoormaker, V. I., Andrade, K. C., Schröter, M. S., Sturm, A., Goya-Maldonado, R., Sämman, P. G., & Czisch, M. (2011). The neural correlates of negative prediction error signaling in human fear conditioning. *NeuroImage*, *54*(3), 2250–2256.
- Sutton, R. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Takemura, H., Samejima, K., Vogels, R., Sakagami, M., & Okuda, J. (2011). Stimulus-dependent adjustment of reward prediction error in the midbrain. *PLOS One*, *6*(12), e28337.
- Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., & Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature Neuroscience*, *7*(8), 887.
- Tanaka, S. C., Samejima, K., Okada, G., Ueda, K., Okamoto, Y., Yamawaki, S., & Doya, K. (2006). Brain mechanism of reward prediction under predictable and unpredictable environmental dynamics. *Neural Networks*, *19*(8), 1233–1241.
- Tobia, M. J., Gläscher, J., & Sommer, T. (2016). Context-specific behavioral surprise is differentially correlated with activity in anterior and posterior brain systems. *Neuroreport*, *27*(9), 677–682.
- Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *NeuroImage*, *16*(3 Pt 1), 765–780.
- Ullsperger, M., & Cramon, D. Y. V. (2003). Error monitoring using external feedback: Specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, *23*, 4308–4314.
- Valentin, V. V., & O'Doherty, J. P. (2009). Overlapping prediction errors in dorsal striatum during instrumental learning with juice and money reward in the human brain. *Journal of Neurophysiology*, *102*(6), 3384–3391.
- Watanabe, N., Sakagami, M., & Haruno, M. (2013). Reward prediction error signal enhanced by striatum-amygdala interaction explains the acceleration of probabilistic reward learning by emotion. *The Journal of Neuroscience*, *33*(10), 4487–4493.
- Wittmann, M. K., Kolling, N., Akaiishi, R., Chau, B. K. H., Brown, J. W., Nelissen, N., & Rushworth, M. F. S. (2016). Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nature Communications*, *7*, 12327.
- Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2009). Neural computations underlying action-based decision making in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(40), 17199–17204.
- Wunderlich, K., Symmonds, M., Bossaerts, P., & Dolan, R. J. (2011). Hedging your bets by learning reward correlations in the human brain. *Neuron*, *71*(6), 1141–1152.
- Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D. F., & Büchel, C. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *The Journal of Neuroscience*, *26*(37), 9530–9537.
- Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *The Journal of Neuroscience*, *24*(28), 6258–6264.
- Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681–692.
- Zalla, T., Koehlin, E., Pietrini, P., Basso, G., Aquino, P., Sirigu, A., & Grafman, J. (2000). Differential amygdala responses to winning and losing: A functional magnetic resonance imaging study in humans. *The European Journal of Neuroscience*, *12*(5), 1764–1770.
- Zhang, S., Mano, H., Ganesh, G., Robbins, T., & Seymour, B. (2016). Dissociable learning processes underlie human pain conditioning. *Current Biology*, *26*(1), 52–58.
- Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron*, *42*(3), 509–517.

**How to cite this article:** Fouragnan E, Retzler C, Philastides MG. Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis. *Hum Brain Mapp*. 2018;39:2887–2906. <https://doi.org/10.1002/hbm.24047>