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

Fouragnan, Elsa

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

SCHOLARONE™
Manuscripts

1 **Title: Separate neural representations of prediction error valence and surprise:**

2 **evidence from an fMRI meta-analysis**

3

4 **Short title:** Separate neural correlates of prediction error valence and surprise

5

6 **Authors:** Elsa Fouragnan[1,2], Chris Retzler[1,3] and Marios G. Philiastides[1]

7

8 **Affiliations:** [1]Institute of Neuroscience & Psychology, University of Glasgow, Glasgow, UK.

9 [2]Department of Experimental Psychology, University of Oxford, Oxford, UK, [3]Department of

10 Behavioural & Social Sciences, University of Huddersfield, Huddersfield, UK.

11

17

18 **Abstract**

19

20 Learning occurs when an outcome differs from expectations, generating a reward prediction

21 error signal (RPE). The RPE signal has been hypothesized to simultaneously embody the

22 valence of an outcome (better or worse than expected) and its surprise (how far from

23 expectations). Nonetheless, growing evidence suggests that separate representations of the

24 two RPE components exist in the human brain. Meta-analyses provide an opportunity to test

25 this hypothesis and directly probe the extent to which the valence and surprise of the error

26 signal are encoded in *separate* or *overlapping* networks. We carried out several meta-

27 analyses on a large set of fMRI studies investigating the neural basis of RPE, locked at

28 decision outcome. We identified two valence learning systems by pooling studies searching

29 for differential neural activity in response to *categorical* positive-vs-negative outcomes. The

30 first valence network (negative > positive) involved areas regulating alertness and switching

31 behaviors such as the midcingulate cortex, the thalamus and the dorsolateral prefrontal

32 cortex whereas the second valence network (positive > negative) encompassed regions of

33 the human reward circuitry such as the ventral striatum and the ventromedial prefrontal

34 cortex. We also found evidence of a largely distinct surprise-encoding network including the

35 anterior cingulate cortex, anterior insula and dorsal striatum. Together with recent animal

36 and electrophysiological evidence this meta-analysis points to a sequential and distributed

37 encoding of different components of the RPE signal, with potentially distinct functional roles.

38

1
2
3   39   **Introduction**
4
5   40
6
7   41   Effective decision-making depends upon accurate outcome representations associated with
8
9   42   potential choices. These representations can be defined through reinforcement learning (RL)
10
11   43   [Rescorla and Wagner, 1972; Sutton, 1998], a modelling framework that uses the reward
12
13   44   prediction error (RPE), the difference between actual and expected outcomes, as a learning
14
15   45   signal to update future outcome expectations. In this framework, RPE is a signed quantity
16
17   46   and learning is driven by two separate components of the RPE signal: its *valence* (i.e. the
18
19   47   sign of the RPE, representing whether an outcome is better [+] or worse [-] than expected)
20
21   48   and its *surprise* (i.e. the modulus of the RPE, representing the degree [high or low] of
22
23   49   deviation from expectations). Whereas the valence informs an agent whether to reinforce or
24
25   50   extinguish a certain behaviour [Fouragnan et al., 2015; Fouragnan et al., 2017; Frank et al.,
26
27   51   2004], the surprise component determines the extent to which the strength of association
28
29   52   between outcome and expectations needs to be adjusted [Collins and Frank, 2016; Niv et
30
31   53   al., 2015; den Ouden et al., 2012].
32
33   54
34
35   55   This modelling framework has received considerable attention in neuroscience since the
36
37   56   early 90's when animal neurophysiological studies identified dopaminergic neurons in the
38
39   57   midbrain, in particular in the ventral tegmental area (VTA), the substantia nigra pars
40
41   58   compacta (SNc) and reticulata (SNr), whose tonic response profile appears to
42
43   59   simultaneously capture both components of the RPE signal outlined above [Montague et al.,
44
45   60   1996; Schultz et al., 1993; Schultz et al., 1997]. Specifically, these neurons show
46
47   61   anticipatory increase and suppression of their tonic activity in response to positive and
48
49   62   negative RPE respectively. While the anticipatory increase is proportional to the magnitude
50
51   63   of positive RPE, the magnitude of negative RPE is encoded by the duration of the basal
52
53   64   tonic suppression.
54
55   65
56
57
58
59
60

66 This discovery was a breakthrough in the field of learning and decision making and has

67 continued to be influential in the field over the past two and half decades (see [Schultz,

68 2016a; Schultz, 2016b] for a review). As a result, this neurophysiological work has strongly

69 motivated human functional magnetic resonance imaging (fMRI) research to identify the

70 corresponding macroscopic Blood-Oxygen-Level-Dependent (BOLD) pattern of the signed

71 RPE. This pattern of activity was expected to be such that the strength of the BOLD would

72 proceed from high positive RPEs > low positive RPEs > low negative RPEs > high negative

73 RPEs. More specifically, studies have employed a model-based fMRI approach, whereby

74 different types of reinforcement-learning models are first fitted to subjects' behavior to yield

75 parametric predictors for signed RPE against which fMRI data are subsequently regressed

76 [Daw et al., 2011; Fouragnan et al., 2013; Gläscher et al., 2010; O'Doherty et al., 2004;

77 O'doherty et al., 2007; Queirazza et al., 2017].

78

79 These fMRI studies have employed different algorithms to derive the signed RPE, ranging

80 from the simple formulation of the temporal difference learning algorithm to incorporating

81 action learning, notably using the Q-learning and SARSA ('state, action, reward, state, and

82 action') algorithms [Schonberg et al., 2010; Seymour et al., 2007; Tanaka et al., 2006].

83 According to qualitative reviews of this previous findings [O'doherty et al., 2007] as well as

84 quantitative, coordinate-based meta-analyses of these studies, the regions correlating with

85 the different formulations of signed RPE have been found to be predominantly subcortical,

86 including the striatum and amygdala, with some cortical regions, such as the ventromedial

87 prefrontal cortex and the cingulate cortex also reported [Bartra et al., 2013; Garrison et al.,

88 2013; Liu et al., 2011]. Additionally, substantial effort has been undertaken to identify how

89 different types of outcomes (primary reward such as food, or secondary reward such as

90 monetary outcomes) can modulate signed RPE in the same regions and the extent to which

91 it can be considered a domain-general, common currency signal [Sescousse et al., 2013].

92

93   While using trial-by-trial estimates of signed RPE from reinforcement-learning models has

94   provided an enormously productive framework for understanding learning and decision-

95   making, a growing number of studies have also discussed the complementary role of

96   surprise, namely the unsigned RPE, which can also be estimated at the single-trial level.

97   These include, but are not limited to, the use of trial-by-trial estimates of the modulus of RPE

98   or Bayesian surprise according to Bayesian learning theory [Hayden et al., 2011; Iglesias et

99   al., 2013]. Additionally, human electroencephalography (EEG) studies, attempting to offer a

100   temporal account of the cortical dynamics associated with RPE processing, did not find a

101   systematic monotonic response profile consistent with a single RPE representation but

102   instead offered evidence suggestive of separate representations for valence and surprise at

103   the macroscopic level of responses recorded on the scalp. Specifically, multiple recent EEG

104   studies combining model-based RPE estimates with single-trial analysis of the EEG revealed

105   an early outcome stage reflecting a purely categorical valence signal and a later processing

106   stage reflecting separate representations for valence and surprise [Fouragnan et al., 2015;

107   Fouragnan et al., 2017; Philiastides et al., 2010b]. These later valence and surprise signals

108   appeared in spatially distinct but temporally overlapping EEG signatures.

109

110   These findings suggest that, in addition to the fully monotonic firing pattern of midbrain

111   neurons, there exist individual representations for valence and surprise, potentially

112   subserving different functional roles during reward-based learning (e.g. approach-avoidance

113   behavior and the speed of learning via varying degrees of attentional engagement,

114   respectively). Here, we conducted an fMRI meta-analysis to explore the possibility that there

115   exist separate neuronal representations encoding valence and surprise promoting reward

116   learning in humans. We discuss the findings of our work in the context of recent reports from

117   animal neurophysiology and human neuroimaging experiments that provide evidence

118   towards a distributed coding of the different facets of the RPE signal [Brischoux et al., 2009;

119   Fouragnan et al., 2015; Fouragnan et al., 2017; Matsumoto and Hikosaka, 2009].

120

Materials and Methods

122

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 twenty papers 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 papers [Bartra et al., 2013; Garrison et al., 2013; Sescousse et al., 2013].

132

Abstracts from the 788 candidate-papers 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 papers 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 papers 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

148   comprised 102 publications with a total of 2316 participants, 144 contrasts, and 991

149   activation foci. The number of participants per study ranged from 8 to 66 (median = 24,

150   interquartile range [IQR] = 7).

151

152   ***Study categorization****.* The goal of this meta-analysis was to separately categorize studies

153   along the three components of RPE, locked at time of outcome, in order to: 1) identify the

154   extent to which there exist distinct neural representations for valence and surprise and 2)

155   identify whether the neural correlates of the signed RPE simply intersect those of valence

156   and surprise (possibly due to colinearities across these components) or appear as unique

157   clusters of activation reflecting the true combined influence of the two measures.

158

159   To group the relevant papers according to the three main RPE components we used the

160   following definitions: 1) *valence* represents the sign of the RPE and as such it is positive

161   when an outcome is better than expected and negative when worse than expected, 2)

162   *surprise* represents the absolute degree of deviation from expectations and is treated as an

163   unsigned quantity and 3) *signed RPE* simultaneously reflects the influence of both valence

164   and surprise and appears as a fully signed parametric signal. According to these definitions,

165   we identified several fMRI statistical analyses conducted in the original studies that fall under

166   each of the three RPE components (Table 1). The main assumptions of these fMRI

167   analyses, with regard to the BOLD signal as a function of each RPE component, are

168   presented schematically in Figure 1.

169

170   [Figure 1]

171

172   For the valence components, the literature has looked at neural responses which vary

173   categorically along positive-negative axes, as represented in patterns A (i) and (ii) of Figure

174   1. We therefore extracted activations exhibiting a relative BOLD signal increase for negative

175   relative to positive outcomes (NEG > POS: pattern A (i)) and greater BOLD for positive

176   relative to negative outcomes (POS > NEG: pattern A (ii)), respectively. We considered six

177   types of fMRI statistical comparisons which reported coordinate results from either: (1) a

178   contrast associated with negative > positive outcomes, (2) a contrast associated with

179   negative > no outcomes, (3) a negative correlation with a trial-by-trial regressor modulated

180   by [+1] for positive outcomes and [-1] for negative outcomes, (4) the positive correlation with

181   the regressor described in (3), (5) a contrast associated with positive > negative outcomes

182   and (6) a contrast associated with positive > no outcomes. We grouped results from

183   contrasts 1-3 (i.e. NEG > POS) and contrasts 3-6 (i.e. POS > NEG) to capture regions

184   yielding greater BOLD activity for negative relative to positive outcomes and a greater

185   activity for positive relative to negative outcomes respectively (Table 1).

186

187   While the fMRI literature on RPE processing has produced a large amount of theoretical and

188   empirical evidence for the valence and the signed RPE components, comparatively little has

189   been done to directly investigate surprise as a separate component. Fewer studies have

190   used fMRI regressors that were parametrically modulated by trial-to-trial changes in surprise

191   using the unsigned RPE [Fouragnan et al., 2017; Hayden et al., 2011; Iglesias et al., 2013].

192   These studies used the terms "surprise", "unsigned RPE", or outcome "salience" to refer to

193   the mathematical modulus of RPE from computational learning models. In addition to these

194   papers, our literature search has revealed a number of other measures (see below), which

195   are highly correlated with outcome surprise, as defined by learning theory. We therefore

196   used these measures as proxies of surprise to gain insights into the spatial extent of the

197   relevant neural responses and the degree to which they overlap with those associated with

198   valence.

199

200   Specifically, a recent line of research has investigated the neural basis of "Bayesian

201   surprise" or "volatility", computed as the direct modulus of Bayesian predictive error [Ide et

202   al., 2013; Iglesias et al., 2013; Mathys et al., 2014; O'Reilly et al., 2013] which correspond to

203   the absolute difference between categorical outcomes and the probabilistic expectation of

204 these outcomes, estimated using Bayesian inference. In the framework of Bayesian learning,

205 the absolute Bayesian RPE plays an important role in learning from rapid changes in

206 behavioral exploration [Courville et al., 2006]. Finally, other studies used the term

207 "associability" which is a parameter in the Pearce-Hall model [Hall and Pearce, 1979; Pearce

208 and Hall, 1980] defined as the degree of divergence between an actual outcome and the

209 original expectation (e.g., the associative strength between a choice and an outcome). We

210 note however, that in the RL framework, associability can also refer to the learning rate. It is

211 clear from these reports that there is a lack of consistent terminology to refer to unsigned

212 RPE, which emphasizes the need for a more unified framework for studying RPE

213 processing.

214

215 To test for consistencies in the neuronal responses across these different reports, and

216 provide initial support for a unified representation of surprise, we grouped fMRI analyses

217 which reported outcome-locked activations resulting from: (1) a positive correlation with a

218 trial-by-trial regressor of the modulus (unsigned) RPE resulting from RL models across both

219 positive and negative outcomes ("surprise" or "unsigned RPE"), (2) a positive correlation

220 with a trial-by-trial regressor of the unsigned RPE resulting from Bayesian modelling

221 ("Bayesian Surprise" or "volatility"), (3) a positive correlation with a trial-by-trial regressor of

222 the free parameter of the Pearce-Hall model ("associability" term), (4) a contrast associated

223 with (high positive outcomes and high negative outcomes) > (low positive outcomes and low

224 negative outcomes OR no outcomes, (5) a positive correlation with a parametric regressor of

225 surprising positive RPE alone and (6) a positive correlation with a parametric regressor of

226 surprising negative RPE alone (Table 1). Figure 1 illustrates the hypothesized pattern of

227 BOLD signal predicted by these contrasts (pattern B), exhibiting a V shaped response profile

228 that is maximal for both highly surprising negative and positive RPEs. Despite possible

229 subtle differences in the definition of these measures we expected that only foci consistently

230 correlating with deviations from reward expectations would be revealed in this analysis.

231

232   One reason the surprise component has not been looked at closely in isolation is because

233   the literature has focused primarily on signed RPE representations instead. This approach

234   was motivated by neurophysiology experiments showing monotonic responses as a function

235   of both valence and surprise and by a theoretical framework suggesting that learning is

236   driven by a single signed RPE representation. To identify the spatial extent of these

237   representations we also looked at fMRI data reporting positive correlations with signed RPE

238   (negative correlation were discarded). Specifically, we combined four types of fMRI

239   analyses, which estimated trial-by-trial signed RPE from different computational models. We

240   used fMRI reports from (1) model-free and (2) model-based RL methods. Model-free

241   methods include Markov Chain Monte Carlo and temporal difference methods [Samson et

242   al., 2010; Seymour et al., 2007]. Model-based methods include dynamic programming and

243   certainty equivalent methods [Daw et al., 2005; Doya et al., 2002]. More on these algorithms

244   can be found in the review by [Kaelbling et al., 1996]. We also included continuous

245   parametric analyses using trial-by-trial signed RPE from (3) Bayesian RL framework

246   described above [Iglesias et al., 2013; Mathys et al., 2014; den Ouden et al., 2012]. Finally,

247   our analysis for signed RPE also contained one type of parametric analysis that employed

248   fixed RPE values (not estimated from RL models) ranked on a scale such that (4) high

249   positive RPEs > low positive RPEs > low negative RPEs > high negative RPEs (Table 1).

250   Figure 1 illustrates the hypothesized pattern of BOLD signal predicted by these contrasts

251   (pattern C) and it is assumed to increase linearly as a function of signed RPE.

252

253   Crucially, we note that an issue requiring closer scrutiny pertains to the difficulty in

254   disambiguating the signed RPE pattern of activity from those associated with valence and

255   surprise. Specifically, pattern C (signed RPE) is generally highly correlated with pattern A (ii),

256   (POS > NEG valence) and in studies in which only positive RPEs are considered, pattern C

257   (signed RPE) and pattern B (surprise) are perfectly correlated. Nonetheless, comparing

258   clusters of activations across the three RPE components could potentially reveal whether or

259   not there exist unique clusters of activations associated with signed RPE.

260

261 [Table 1]

262

2.1. Activation Likelihood Estimation (ALE) analysis

264

265 We conducted the meta-analysis using the GingerALE software (version 2.3.6) [Eickhoff et

266 al., 2009] that employs a revised (and rectified [Eickhoff et al., 2017]) version of the

267 activation likelihood estimation (ALE) algorithm [Laird et al., 2005; Turkeltaub et al., 2002],

268 which identifies common areas of activation across studies. This method performs

269 coordinate based meta-analysis which considers each reported foci as a 3D Gaussian

270 probability distribution, centred at the coordinates provided by each study reflecting the

271 spatial uncertainty associated with each reported set of coordinates. Note that each contrast

272 provided to the ALE algorithm is treated as a separate experiment. The probabilities

273 distributions are then combined to create a modelled activation map, namely an ALE map for

274 that contrast. Studies are weighted according to the number of subjects they contain by

275 adjusting the full width at half maximum of the Gaussian distributions. The convergence of

276 results across the whole brain is obtained by computing the union of all resulting voxel-wise

277 ALE scores. To distinguish meaningful convergence from random noise, statistics are

278 computed by comparing ALE scores with an empirical null-distribution representing a

279 random spatial association between studies. To infer true convergence, a random-effect

280 inference is applied to capitalize on the differences between studies rather than between foci

281 within a particular study. The null-hypothesis is modelled by randomly sampling voxels from

282 each of the ALE maps from which the union is obtained. The ALE maps are assessed

283 against the null distribution using a cluster level threshold of specific p-values. Contrast

284 analyses between categories of the entire dataset are determined by ALE subtraction

285 method, including a correction for differences in sample size between the categories.

286

287   Here, we manually extracted all coordinates from the studies shown in Table 1 and entered

288   them into separate files for each of the three RPE components in preparation for the ALE

289   analyses. Any studies that provided coordinates in Talairach space were converted into MNI

290   space by the Matlab (MathWorks, Natick, Massachusetts) function *tal2mni* in the fieldtrip

291   toolbox [Oostenveld et al., 2011]. We conducted ALE analyses for each of the three

292   components of RPE individually. Along the valence component, we looked at both patterns A

293   (i) and A (ii) in Figure 1 (i.e. to identify activations for negative > positive RPE and vice

294   versa, respectively). Accordingly, we ran separate ALE analyses for each of the two

295   patterns. In addition, we performed two conjunction analyses – one between the valence and

296   surprise components to investigate our hypothesis of largely separate neural representations

297   and another between all three RPE components to identify regions that simultaneously

298   encode these representations. Subsequently, we also performed all possible pairwise

299   contrast analyses between the three patterns (A, B and C), using the individual maps

300   associated with each pattern.

301   A total of 402 foci from 66 contrasts were used with 262 foci from 31 contrasts for Pattern A

302   (i) revealing BOLD patterns greater for negative than positive outcomes and 205 from 35

303   contrasts for Pattern A (ii) (e.g. the opposite contrast). For the surprise (Pattern B) and

304   signed RPE (Pattern C) analyses, we applied individual ALE analyses, with 284 foci from 40

305   contrasts for surprise and 240 foci from 38 contrasts for signed RPE. Overall, the number of

306   contrasts used for each separate outcome component was large enough (> 30) to allow

307   sufficient power for the required statistical tests [Eickhoff and Etkin, 2016]. Finally, we

308   transformed the resulting ALE maps from the Colins MNI individual brain space

309   (Colin27_T1_seg_MNI) to the MNI normalized brain space (MNI ICBM152 template) by

310   applying an affine transformation using the FSL *flirt* program [Jenkinson et al., 2002], prior to

311   overlaying onto the canonical MNI template for visualization.

312

313   **3. Results**

314

315   All coordinates used for the following ALE analyses were collated from fMRI studies in which

316   the components of RPE have been regressed onto BOLD activity time-locked to outcome

317   presentation. We report ALE maps with clusters surviving the False Discovery Rate (FDR)

318   yielding two p-value thresholds. The most conservative FDR correction yields a p-value with

319   no assumptions about how the data is correlated (FRN), and the least conservative FDR

320   correction assumes independence or positive dependence (FID) with $p < 0.05$ and a

321   minimum volume clustering value of 50 mm^3. Note that, using a cluster-level family-wise

322   error (FWE) correction implemented with a cluster-extent threshold of $p < 0.05$ and a cluster-

323   forming threshold of $p < 0.001$ revealed virtually identical results (compared with FRN)

324   [Eickhoff et al., 2017] as per previous reports [R Garrison et al., 2017]. For all tables

325   presenting ALE cluster results, the size of each cluster is provided in mm^3 along with the

326   associated MNI coordinates and maximum ALE score. The ALE score indicates the relative

327   effect size for each peak voxel within each ALE analysis.

328

329   **3.1.   Outcome Valence**

330

331   The first two ALE analyses were conducted to identify regions in which BOLD signals

332   correlate with outcome valence. Specifically, we looked at activations that yielded greater

333   BOLD for negative relative to positive outcomes (NEG > POS; pattern A (i) in Figure 1) and

334   greater BOLD for negative relative to positive outcomes (POS > NEG; pattern A (ii) in Figure

335   1), respectively. Accordingly, we considered all fMRI studies, which assumed BOLD

336   responses varying categorically along a positive-negative axis for outcome valence.

337

338   The findings of the two valence ALE analyses are shown in Figure 2. The resulting maps

339   revealed a highly distributed network of brain activations encompassing several cortical

340   regions and sub-cortical structures. More precisely, NEG > POS valence clusters were found

341   in a network encompassing the anterior and dorsal part of the mid-cingulate cortex (aMCC

342   and dMCC) including the pre supplementary motor area (pre-SMA), the bilateral anterior and

343   middle insular cortex (aINS, mINS), the bilateral dorsolateral prefrontal cortex (dlPFC), the

344   bilateral thalamus, right amygdala, left inferior parietal lobule (IPL) and the habenula.

345

346   POS > NEG valence clusters were found in the bilateral ventral striatum (vSTR), the

347   ventromedial prefrontal cortex (vmPFC), the posterior part of the cingulate cortex (PCC), as

348   well as the ventrolateral orbitofrontal cortex (vlOFC). At a lower threshold (uncorrected p-

349   value of 0.001), we also found the midbrain as part of this network, encompassing the VTA,

350   which is commonly associated with the delivery of reward [D'Ardenne et al., 2008]. Table 2

351   contains the complete list of regions, coordinates, and statistics of these two ALE analyses.

352

353   [Figure 2], [Table 2]

354

355   **3.2.   Surprise**

356

357   FMRI investigations of RPE have focused primarily on the valence components while

358   neglecting potential contributions from possible separate representations along the surprise

359   component, defined as the degree by which outcomes deviate from expectations and

360   mathematically expressed as the modulus of RPE. A major goal of this work was to explore

361   the possibility that there exist largely separate neuronal representations encoding surprise.

362   To this end, we conducted a new ALE analysis in which the few empirical fMRI studies

363   making use of the surprise from RL models were combined with other fMRI measures

364   correlated with the surprise as defined by RL models (Table 1).

365

366   Figure 3 shows the areas in which BOLD signal correlated with surprise. We found evidence

367   for activations in a distributed network encompassing the aMCC, dMCC, the pre-SMA the

368   bilateral dorsal striatum (dSTR), the bilateral aINS, the MTG and the midbrain. Crucially, this

369 activation map shows that the neural network associated with surprise is largely distinct from

370 that of valence. This finding provides initial support for the notion that these two RPE

371 components are encoded in separate brain areas and, as such, they might be contributing

372 individually to promote learning. The full results of the surprise ALE analysis are also

373 summarized in Table 3.

374

375 [Figure 3], [Table 3]

376

377 **3.3.   Valence and surprise conjunction and contrast analyses**

378

379 The activation maps for valence (NEG > POS and POS > NEG) and surprise ALE analyses

380 conducted above revealed little overlap between the spatial representations of these two

381 RPE components. To formally quantify the degree of overlap between the valence and

382 surprise networks, we next ran a conjunction analysis between the two components. The

383 statistical map resulting from this conjunction analysis and the two separate statistical maps

384 of valence and surprise (as already reported in Figures 2 and 3) are overlaid in Figure 4.

385

386 [Figure 4], [Table 4]

387

388 Contrast analyses were conducted for each possible pairing between any dimensions of

389 valence (POS > NEG [positive]; NEG > POS [negative] and POS + NEG [all valence]) and

390 surprise. These analyses allowed us to identify the areas that were unique and specific to

391 each individual outcome and RPE-related component. The positive valence (pattern A (ii))

392 minus surprise (pattern B) contrast revealed two main clusters in the vSTR and vmPFC

393 whereas the reverse contrast revealed a network of clusters including preSMA, aINS, and

394 MTG. Contrasting negative valence (pattern A (i)) and surprise also exposed separate

395 networks of areas for each subtraction. Specifically, this contrast revealed a network

396 encompassing the thalamus, the habenula, the right mINS and the dMCC, whereas the

397 reverse contrast showed clusters in the dorsal portion of the STR and the dlPFC. The

398 statistical maps resulting from these contrast analyses are presented in Figure 5.

399

400 [Figure 5], [Table 5]

401

402 **3.4.   Signed RPE**

403

404 A major goal of this work was to investigate the spatial profile of the signed RPE component

405 and to scrutinise more closely the extent to which it overlaps with the separate

406 representations identified for valence (NEG > POS and POS > NEG) and surprise. The

407 fMRI-RPE literature has focused on this component largely due to neurophysiological

408 evidence suggesting that RPE-like learning is driven by a single, theoretically unified

409 representation of both POS > NEG valence and surprise (Table 1).

410

411 Results from this ALE analysis revealed very few unique activations for signed RPE

412 compared to valence and surprise. Instead, brain areas identified in this analysis overlapped

413 mostly with areas appearing in the POS > NEG valence component and, to a lesser extent,

414 surprise (Figure 6). Specifically, a large overlap between signed RPE and the POS > NEG

415 valence component was found in the STR and a smaller one in the vmPFC. Similarly, areas

416 appearing in the singed RPE analysis that overlapped with the surprise component were

417 also found, albeit only in small clusters comprising the aMCC and dorsal STR. Taken

418 together, these findings emphasize the potential collinearities between the BOLD predictors

419 used to identify neural representations associated with the three RPE components and

420 highlight the need for developing a methodology for properly disentangling their individual

421 contributions.

422

423 [Figure 6], [Table 6]

424

1
2
3   **3.5.   Putting it all together**
4
5
6
7   Subsequently, to formally test for the overlap between all three RPE components and
8
9   identify potential regions integrating valence and surprise either into a signed RPE
10
11   representation or a linear superposition of the two signals [Fouragnan et al., 2017], we
12
13   performed a conjunction analysis between the valence (pattern A), the surprise (pattern B)
14
15   and signed RPE (pattern C) signals. We summarize our conjunction results in Figure 7,
16
17   which revealed a major overlap between all activations associated with signed RPE and
18
19   each of the other two RPE representations in the central part of the STR. Thus, one
20
21   possibility is that the STR meets the requirement that a full monotonic representation of the
22
23   error signal also simultaneously encodes valence and surprise, as per our last ALE analysis.
24
25
26
27   [Figure 7]
28
29
30
31   Another possibility is that the overlap between all components of outcomes in the STR is
32
33   arising, at least in part, due to collinearities across the different outcome representations,
34
35   particularly between the positive categorical nature of outcome valence (pattern A (ii)) and
36
37   the signed RPE. To formally test this hypothesis, we performed a new series of contrast
38
39   analyses between signed RPE and all dimensions of categorical valence and surprise.
40
41   Particularly, we performed contrast analyses between patterns C-A(i), C-A(ii), C-A and C-B
42
43   (and vice versa). The results are summarized in Figure 8. Particularly, we did not find any
44
45   area unique to signed RPE when looking at each of the individual comparisons of signed
46
47   RPE with the other three patterns. In fact, when comparing signed RPE to positive valence
48
49   (pattern A (ii)), no clusters were found to be significantly different than those found with the
50
51   categorical outcome valence (POS > NEG). Conversely, the STR was found for all the other
52
53   signed RPE comparisons (signed RPE > negative; signed RPE > surprise). Finally, the
54
55   unique network related to negative valence (pattern A (i)) was found in the dMCC, thalamus
56
57
58
59
60

452   and mINS, the unique cluster related to positive valence was found in the vmPFC and the

453   unique network related to surprise was found in the aMCC, preSMA and the aINS.

454

455   [Figure 8], [table 7]

456

457   **Discussion**

458

459   In this fMRI meta-analysis work, we demonstrated that reward learning in humans involves

460   separate neuronal signatures of RPE, comprising distinct representations for valence and

461   surprise. Together with recent neurophysiological and EEG evidence (including studies

462   using simultaneous EEG and fMRI), these findings point to a potentially sequential and

463   distributed encoding of different RPE components with potentially functionally distinct roles.

464

465   **Valence networks**

466

467   The ALE analyses related to valence revealed two distributed set of activations correlating

468   with both pattern A (i) and (ii) in Figure 1. Foci for which the BOLD signal was greater for

469   negative than positive outcomes showed significant clustering in a large network of areas

470   including the thalamus, the aMCC and dMCC, the aINS, mINS and the dlPFC. Conversely,

471   foci for which the BOLD signal was greater for positive than negative outcomes showed

472   significant clusters in a separate network including vmPFC, vSTR, PCC, and vlOFC. These

473   findings clearly suggest the presence of multiple systems responding to the categorical

474   nature of valence which supports the notion that separate valuation systems shape learning

475   in the human brain [Fiorillo, 2013; Fouragnan et al., 2013], although their functional role

476   remain debated. More specifically, the debate focuses on the number and exact nature of

477   the neural systems assigning value to decision outcomes and driving behaviors that are

478   evolutionarily appropriate in response to changes in the environment.

479

480   A first theory describes two distinct valence systems invoking two orthogonal axes of

481   decision-making: alertness (involving the implementation of action) and learning (including

482   the updates of value expectations for future avoidance and approach behaviors). In this

483   framework, the first system is thought to monitor on-going activity and interrupt it when

484   needed to trigger switching behaviors (e.g. following negative RPEs). In contrast, the second

485   system uses both negative and positive RPE values for decreasing or increasing internal

486   value representations associated with decisions to ultimately drive avoidance and approach

487   learning, respectively [Boureau and Dayan, 2011; Cools et al., 2011; Elliot, 2006; Fiorillo,

488   2013; Fouragnan et al., 2015; Gray and McNaughton, 2003; Guitart-Masip et al., 2012].

489

490   A second (not mutually exclusive) proposition supports the idea that there are at least two

491   separate systems responsible for aversive and appetitive reinforcements such that

492   punishments and rewards are encoded separately (i.e. a punishment space and a reward

493   space [Morrens, 2014]). This proposition was developed on the basis of neurophysiological

494   evidence showing that different types of neurons exhibit differential activity in response to

495   punishing vs. non-punishing outcomes and rewarding vs. non-rewarding outcomes,

496   respectively [Fiorillo et al., 2003; Fiorillo, 2013; Schultz et al., 1992; Schultz, 1998]. In this

497   second theory, the punishment space is responsible for avoidance behaviors as well as

498   avoidance learning and the reward space is responsible for approach behaviors and

499   approach learning.

500

501   It is noteworthy that our meta-analysis on itself cannot directly distinguish between the two

502   theories because the results do not reveal whether the relevant activations respond

503   exclusively to either positive or negative outcomes or are modulated by both outcomes in

504   opposite directions. This distinction is critical because the former response profile would

505   suggest the presence of separate approach and avoidance systems that might not

506   necessarily be linked to the learning processes as such, while the latter might point to both

507   up- and down-regulation of activity consistent with learning and updating of reward

508   expectations. Nonetheless, the meta-analysis results suggest that two main networks

509   process valence. The network encompassing aINS, aMCC, thalamus and dlPFC could

510   regulate on-going activity and alertness or could represent the punishment space in

511   accordance to the first and the second theories respectively. Conversely, the network of

512   regions encompassing the vmPFC, vSTR, PCC and vlOFC could represent the learning

513   system depicted in the first theory or could represent the reward space depicted in the

514   second theory. Further research is required to tease apart the roles of these systems,

515   especially by investigating their precise response profiles in the appetitive (where rewarding

516   and non-rewarding outcomes are manipulated) and in a true aversive (where punishing and

517   non-punishing outcomes are manipulated) domains, respectively.

518

519   **Surprise network**

520

521   Emerging evidence indicates that the brain encodes the unsigned RPE signal (surprise),

522   which alerts the organism of relative deviations from expectations, regardless of the outcome

523   value. However, to date, only few papers have modelled surprise as such to search for

524   independent neural representations, with the exception of recent neurophysiological

525   developments [Brischoux et al., 2009; Matsumoto and Hikosaka, 2009], recent EEG work

526   [Philiastides et al., 2010b; Yeung and Sanfey, 2004] and an increasing number of fMRI

527   studies [Fouragnan et al., 2017; Gläscher et al., 2010; Li and Daw, 2011; Metereau and

528   Dreher, 2013]. Nevertheless, other fMRI studies used variables highly correlated with

529   surprise that can be employed as proxies [Behrens et al., 2007; Iglesias et al., 2013; Nassar

530   et al., 2012; den Ouden et al., 2012; Yu and Dayan, 2005]. These studies share the

531   assumption that the corresponding BOLD response profile is maximal for high positive and

532   high negative RPE and minimal for no RPE, resembling a V-shape, as illustrated with

533   Pattern B in Figure 1. By combining these fMRI results into a single ALE-analysis, we

534   expose for the first time the network associated with surprise while stressing the need for a

535   common lexicon for this learning component to guide subsequent research in the field.

536

The surprise ALE-analysis revealed a large network including cortical and sub-cortical areas such as aMCC, bilateral aINS, dSTR and midbrain, 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 et al., 2011; Zink et al., 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 and 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 et al., 2006; Kahnt and Tobler, 2013] and areas involved in visual and motor preparation such as the supplementary motor area [Wunderlich et al., 2009] or the supplementary eye field [Middlebrooks and Sommer, 2012; So and Stuphorn, 2012].

552

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 behavioral adjustments [Dayan and Balleine, 2002; Fouragnan et al., 2017; Kolling et al., 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 behavioral adjustment [Menon and 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 et al., 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

564 of learning, some authors argue that surprise can also be considered as a signal predicting

565 the level of risk associated with a future decision outcome, and thus reflect a risk RPE

566 [Fiorillo et al., 2003; Preuschoff et al., 2008; Rudorf et al., 2012].

567

568 **Neuromodulatory pathways encoding multicomponent RPE signals**

569

570 Supporting the idea of separate neural systems for valence and surprise, recent

571 electrophysiological work has revealed both signals existing in neighbouring groups of

572 neurons. The first study of this kind observed the response of dopaminergic neurons in

573 ventral and dorsal areas of the SNc and reported two categories of dopamine neurons

574 [Matsumoto and Hikosaka, 2009]. Some dopamine neurons increase their phasic firing

575 activity in response to valence while others responded only to the changes in unsigned RPE,

576 regardless of the valence component. The latter population of neurons was located more

577 dorsolaterally in the SNc, whilst the neurons encoding valence were located more

578 ventromedially, including the VTA. Interestingly, the dorsolateral SNc projects mainly to the

579 dorsal STR, whereas the ventral SNc and VTA project to the ventral STR, which matches

580 the results of our last conjunction analysis (Figure 7). We found that the only region that

581 encodes the full monotonic representation of the RPE as well as the separate valence and

582 surprise components of RPE seems to be the central part of the STR as shown in Figure 7.

583 This result aligns with the assumption that this region receives direct projections from the

584 midbrain dopaminergic neurons encoding a fully monotonic signed RPE signal [Schultz et

585 al., 1997]. Additionally, the meta-analysis also revealed that both the valence (POS > NEG)

586 and surprise networks include activity in the midbrain, confirming this hypothesis.

587

588 It is important to note that identifying neural activity associated with valence and surprise

589 signals is challenging because in many experimental paradigms both components are highly

590 correlated. For example, when positive RPE are manipulated in isolation, valence (POS >

591 NEG) strongly correlates with surprise. Additionally, whether positive or negative, an

592 unexpected outcome attracts more attention, leads to higher levels of emotional arousal and

593 involves higher levels of motor preparation compared to no RPE [Matsumoto and Hikosaka,

594 2009; Maunsell, 2004; Roesch and Olson, 2004]. Consequently, to disentangle these

595 signals, one needs to design tasks in which the level of valence and surprise can

596 independently be controlled and decoupled [Kahnt, 2017; Kahnt and Tobler, 2013] or

597 capitalize on the variability of physiologically-derived responses (i.e. endogenous variability)

598 associated with valence and surprise [Fouragnan et al., 2015; Fouragnan et al., 2017;

599 Pisauro et al., 2017].

600

601 It is important to note that since the problem of collinearity and functional specificity of some

602 brain regions is already present in single studies, it will inevitably be carried over to studies

603 performing conjunction meta-analyses. Virtually every experimental design engages a large

604 number of cognitive operations and, thereby, activates functional neural networks that may

605 be irrelevant to a particular regressor (psychological construct) of interest. For example in

606 our study, regions related to outcome value and surprise might share variance with outcome

607 confidence [Gherman and Philiastides, 2015; Gherman and Philiastides, 2017; Lebreton et

608 al., 2015; Philiastides et al., 2014]. Despite this general limitation and the difficulty of

609 interpreting conjunction results, aggregating results across a large number of experiments

610 allows one to expose convergence of findings across studies and increasing the

611 generalizability of the conclusions. In particular, this meta-analysis, capitalizing on both

612 individual maps of activations as well as contrasts between different outcome components,

613 points to a distributed encoding of valence and surprise, with potentially distinct functional

614 roles.

615

616 **Temporally specific components of RPE processing**

617

618 The presence of separate RPE-related neural systems raises the question of how these

619 systems unfold in time. Capitalizing on the high temporal resolution of EEG, three recent

620   studies using simultaneous EEG-fMRI have started to shed light on the spatiotemporal

621   characterisation of the RPE components. First, these studies have revealed two temporally

622   specific EEG components discriminating between positive and negative RPEs peaking

623   around 220ms and 300ms respectively, largely consistent with the timing of the feedback-

624   related negativity and feedback-related positivity ERP components [Cohen et al., 2007;

625   Hajcak et al., 2006; Yeung and Sanfey, 2004]. Additionally, the studies also revealed a late

626   unsigned RPE component which overlaps temporally with the late valence signal

627   [Philiastides et al., 2010b] but appears in a largely separate and distributed neural network

628   [Fouragnan et al., 2017].

629

630   Based on these previous studies and the current meta-analysis, we propose that the early

631   and late EEG valence components might reflect the separate contributions of the two

632   networks of areas found for the ALE-valence analyses. This proposal assumes that an early

633   network processes mainly negative RPEs in order to initiate a fast alertness response in the

634   presence of negative outcomes. Conversely, a later network – associated with the brain's

635   reward circuitry – is modulated by both positive and negative RPEs, consistent with a role in

636   approach/avoidance learning and value updating [Philiastides et al., 2010a]. We also

637   propose that the surprise network unfolds near simultaneously with the late valence

638   component and thus influences learning through largely distinct spatial representations of

639   the two outcomes signals, which happen to form a composite signal in overlapping areas

640   [Fouragnan et al., 2017].

641

642   **Full representation of a monotonic signed RPE signal**

643

644   To examine the spatial profile of a true monotonic signed RPE representation in the human

645   brain, we pooled results from fMRI studies, which hypothesized that RPE-like learning is

646   driven by a simultaneous representation of both categorical valence and surprise. These

647 fMRI studies are based on the influential assumption that BOLD signal increases

648 monotonically as a function of signed RPE, as illustrated in pattern C (Fig. 1), equivalent to

649 the teaching signal that is predicted in the Rescorla–Wagner model of RL [Rescorla and

650 Wagner, 1972]. Additionally, we combined the valence and surprise networks and

651 subsequently compared it with the signed RPE to test the requirement that the signed RPE

652 simultaneous encodes both components. This conjunction analysis revealed that the only

653 brain region that seems to encode a true monotonic signal is the STR in the basal ganglia,

654 which could explain why such a signal is not tractable with EEG recordings as highlighted

655 earlier. This result confirms the long standing view that the BOLD activity in STR mirrors the

656 dopaminergic signalling of the mesolimbic neurons [Delgado et al., 2000; Haber et al., 1995;

657 O'Doherty et al., 2004; Pagnoni et al., 2002] that fully encode the RL prediction error signal

658 of the Rescorla-Wagner rule [Ikemoto, 2007; Schultz et al., 1992].

659

660 Nonetheless, the ALE contrast analyses between valence (the positive correlation with

661 pattern A (ii)) and signed RPE revealed no significant activation, whereas the reverse

662 contrast revealed a denser cluster of activity in vmPFC for valence than signed RPE. Given

663 the evidence presented above that the signed RPE may only be encoded in the STR, we

664 suggest that this result may arise due to collinearities between valence and signed RPE or

665 surprise and signed RPE. More precisely, a parametric predictor for signed RPE would be

666 positively correlated with the contrast positive > negative outcomes whereas the signed RPE

667 and surprise would be perfectly correlated in the positive (appetitive) domain.

668

669 **Conclusion**

670

671 In conclusion, the current meta-analysis points to a framework whereby heterogeneous

672 signals are involved in RPE processing. The proposal of a temporally distinct and spatially

673 distributed representation of valence and surprise is open to debate and many questions

674 remain about how these signals interact and how they correspond to the computations made

675 in the brain. For example, it is currently unclear whether valence and surprise encoding

676 occur before the computation of the signed RPE, or whether these three computations are

677 performed in parallel. Nevertheless the taxonomy proposed is conceptually useful because it

678 breaks down the learning and valuation processes into testable components and organizes

679 the RPE literature in terms of the computations that are potentially involved. It will require

680 additional experiments to validate the current proposal and to better understand the

681 complexity of RPE processing.

682

References

683 Abler B, Walter H, Erk S, Kammerer H, Spitzer M (2006): Prediction error as a linear function
684 of reward probability is coded in human nucleus accumbens. NeuroImage 31:790–
685 795.
687 Allen M, Fardo F, Dietz MJ, Hillebrandt H, Friston KJ, Rees G, Roepstorff A (2016): Anterior
688 insula coordinates hierarchical processing of tactile mismatch responses.
689 NeuroImage 127:34–43.
690 Amado C, Hermann P, Kovács P, Grotheer M, Vidnyánszky Z, Kovács G (2016): The
691 contribution of surprise to the prediction based modulation of fMRI responses.
692 Neuropsychologia 84:105–112.
693 Amiez C, Sallet J, Procyk E, Petrides M (2012): Modulation of feedback related activity in the
694 rostral anterior cingulate cortex during trial and error exploration. NeuroImage
695 63:1078–1090.
696 Aron AR, Shohamy D, Clark J, Myers C, Gluck MA, Poldrack RA (2004): Human Midbrain
697 Sensitivity to Cognitive Feedback and Uncertainty During Classification Learning. J
698 Neurophysiol 92:1144–1152.
699 Bartra O, McGuire JT, Kable JW (2013): The valuation system: a coordinate-based meta-
700 analysis of BOLD fMRI experiments examining neural correlates of subjective value.
701 NeuroImage 76:412–427.
702 Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS (2007): Learning the value of
703 information in an uncertain world. Nat Neurosci 10:1214–1221.
704 Bickel WK, Pitcock JA, Yi R, Angtuaco EJC (2009): Congruence of BOLD Response across
705 Intertemporal Choice Conditions: Fictive and Real Money Gains and Losses. J
706 Neurosci 29:8839–8846.
707 Boll S, Gamer M, Gluth S, Finsterbusch J, Büchel C (2013): Separate amygdala subregions
708 signal surprise and predictiveness during associative fear learning in humans. Eur J
709 Neurosci 37:758–767.
710 van den Bos W, Cohen MX, Kahnt T, Crone EA (2012): Striatum-medial prefrontal cortex
711 connectivity predicts developmental changes in reinforcement learning. Cereb
712 Cortex N Y N 1991 22:1247–1255.
713 Boureau Y-L, Dayan P (2011): Opponency Revisited: Competition and Cooperation Between
714 Dopamine and Serotonin. Neuropsychopharmacology 36:74–97.
715 Brischoux F, Chakraborty S, Brierley DI, Ungless MA (2009): Phasic excitation of dopamine
716 neurons in ventral VTA by noxious stimuli. Proc Natl Acad Sci U S A 106:4894–4899.
717 Browning M, Holmes EA, Murphy SE, Goodwin GM, Harmer CJ (2010): Lateral Prefrontal
718 Cortex Mediates the Cognitive Modification of Attentional Bias. Biol Psychiatry
719 67:919–925.
720 de Bruijn E, de Lange F, Cramon D, Ullsperger M (2009): When Errors Are Rewarding. J
721 Neurosci Off J Soc Neurosci 29:12183–6.
722 Canessa N, Crespi C, Motterlini M, Baud-Bovy G, Chierchia G, Pantaleo G, Tettamanti M,
723 Cappa SF (2013): The Functional and Structural Neural Basis of Individual Differences
724 in Loss Aversion. J Neurosci 33:14307–14317.
725 Chumbley JR, Burke CJ, Stephan KE, Friston KJ, Tobler PN, Fehr E (2014): Surprise beyond
726 prediction error. Hum Brain Mapp 35:4805–4814.

727 Cohen MX, Elger CE, Ranganath C (2007): Reward Expectation Modulates Feedback-Related
728   Negativity and EEG Spectra. NeuroImage 35:968–978.
729 Cohen MX, Ranganath C (2007): Reinforcement Learning Signals Predict Future Decisions. J
730   Neurosci 27:371–378.
731 Collins AGE, Frank MJ (2016): Surprise! Dopamine signals mix action, value and error. Nat
732   Neurosci 19:3–5.
733 Cools R, Nakamura K, Daw ND (2011): Serotonin and Dopamine: Unifying Affective,
734   Activational, and Decision Functions. Neuropsychopharmacology 36:98–113.
735 Courville AC, Daw ND, Touretzky DS (2006): Bayesian theories of conditioning in a changing
736   world. Trends Cogn Sci 10:294–300.
737 Daniel R, Wagner G, Koch K, Reichenbach JR, Sauer H, Schlösser RGM (2011): Assessing the
738   neural basis of uncertainty in perceptual category learning through varying levels of
739   distortion. J Cogn Neurosci 23:1781–1793.
740 D'Ardenne K, McClure SM, Nystrom LE, Cohen JD (2008): BOLD responses reflecting
741   dopaminergic signals in the human ventral tegmental area. Science 319:1264–1267.
742 Daw ND, Niv Y, Dayan P (2005): Uncertainty-based competition between prefrontal and
743   dorsolateral striatal systems for behavioral control. Nat Neurosci 8:1704–1711.
744 Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011): Model-Based Influences on
745   Humans' Choices and Striatal Prediction Errors. Neuron 69:1204–1215.
746 Dayan P, Balleine BW (2002): Reward, motivation, and reinforcement learning. Neuron
747   36:285–298.
748 Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA (2000): Tracking the Hemodynamic
749   Responses to Reward and Punishment in the Striatum. J Neurophysiol 84:3072–
750   3077.
751 Delgado MR (2007): Reward-Related Responses in the Human Striatum. Ann N Y Acad Sci
752   1104:70–88.
753 Demos KE, Heatherton TF, Kelley WM (2012): Individual Differences in Nucleus Accumbens
754   Activity to Food and Sexual Images Predict Weight Gain and Sexual Behavior. J
755   Neurosci 32:5549–5552.
756 Diederen KMJ, Ziauddeen H, Vestergaard MD, Spencer T, Schultz W, Fletcher PC (2017):
757   Dopamine Modulates Adaptive Prediction Error Coding in the Human Midbrain and
758   Striatum. J Neurosci Off J Soc Neurosci 37:1708–1720.
759 Diuk C, Tsai K, Wallis J, Botvinick M, Niv Y (2013): Hierarchical Learning Induces Two
760   Simultaneous, But Separable, Prediction Errors in Human Basal Ganglia. J Neurosci
761   33:5797–5805.
762 Doya K, Samejima K, Katagiri K, Kawato M (2002): Multiple model-based reinforcement
763   learning. Neural Comput 14:1347–1369.
764 Dreher J-C (2013): Neural coding of computational factors affecting decision making. Prog
765   Brain Res 202:289–320.
766 van Duijvenvoorde ACK, Op de Macks ZA, Overgaauw S, Gunther Moor B, Dahl RE, Crone EA
767   (2014): A cross-sectional and longitudinal analysis of reward-related brain activation:
768   Effects of age, pubertal stage, and reward sensitivity. Brain Cogn 89:3–14.
769 Dunne S, D'Souza A, O'Doherty JP (2016): The involvement of model-based but not model-
770   free learning signals during observational reward learning in the absence of choice. J
771   Neurophysiol 115:3195–3203.
772 Eickhoff SB, Etkin A (2016): Going Beyond Finding the "Lesion": A Path for Maturation of
773   Neuroimaging. Am J Psychiatry 173:302–303.

774 Eickhoff SB, Laird AR, Fox PM, Lancaster JL, Fox PT (2017): Implementation errors in the
775   GingerALE Software: Description and recommendations. Hum Brain Mapp 38:7–11.
776 Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT (2009): Coordinate-based
777   activation likelihood estimation meta-analysis of neuroimaging data: a random-
778   effects approach based on empirical estimates of spatial uncertainty. Hum Brain
779   Mapp 30:2907–2926.
780 Elliot AJ (2006): The Hierarchical Model of Approach-Avoidance Motivation. Motiv Emot
781   30:111–116.
782 Elliott R, Friston KJ, Dolan RJ (2000): Dissociable neural responses in human reward systems.
783   J Neurosci Off J Soc Neurosci 20:6159–6165.
784 Elward RL, Vilberg KL, Rugg MD (2015): Motivated Memories: Effects of Reward and
785   Recollection in the Core Recollection Network and Beyond. Cereb Cortex 25:3159–
786   3166.
787 Ernst M, Nelson EE, McClure EB, Monk CS, Munson S, Eshel N, Zarahn E, Leibenluft E,
788   Zametkin A, Towbin K, Blair J, Charney D, Pine DS (2004): Choice selection and
789   reward anticipation: an fMRI study. Neuropsychologia 42:1585–1597.
790 Ferdinand NK, Opitz B (2014): Different aspects of performance feedback engage different
791   brain areas: Disentangling valence and expectancy in feedback processing. Sci Rep
792   4:srep05986.
793 Fiorillo CD (2013): Two Dimensions of Value: Dopamine Neurons Represent Reward But Not
794   Aversiveness. Science 341:546–549.
795 Fiorillo CD, Tobler PN, Schultz W (2003): Discrete Coding of Reward Probability and
796   Uncertainty by Dopamine Neurons. Science 299:1898–1902.
797 Forster SE, Brown JW (2011): Medial prefrontal cortex predicts and evaluates the timing of
798   action outcomes. NeuroImage 55:253–265.
799 Fouragnan E, Chierchia G, Greiner S, Neveu R, Avesani P, Coricelli G (2013): Reputational
800   Priors Magnify Striatal Responses to Violations of Trust. J Neurosci 33:3602–3611.
801 Fouragnan E, Queirazza F, Retzler C, Mullinger K, Philiastides M (2017): Spatiotemporal
802   neural characterization of prediction error valence and surprise during reward
803   learning in humans. Sci Rep. https://www.nature.com/srep/.
804 Fouragnan E, Retzler C, Mullinger K, Philiastides MG (2015): Two spatiotemporally distinct
805   value systems shape reward-based learning in the human brain. Nat Commun
806   6:8107.
807 Frank MJ, Seeberger LC, O'reilly RC (2004): By carrot or by stick: cognitive reinforcement
808   learning in parkinsonism. Science 306:1940–1943.
809 Friston K (2009): The free-energy principle: a rough guide to the brain? Trends Cogn Sci
810   13:293–301.
811 Friston K, Kilner J, Harrison L (2006): A free energy principle for the brain. J Physiol Paris
812   100:70–87.
813 Fujiwara J, Tobler PN, Taira M, Iijima T, Tsutsui K-I (2009): Segregated and integrated coding
814   of reward and punishment in the cingulate cortex. J Neurophysiol 101:3284–3293.
815 Garrison J, Erdeniz B, Done J (2013): Prediction error in reinforcement learning: A meta-
816   analysis of neuroimaging studies. Neurosci Biobehav Rev 37:1297–1310.
817 Gherman S, Philiastides MG (2015): Neural representations of confidence emerge from the
818   process of decision formation during perceptual choices. NeuroImage 106:134–143.
819 Gherman S, Philiastides MG (2017): Human VMPFC encodes early signatures of confidence
820   in perceptual decisions. bioRxiv:224337.

821 Gläscher J, Daw N, Dayan P, O'Doherty JP (2010): States versus rewards: dissociable neural
822   prediction error signals underlying model-based and model-free reinforcement
823   learning. Neuron 66:585–595.
824 Gläscher J, Hampton AN, O'Doherty JP (2009): Determining a Role for Ventromedial
825   Prefrontal Cortex in Encoding Action-Based Value Signals During Reward-Related
826   Decision Making. Cereb Cortex 19:483–495.
827 Gray JA, McNaughton N (2003): The Neuropsychology of Anxiety: An Enquiry Into the
828   Function of the Septo-hippocampal System. Oxford University Press.
829 Guitart-Masip M, Huys QJM, Fuentemilla L, Dayan P, Duzel E, Dolan RJ (2012): Go and no-go
830   learning in reward and punishment: interactions between affect and effect.
831   NeuroImage 62:154–166.
832 Guo R, Böhmer W, Hebart M, Chien S, Sommer T, Obermayer K, Gläscher J (2016):
833   Interaction of Instrumental and Goal-Directed Learning Modulates Prediction Error
834   Representations in the Ventral Striatum. J Neurosci 36:12650–12660.
835 Haber SN, Kunishio K, Mizobuchi M, Lynd-Balta E (1995): The orbital and medial prefrontal
836   circuit through the primate basal ganglia. J Neurosci Off J Soc Neurosci 15:4851–
837   4867.
838 Hajcak G, Moser JS, Holroyd CB, Simons RF (2006): The feedback-related negativity reflects
839   the binary evaluation of good versus bad outcomes. Biol Psychol 71:148–154.
840 Hall G, Pearce JM (1979): Latent inhibition of a CS during CS-US pairings. J Exp Psychol Anim
841   Behav Process 5:31–42.
842 Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A (2008): Dissociating the Role of the
843   Orbitofrontal Cortex and the Striatum in the Computation of Goal Values and
844   Prediction Errors. J Neurosci 28:5623–5630.
845 Haruno M, Kuroda T, Doya K, Toyama K, Kimura M, Samejima K, Imamizu H, Kawato M
846   (2004): A neural correlate of reward-based behavioral learning in caudate nucleus: a
847   functional magnetic resonance imaging study of stochastic decision task. J Neurosci
848   Off J Soc Neurosci 24:1660–1665.
849 Häusler AN, Oroz Artigas S, Trautner P, Weber B (2016): Gain- and Loss-Related Brain
850   Activation Are Associated with Information Search Differences in Risky Gambles: An
851   fMRI and Eye-Tracking Study. eNeuro 3.
852   https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5032244/.
853 Hayden BY, Heilbronner SR, Pearson JM, Platt ML (2011): Surprise Signals in Anterior
854   Cingulate Cortex: Neuronal Encoding of Unsigned Reward Prediction Errors Driving
855   Adjustment in Behavior. J Neurosci 31:4178–4187.
856 Hester R, Barre N, Murphy K, Silk TJ, Mattingley JB (2008): Human Medial Frontal Cortex
857   Activity Predicts Learning from Errors. Cereb Cortex 18:1933–1940.
858 Hester R, Murphy K, Brown FL, Skilleter AJ (2010): Punishing an Error Improves Learning: The
859   Influence of Punishment Magnitude on Error-Related Neural Activity and Subsequent
860   Learning. J Neurosci 30:15600–15607.
861 Huettel SA, Stowe CJ, Gordon EM, Warner BT, Platt ML (2006): Neural signatures of
862   economic preferences for risk and ambiguity. Neuron 49:765–775.
863 Ide JS, Shenoy P, Yu AJ, Li CR (2013): Bayesian Prediction and Evaluation in the Anterior
864   Cingulate Cortex. J Neurosci 33:2039–2047.
865 Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HEM, Stephan KE
866   (2013): Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory
867   Learning. Neuron 80:519–530.

868 Ikemoto S (2007): Dopamine reward circuitry: Two projection systems from the ventral
869   midbrain to the nucleus accumbens–olfactory tubercle complex. Brain Res Rev
870   56:27–78.
871 Jenkinson M, Bannister P, Brady M, Smith S (2002): Improved optimization for the robust
872   and accurate linear registration and motion correction of brain images. NeuroImage
873   17:825–841.
874 Jensen J, Smith AJ, Willeit M, Crawley AP, Mikulis DJ, Vitcu I, Kapur S (2007): Separate brain
875   regions code for salience vs. valence during reward prediction in humans. Hum Brain
876   Mapp 28:294–302.
877 Jocham G, Brodersen KH, Constantinescu AO, Kahn MC, Ianni AM, Walton ME, Rushworth
878   MFS, Behrens TEJ (2016): Reward-Guided Learning with and without Causal
879   Attribution. Neuron 90:177–190.
880 Kaelbling LP, Littman ML, Moore AW (1996): Reinforcement Learning: A Survey.
881   arXiv:cs/9605103. http://arxiv.org/abs/cs/9605103.
882 Kahnt T (2017): A decade of decoding reward-related fMRI signals and where we go from
883   here. NeuroImage.
884   http://www.sciencedirect.com/science/article/pii/S1053811917304688.
885 Kahnt T, Heinzle J, Park SQ, Haynes J-D (2010): The neural code of reward anticipation in
886   human orbitofrontal cortex. Proc Natl Acad Sci 107:6010–6015.
887 Kahnt T, Tobler PN (2013): Salience Signals in the Right Temporoparietal Junction Facilitate
888   Value-Based Decisions. J Neurosci 33:863–869.
889 Katahira K, Matsuda Y-T, Fujimura T, Ueno K, Asamizuya T, Suzuki C, Cheng K, Okanoya K,
890   Okada M (2015): Neural basis of decision making guided by emotional outcomes. J
891   Neurophysiol 113:3056–3068.
892 Klein-Flügge MC, Hunt LT, Bach DR, Dolan RJ, Behrens TEJ (2011): Dissociable reward and
893   timing signals in human midbrain and ventral striatum. Neuron 72:654–664.
894 Knutson B, Adams CM, Fong GW, Hommer D (2001): Anticipation of increasing monetary
895   reward selectively recruits nucleus accumbens. J Neurosci Off J Soc Neurosci
896   21:RC159.
897 Knutson B, Westdorp A, Kaiser E, Hommer D (2000): FMRI visualization of brain activity
898   during a monetary incentive delay task. NeuroImage 12:20–27.
899 Koch K, Schachtzabel C, Wagner G, Reichenbach JR, Sauer H, Schlösser R (2008): The neural
900   correlates of reward-related trial-and-error learning: An fMRI study with a
901   probabilistic learning task. Learn Mem 15:728–732.
902 Kolling N, Behrens TEJ, Mars RB, Rushworth MFS (2012): Neural Mechanisms of Foraging.
903   Science 336:95–98.
904 Kotz SA, Dengler R, Wittfoth M (2015): Valence-specific conflict moderation in the dorso-
905   medial PFC and the caudate head in emotional speech. Soc Cogn Affect Neurosci
906   10:165–171.
907 Kurniawan IT, Guitart-Masip M, Dayan P, Dolan RJ (2013): Effort and valuation in the brain:
908   the effects of anticipation and execution. J Neurosci Off J Soc Neurosci 33:6160–
909   6169.
910 Laird AR, Fox PM, Price CJ, Glahn DC, Uecker AM, Lancaster JL, Turkeltaub PE, Kochunov P,
911   Fox PT (2005): ALE meta-analysis: controlling the false discovery rate and performing
912   statistical contrasts. Hum Brain Mapp 25:155–164.
913 Lebreton M, Abitbol R, Daunizeau J, Pessiglione M (2015): Automatic integration of
914   confidence in the brain valuation signal. Nat Neurosci 18:1159–1167.

915 Leknes S, Lee M, Berna C, Andersson J, Tracey I (2011): Relief as a Reward: Hedonic and
916   Neural Responses to Safety from Pain. PLOS ONE 6:e17870.

917 Leong YC, Radulescu A, Daniel R, DeWoskin V, Niv Y (2017): Dynamic Interaction between
918   Reinforcement Learning and Attention in Multidimensional Environments. Neuron
919   93:451–463.

920 Li J, Daw ND (2011): Signals in human striatum are appropriate for policy update rather than
921   value prediction. J Neurosci Off J Soc Neurosci 31:5504–5511.

922 Li J, Zhang Y (2006): Prediction error method-based second-order structural identification
923   algorithm in stochastic state space formulation. Earthq Eng Struct Dyn 35:761–779.

924 Lin A, Adolphs R, Rangel A (2012): Social and monetary reward learning engage overlapping
925   neural substrates. Soc Cogn Affect Neurosci 7:274–281.

926 Litt A, Plassmann H, Shiv B, Rangel A (2011): Dissociating valuation and saliency signals
927   during decision-making. Cereb Cortex N Y N 1991 21:95–102.

928 Liu X, Hairston J, Schrier M, Fan J (2011): Common and distinct networks underlying reward
929   valence and processing stages: a meta-analysis of functional neuroimaging studies.
930   Neurosci Biobehav Rev 35:1219–1236.

931 Losecaat Vermeer A, Boksem M, G Sanfey A (2014): Neural mechanisms underlying context-
932   dependent shifts in risk preferences. NeuroImage 103.

933 Luking KR, Luby JL, Barch DM (2014): Kids, candy, brain and behavior: Age differences in
934   responses to candy gains and losses. Dev Cogn Neurosci 9:82–92.

935 Manza P, Hu S, Ide JS, Farr OM, Zhang S, Leung H-C, Li CR (2016): The effects of
936   methylphenidate on cerebral responses to conflict anticipation and unsigned
937   prediction error in a stop-signal task. J Psychopharmacol Oxf Engl 30:283–293.

938 Marsh R, Hao X, Xu D, Wang Z, Duan Y, Liu J, Kangarlu A, Martinez D, Garcia F, Tau GZ, Yu S,
939   Packard MG, Peterson BS (2010): A virtual reality-based FMRI study of reward-based
940   spatial learning. Neuropsychologia 48:2912–2921.

941 Mathys CD, Lomakina EI, Daunizeau J, Iglesias S, Brodersen KH, Friston KJ, Stephan KE
942   (2014): Uncertainty in perception and the Hierarchical Gaussian Filter. Front Hum
943   Neurosci 8. http://journal.frontiersin.org/article/10.3389/fnhum.2014.00825/full.

944 Matsumoto M, Hikosaka O (2009): Two types of dopamine neuron distinctly convey positive
945   and negative motivational signals. Nature 459:837–841.

946 Mattfeld AT, Gluck MA, Stark CEL (2011): Functional specialization within the striatum along
947   both the dorsal/ventral and anterior/posterior axes during associative learning via
948   reward and punishment. Learn Mem Cold Spring Harb N 18:703–711.

949 Maunsell JHR (2004): Neuronal representations of cognitive state: reward or attention?
950   Trends Cogn Sci 8:261–265.

951 McClure SM, Berns GS, Montague PR (2003): Temporal prediction errors in a passive
952   learning task activate human striatum. Neuron 38:339–346.

953 Menon V, Levitin DJ (2005): The rewards of music listening: response and physiological
954   connectivity of the mesolimbic system. NeuroImage 28:175–184.

955 Metereau E, Dreher J-C (2013): Cerebral correlates of salient prediction error for different
956   rewards and punishments. Cereb Cortex N Y N 1991 23:477–487.

957 Metereau E, Dreher J-C (2015): The medial orbitofrontal cortex encodes a general unsigned
958   value signal during anticipation of both appetitive and aversive events. Cortex J
959   Devoted Study Nerv Syst Behav 63:42–54.

960 Meyniel F, Dehaene S (2017): Brain networks for confidence weighting and hierarchical
961   inference during probabilistic learning. Proc Natl Acad Sci 114:E3859–E3868.

962 Middlebrooks PG, Sommer MA (2012): Neuronal correlates of metacognition in primate
963   frontal cortex. Neuron 75:517–530.
964 Montague PR, Dayan P, Sejnowski TJ (1996): A framework for mesencephalic dopamine
965   systems based on predictive Hebbian learning. J Neurosci Off J Soc Neurosci
966   16:1936–1947.
967 Morrens J (2014): Dopamine neurons coding prediction errors in reward space, but not in
968   aversive space: a matter of location? J Neurophysiol 112:1021–1024.
969 Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI (2012): Rational regulation of
970   learning dynamics by pupil-linked arousal systems. Nat Neurosci 15:1040–1046.
971 Nieuwenhuis S, Slagter HA, von Geusau NJA, Heslenfeld DJ, Holroyd CB (2005): Knowing
972   good from bad: differential activation of human cortical areas by positive and
973   negative outcomes. Eur J Neurosci 21:3161–3168.
974 Niv Y, Daniel R, Geana A, Gershman SJ, Leong YC, Radulescu A, Wilson RC (2015):
975   Reinforcement Learning in Multidimensional Environments Relies on Attention
976   Mechanisms. J Neurosci 35:8145–8157.
977 Noonan MP, Mars RB, Rushworth MFS (2011): Distinct roles of three frontal cortical areas in
978   reward-guided behavior. J Neurosci Off J Soc Neurosci 31:14399–14412.
979 O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001): Abstract reward and
980   punishment representations in the human orbitofrontal cortex. Nat Neurosci 4:95–
981   102.
982 O'Doherty J, Critchley H, Deichmann R, Dolan RJ (2003): Dissociating valence of outcome
983   from behavioral control in human orbital and ventral prefrontal cortices. J Neurosci
984   Off J Soc Neurosci 23:7931–7939.
985 O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004): Dissociable Roles
986   of Ventral and Dorsal Striatum in Instrumental Conditioning. Science 304:452–454.
987 O'doherty JP, Hampton A, Kim H (2007): Model-Based fMRI and Its Application to Reward
988   Learning and Decision Making. Ann N Y Acad Sci 1104:35–53.
989 Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011): FieldTrip: Open Source Software for
990   Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Research
991   article. Computational Intelligence and Neuroscience.
992   https://www.hindawi.com/journals/cin/2011/156869/.
993 O'Reilly JX, Schüffelgen U, Cuell SF, Behrens TEJ, Mars RB, Rushworth MFS (2013):
994   Dissociable effects of surprise and model update in parietal and anterior cingulate
995   cortex. Proc Natl Acad Sci 110:E3660–E3669.
996 den Ouden HEM, Kok P, de Lange FP (2012): How Prediction Errors Shape Perception,
997   Attention, and Motivation. Front Psychol 3.
998   http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3518876/.
999 Pagnoni G, Zink CF, Montague PR, Berns GS (2002): Activity in human ventral striatum
1000   locked to errors of reward prediction. Nat Neurosci 5:97–98.
1001 Paschke LM, Walter H, Steimke R, Ludwig VU, Gaschler R, Schubert T, Stelzel C (2015):
1002   Motivation by potential gains and losses affects control processes via different
1003   mechanisms in the attentional network. NeuroImage 111:549–561.
1004 Pearce JM, Hall G (1980): A model for Pavlovian learning: Variations in the effectiveness of
1005   conditioned but not of unconditioned stimuli. Psychol Rev 87:532–552.
1006 Pessiglione M, Petrovic P, Daunizeau J, Palminteri S, Dolan RJ, Frith CD (2008): Subliminal
1007   instrumental conditioning demonstrated in the human brain. Neuron 59:561–567.

1008 Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006): Dopamine-dependent
1009   prediction errors underpin reward-seeking behaviour in humans. Nature 442:1042–
1010   1045.
1011 Philiastides MG, Biele G, Heekeren HR (2010a): A mechanistic account of value computation
1012   in the human brain. Proc Natl Acad Sci U S A 107:9430–9435.
1013 Philiastides MG, Biele G, Vavatzanidis N, Kazzer P, Heekeren HR (2010b): Temporal dynamics
1014   of prediction error processing during reward-based decision making. NeuroImage
1015   53:221–232.
1016 Philiastides MG, Heekeren HR, Sajda P (2014): Human Scalp Potentials Reflect a Mixture of
1017   Decision-Related Signals during Perceptual Choices. J Neurosci 34:16877–16889.
1018 Pisauro A, Fouragnan E, Retzler C, Philiastides M (2017): Neural correlates of evidence
1019   accumulation during value-based decisions revealed via simultaneous EEG-fMRI. Nat
1020   Commun. http://eprints.hud.ac.uk/31834/.
1021 Poudel GR, Innes CRH, Jones RD (2013): Distinct neural correlates of time-on-task and
1022   transient errors during a visuomotor tracking task after sleep restriction.
1023   NeuroImage 77:105–113.
1024 Preuschoff K, Quartz SR, Bossaerts P (2008): Human Insula Activation Reflects Risk
1025   Prediction Errors As Well As Risk. J Neurosci 28:2745–2752.
1026 Queirazza F, Fouragnan E, Steele JD, Cavanagh J, Philiastides M (2017): Dorsomedial
1027   prefrontal cortex activity during reinforcement learning discriminates response to
1028   Cognitive Behavioural Therapy in depression. bioRxiv:224410.
1029 R Garrison J, Done J, S Simons J (2017): Interpretation of published meta-analytical studies
1030   affected by implementation errors in the GingerALE software.
1031 Rescorla R, Wagner A (1972): A theory of Pavlovian conditioning: Variations in the
1032   effectiveness of reinforcement and nonreinforcement. In: Black, A, Prokasy, W,
1033   editors. Classical Conditioning II: Current Research and Theory. Appleton-Century-
1034   Crofts. pp 64–99.
1035 Ribas-Fernandes JJF, Solway A, Diuk C, McGuire JT, Barto AG, Niv Y, Botvinick MM (2011): A
1036   neural signature of hierarchical reinforcement learning. Neuron 71:370–379.
1037 Rodriguez PF (2009): Stimulus-outcome learnability differentially activates anterior cingulate
1038   and hippocampus at feedback processing. Learn Mem Cold Spring Harb N 16:324–
1039   331.
1040 Roesch MR, Olson CR (2004): Neuronal Activity Related to Reward Value and Motivation in
1041   Primate Frontal Cortex. Science 304:307–310.
1042 Rohe T, Noppeney U (2015): Cortical Hierarchies Perform Bayesian Causal Inference in
1043   Multisensory Perception. PLOS Biol 13:e1002073.
1044 Rohe T, Weber B, Fliessbach K (2012): Dissociation of BOLD responses to reward prediction
1045   errors and reward receipt by a model comparison. Eur J Neurosci 36:2376–2382.
1046 Rolls ET, Grabenhorst F, Parris BA (2008): Warm pleasant feelings in the brain. NeuroImage
1047   41:1504–1513.
1048 Rudorf S, Preuschoff K, Weber B (2012): Neural Correlates of Anticipation Risk Reflect Risk
1049   Preferences. J Neurosci 32:16683–16692.
1050 Samson RD, Frank MJ, Fellous J-M (2010): Computational models of reinforcement learning:
1051   the role of dopamine as a reward signal. Cogn Neurodyn 4:91–105.
1052 Sarinopoulos I, Grupe DW, Mackiewicz KL, Herrington JD, Lor M, Steege EE, Nitschke JB
1053   (2010): Uncertainty during anticipation modulates neural responses to aversion in
1054   human insula and amygdala. Cereb Cortex N Y N 1991 20:929–940.

1055   Schlagenhauf F, Rapp MA, Huys QJM, Beck A, Wüstenberg T, Deserno L, Buchholz H-G,
1056   Kalbitzer J, Buchert R, Bauer M, Kienast T, Cumming P, Plotkin M, Kumakura Y, Grace
1057   AA, Dolan RJ, Heinz A (2013): Ventral striatal prediction error signaling is associated
1058   with dopamine synthesis capacity and fluid intelligence. Hum Brain Mapp 34:1490–
1059   1499.
1060   Scholl J, Kolling N, Nelissen N, Wittmann MK, Harmer CJ, Rushworth MFS (2015): The Good,
1061   the Bad, and the Irrelevant: Neural Mechanisms of Learning Real and Hypothetical
1062   Rewards and Effort. J Neurosci Off J Soc Neurosci 35:11233–11251.
1063   Schonberg T, O'Doherty JP, Joel D, Inzelberg R, Segev Y, Daw ND (2010): Selective
1064   impairment of prediction error signaling in human dorsolateral but not ventral
1065   striatum in Parkinson's disease patients: evidence from a model-based fMRI study.
1066   NeuroImage 49:772–781.
1067   Schultz W, Apicella P, Ljungberg T (1993): Responses of monkey dopamine neurons to
1068   reward and conditioned stimuli during successive steps of learning a delayed
1069   response task. J Neurosci Off J Soc Neurosci 13:900–913.
1070   Schultz W, Apicella P, Scarnati E, Ljungberg T (1992): Neuronal activity in monkey ventral
1071   striatum related to the expectation of reward. J Neurosci Off J Soc Neurosci
1072   12:4595–4610.
1073   Schultz W (1998): Predictive Reward Signal of Dopamine Neurons. J Neurophysiol 80:1–27.
1074   Schultz W (2016a): Dopamine reward prediction-error signalling: a two-component
1075   response. Nat Rev Neurosci 17:183–195.
1076   Schultz W (2016b): Dopamine reward prediction error coding. Dialogues Clin Neurosci
1077   18:23–32.
1078   Schultz W, Dayan P, Montague PR (1997): A Neural Substrate of Prediction and Reward.
1079   Science 275:1593–1599.
1080   Schwartenbeck P, FitzGerald THB, Dolan R (2016): Neural signals encoding shifts in beliefs.
1081   NeuroImage 125:578–586.
1082   Scimeca JM, Katzman PL, Badre D (2016): Striatal prediction errors support dynamic control
1083   of declarative memory decisions. Nat Commun 7:ncomms13061.
1084   Serences JT (2008): Value-based modulations in human visual cortex. Neuron 60:1169–
1085   1181.
1086   Sescousse G, Caldú X, Segura B, Dreher J-C (2013): Processing of primary and secondary
1087   rewards: A quantitative meta-analysis and review of human functional neuroimaging
1088   studies. Neurosci Biobehav Rev 37:681–696.
1089   Seymour B, Daw N, Dayan P, Singer T, Dolan R (2007): Differential Encoding of Losses and
1090   Gains in the Human Striatum. J Neurosci Off J Soc Neurosci 27:4826–4831.
1091   Silvetti M, Verguts T (2012): Reinforcement Learning, High-Level Cognition, and the Human
1092   Brain. http://www.intechopen.com/books/neuroimaging-cognitive-and-clinical-
1093   neuroscience/reinforcement-learning-high-level-cognition-and-the-human-brain.
1094   So N, Stuphorn V (2012): Supplementary Eye Field Encodes Reward Prediction Error. J
1095   Neurosci 32:2950–2963.
1096   Späti J, Chumbley J, Brakowski J, Dörig N, Grosse Holtforth M, Seifritz E, Spinelli S (2014):
1097   Functional lateralization of the anterior insula during feedback processing. Hum
1098   Brain Mapp 35:4428–4439.
1099   Spicer J, Galvan A, Hare TA, Voss H, Glover G, Casey B (2007): Sensitivity of the Nucleus
1100   Accumbens to Violations in Expectation of Reward. NeuroImage 34:455–461.

1101 Spoormaker VI, Andrade KC, Schröter MS, Sturm A, Goya-Maldonado R, Sämann PG, Czisch
1102   M (2011): The neural correlates of negative prediction error signaling in human fear
1103   conditioning. NeuroImage 54:2250–2256.
1104 Sutton R (1998): Reinforcement Learning: An Introduction. Cambridge, Mass: MIT Press.
1105 Takemura H, Samejima K, Vogels R, Sakagami M, Okuda J (2011): Stimulus-Dependent
1106   Adjustment of Reward Prediction Error in the Midbrain. PLOS ONE 6:e28337.
1107 Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S (2004): Prediction of
1108   immediate and future rewards differentially recruits cortico-basal ganglia loops. Nat
1109   Neurosci 7:nn1279.
1110 Tanaka SC, Samejima K, Okada G, Ueda K, Okamoto Y, Yamawaki S, Doya K (2006): Brain
1111   mechanism of reward prediction under predictable and unpredictable environmental
1112   dynamics. Neural Netw Off J Int Neural Netw Soc 19:1233–1241.
1113 Tobia MJ, Gläscher J, Sommer T (2016): Context-specific behavioral surprise is differentially
1114   correlated with activity in anterior and posterior brain systems. Neuroreport 27:677–
1115   682.
1116 Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA (2002): Meta-analysis of the functional
1117   neuroanatomy of single-word reading: method and validation. NeuroImage 16:765–
1118   780.
1119 Ullsperger M, Cramon DY von (2003): Error Monitoring Using External Feedback: Specific
1120   Roles of the Habenular Complex, the Reward System, and the Cingulate Motor Area
1121   Revealed by Functional Magnetic Resonance Imaging. J Neurosci 23:4308–4314.
1122 Valentin VV, O'Doherty JP (2009): Overlapping prediction errors in dorsal striatum during
1123   instrumental learning with juice and money reward in the human brain. J
1124   Neurophysiol 102:3384–3391.
1125 Watanabe N, Sakagami M, Haruno M (2013): Reward Prediction Error Signal Enhanced by
1126   Striatum–Amygdala Interaction Explains the Acceleration of Probabilistic Reward
1127   Learning by Emotion. J Neurosci 33:4487–4493.
1128 Wittmann MK, Kolling N, Akaishi R, Chau BKH, Brown JW, Nelissen N, Rushworth MFS
1129   (2016): Predictive decision making driven by multiple time-linked reward
1130   representations in the anterior cingulate cortex. Nat Commun 7:12327.
1131 Wunderlich K, Rangel A, O'Doherty JP (2009): Neural computations underlying action-based
1132   decision making in the human brain. Proc Natl Acad Sci 106:17199–17204.
1133 Wunderlich K, Symmonds M, Bossaerts P, Dolan RJ (2011): Hedging your bets by learning
1134   reward correlations in the human brain. Neuron 71:1141–1152.
1135 Yacubian J, Gläscher J, Schroeder K, Sommer T, Braus DF, Büchel C (2006): Dissociable
1136   systems for gain- and loss-related value predictions and errors of prediction in the
1137   human brain. J Neurosci Off J Soc Neurosci 26:9530–9537.
1138 Yeung N, Sanfey AG (2004): Independent Coding of Reward Magnitude and Valence in the
1139   Human Brain. J Neurosci 24:6258–6264.
1140 Yu AJ, Dayan P (2005): Uncertainty, Neuromodulation, and Attention. Neuron 46:681–692.
1141 Zalla T, Koechlin E, Pietrini P, Basso G, Aquino P, Sirigu A, Grafman J (2000): Differential
1142   amygdala responses to winning and losing: a functional magnetic resonance imaging
1143   study in humans. Eur J Neurosci 12:1764–1770.
1144 Zhang S, Mano H, Ganesh G, Robbins T, Seymour B (2016): Dissociable Learning Processes
1145   Underlie Human Pain Conditioning. Curr Biol 26:52–58.
1146 Zink CF, Pagnoni G, Martin-Skurski ME, Chappelow JC, Berns GS (2004): Human striatal
1147   responses to monetary reward depend on saliency. Neuron 42:509–517.

1148

1149 **Table 1.** Categorisation of fMRI studies into the three RPE components (valence, surprise,

1150 signed RPE) and broken down by the relevant fMRI contrast/regressor.

Statistical comparisons	Number	Total	Reference
Valence *Pattern A i (NEG>POS)*		32	[de Bruijn et al., 2009; Daniel et al., 2011; Demos et al., 2012; van Duijvenvoorde et al., 2014; Elward et al., 2015; Ferdinand and Opitz, 2014; Fouragnan et al., 2015; Gläscher et al., 2009; Haruno et al., 2004; Häusler et al., 2016; Jocham et al., 2016; Kahnt et al., 2010; Katahira et al., 2015; Klein-Flügge et al., 2011; Klein-Flügge et al., 2011; Knutson et al., 2000; Knutson et al., 2001; Koch et al., 2008; Leknes et al., 2011; Losecaat Vermeer et al., 2014; Marsh et al., 2010; Mattfeld et al., 2011; Noonan et al., 2011; O'Doherty et al., 2001; O'Doherty et al., 2003; Rodriguez, 2009; Rolls et al., 2008; Scholl et al., 2015; Seymour et al., 2007; Spicer et al., 2007; Spoormaker et al., 2011; Ullsperger and Cramon, 2003; Yacubian et al., 2006]
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		
Valence *Pattern A ii (POS>NEG)*		33	[Amiez et al., 2012; Aron et al., 2004; Bickel et al., 2009; de Bruijn et al., 2009; Canessa et al., 2013; Daniel et al., 2011; van Duijvenvoorde et al., 2014; Elliott et al., 2000; Ernst et al., 2004; Forster and Brown, 2011; Fouragnan et al., 2015; Fujiwara et al., 2009; Häusler et al., 2016; Hester et al., 2008; Hester et al., 2010; Jocham et al., 2016; Katahira et al., 2015; Knutson et al., 2000; Knutson et al., 2001; Knutson et al., 2001; Kurniawan et al., 2013; Losecaat Vermeer et al., 2014; Luking et al., 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; Spoormaker et al., 2011; Ullsperger and Cramon, 2003]
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		
Surprise *Pattern B*		41	[Allen et al., 2016; Amado et al., 2016; Amiez et al., 2012; Boll et al., 2013; Browning et al., 2010; Chumbley et al., 2014; Daw et al., 2011; Dreher, 2013; Ferdinand and Opitz, 2014; Forster and 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 et al., 2015; Leong et al., 2017; Losecaat Vermeer et al., 2014; Manza et al., 2016; McClure et al., 2003; Metereau and Dreher, 2013; Metereau and Dreher, 2015; Meyniel and Dehaene, 2017; Nieuwenhuis et al., 2005; O'Reilly et al., 2013; den Ouden et al., 2012; Poudel et al., 2013; Rodriguez, 2009; Rohe et al., 2012; Rohe and Noppeney, 2015; Rohe and Noppeney, 2015; Rolls et al., 2008; Schwartenbeck et al., 2016; Silvetti and Verguts, 2012; Tobia et al., 2016; Watanabe et al., 2013; Wunderlich et al., 2009; Wunderlich et al., 2011; Yacubian et al., 2006; Zalla et al., 2000; Zhang et al., 2016]
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		
Signed RPE *Pattern C*		38	[Abler et al., 2006; Behrens et al., 2007; van den Bos et al., 2012; Cohen and Ranganath, 2007; Daw et al., 2011; Delgado et al., 2000; Delgado, 2007; Diederen et al., 2017; Diuk et al., 2013; Dunne et al., 2016; Gläscher et al., 2010; Guo et al., 2016; Hare et al., 2008; Ide et al., 2013; Katahira et al., 2015; Leong et al., 2017; Li and Zhang, 2006; Lin et al., 2012; Mattfeld et al., 2011; McClure et al., 2003; Metereau and Dreher, 2013; Metereau and Dreher, 2015; O'Doherty et al., 2003; Pessiglione et al., 2006; Pessiglione et al., 2008; Ribas-Fernandes et al., 2011; Rolls et al., 2008; Schlagenhauf et al., 2013; Schonberg et al., 2010; Scimeca et al., 2016; Seymour et al., 2007; Takemura et al., 2011; Tanaka et al., 2004; Tanaka et al., 2006; Valentin and O'Doherty, 2009; Watanabe et al., 2013; Wunderlich et al., 2011]
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		

1151

1152

1153   **Table 2.** ALE cluster results for the valence analysis: Pattern A (i) and (ii) (FDR-ID P < 0.05,

1154   with a minimum volume cluster size of 50 mm^3.

Region	R/L	x	y	z	Cluster size	ALE score
Pattern A (i) NEG > POS						
Dorsomedial cingulate cortex (dMCC)	R	2	24	36	12712	0.051
Anterior Insula (aINS)	R	32	24	-2	6120	0.062
-	L	-32	22	-4	4880	0.056
Pallidum	R	12	8	4	3360	0.04
-	L	-14	6	2	2520	0.029
Middle Frontal Gyrus	R	38	4	32	3152	0.029
-	R	30	10	56	488	0.021
-	L	-28	12	60	104	0.019
Inferior Parietal Lobule (IPL)	R	40	-48	42	2416	0.039
-	L	-38	-48	42	2216	0.043
Middle Temporal Gyrus (MTG)	R	60	-28	-6	1192	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
Pattern A (ii) (POS > NEG)						
Ventral striatum (vSTR)	L	-12	8	-4	4880	0.052
-	R	8	8	-2	2880	0.038
Ventromedial Prefrontal Cortex (vmPFC)	L	-2	42	0	3416	0.037
Posterior Cingulate Cortex (PCC)	L	0	-32	36	240	0.016
-	L	0	-36	26	88	0.014
Ventrolateral OFC (vlOFC)	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

1155

1156

1157   **Table 3.** ALE clusters results for the surprise analysis (FDR-ID P < 0.05, with a minimum

1158   volume cluster size of 50 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

1159

1160

1161 **Table 4.** ALE cluster results for the conjunction analysis of valence and surprise (FDR-ID p <

1162 0.05, with a minimum volume cluster size of 50 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

1163

1164 **Table 5.** ALE cluster results for the contrast analyses of valence and surprise (FDR-pN p <

1165 0.05, with a minimum volume cluster size of 50 mm^3).

Region	R/L	x	y	z	Cluster size	ALE score
Valence vs. Surprise						
Ventral Striatum (vSTR)	L	-10	8	-10	1096	3.29
ventromedial prefrontal cortex (vmPFC)	L	-2	44	0	256	3.29
Positive vs. Surprise						
Ventral Striatum (vSTR)	L	-12	-8	-8	1872	3.29
ventromedial prefrontal cortex (vmPFC)	R	0	46	0	512	3.29
Ventral Striatum (vSTR)	R	8	8	-6	168	3.29
Negative vs. Surprise						
Middle Insula (mINS)	R	40	10	2	544	3.29
Mid Cingulate Cortex (MCC)	R	6	20	42	144	3.29
Surprise vs. Valence						
Anterior Insula (aINS)	R	32	24	-4	1224	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
Surprise vs. Positive						
Anterior Insula (aINS)	R	32	22	-2	1648	3.29
Middle Temporal Gyrus (MTG)	R	40	-46	42	1184	3.29
Anterior Insula (aINS)	L	-32	22	-2	1016	3.29
Inferior Frontal Gyrus	R	52	10	18	184	3.29
Supplementary Motor Area (SMA)	L	-2	12	52	160	3.29
Surprise vs. Negative						
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

1166

1167 **Table 6.** ALE clusters results for the signed RPE studies (FDR-ID p < 0.05, with a minimum

1168 volume cluster size of 50 mm^3).

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

1169

1170 **Table 7.** ALE cluster results for the contrast analyses of signed RPE and valence as well as

1171 signed RPE and surprise (FDR-pN p < 0.05, with a minimum volume cluster size of 50 mm^3).

Region	R/L	x	y	z	Cluster size	ALE score
Positive – Signed RPE						
Ventromedial Prefrontal Cortex (vmPFC)	R	2	44	-15	160	3.29
Signed RPE - Positive						
No significant						
Negative – Signed RPE						
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
Signed RPE - Negative						
Ventral Striatum (vSTR)	R	10	10	-6	2208	3.29
Valence – Signed RPE						
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
Signed RPE - Valence						
Ventral Striatum (vSTR)	R	12	16	-2	184	3.29
Surprise – Signed RPE						
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

Signed RPE - Surprise						
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

1172

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1173 **Figure Legends**

1174

1175 **Figure 1.** Hypothesized profiles for BOLD responses as function of the three RPE

1176 components. Pattern A (i and ii) describe the two categorical valence responses (orange and

1177 blue colours indicate (i) responses being greater for negative compared to positive outcomes

1178 [NEG > POS] and (ii) responses being greater for positive compared to negative outcomes

1179 [POS > NEG]). Pattern B captures surprise effects with greater responses to higher outcome

1180 deviations from expectations, independent of the sign (valence) of the RPE. Pattern C shows

1181 a monotonically increasing response profile consistent with a signed RPE representation.

1182

1183 **Figure 2.** Results of whole-brain ALE analysis along the valence component. Overlays of

1184 brain areas activated by correlations with NEG > POS (blue) and POS > NEG (orange)

1185 (Pattern A (i) and (ii), respectively; Fig. 1) (P-values corrected with FDR-ID [FID] and FDR-

1186 pN [FRN] < 0.05 and a minimum cluster volume of 50 mm^3). Representative slices are

1187 shown with MNI coordinates given below each image.

1188

1189 **Figure 3.** Results of the whole brain ALE analysis for the surprise component of RPE

1190 (pattern B, Figure 1). Overlay of brain areas activated by all analyses representing direct or

1191 indirect measures of the surprise component of RPE (P-values corrected with FDR-ID [FID]

1192 and FDR-pN [FRN] < 0.05 and a minimum cluster volume of 50 mm^3). Representative slices

1193 are shown with MNI coordinates given below each image.

1194

1195 **Figure 4.** Results of the ALE conjunction analysis between valence and surprise (purple).

1196 The regions identified earlier with separate ALE analyses along the valence (NEG > POS:

1197 blue, POS > NEG: orange) and surprise (green) components are shown for comparison

1198 purposes. P-values were corrected with FDR-pN [FRN] < 0.05 and a minimum cluster

1199 volume of 50 mm^3 for the initial maps. Representative slices are shown with MNI coordinates

1200 given bellow each image.

1201 **Figure 5.** Results of the ALE contrast analyses for [valence – surprise] (left panel) and

1202 [surprise – valence]. P-values were corrected with FDR-pN [FRN] < 0.05 and a minimum

1203 cluster volume of 50 mm^3 for the initial maps. Representative slices are shown with MNI

1204 coordinates given bellow each image.

1205

1206 **Figure 6.** Results of whole brain ALE analysis for signed RPE. Overlay of brain areas

1207 activated by positive correlation with signed RPE (P-values corrected with FDR-ID [FID] and

1208 FDR-pN [FRN] < 0.05 and a minimum cluster volume of 50 mm^3). Representative slices are

1209 shown with MNI coordinates given bellow each image.

1210

1211 **Figure 7.** Results of the ALE conjunction analysis for all components of RPE. Overlay of

1212 brain areas individually activated by (1) valence (orange), (2) surprise (green), and (3)

1213 signed RPE (red), with P-values corrected with FDR-pN [FRN] < 0.05 and a minimum cluster

1214 volume of 50 mm^3 for the initial maps. Importantly, the overlap between the three analyses,

1215 shown in white, also corresponds to the only cluster found for the ALE conjunction analysis

1216 between valence/surprise vs. signed RPE. MNI coordinates are given below each image.

1217

1218 **Figure 8.** Results of the ALE contrast analyses for [signed RPE – positive valence] (left

1219 panel), [signed RPE – negative valence] (middle panel) and [signed RPE – (positive +

1220 negative valence)] (right panel). P-values were corrected with FDR-pN [FRN] < 0.05 and a

1221 minimum cluster volume of 50 mm^3 for the initial maps. Representative slices are shown with

1222 MNI coordinates given bellow each image.
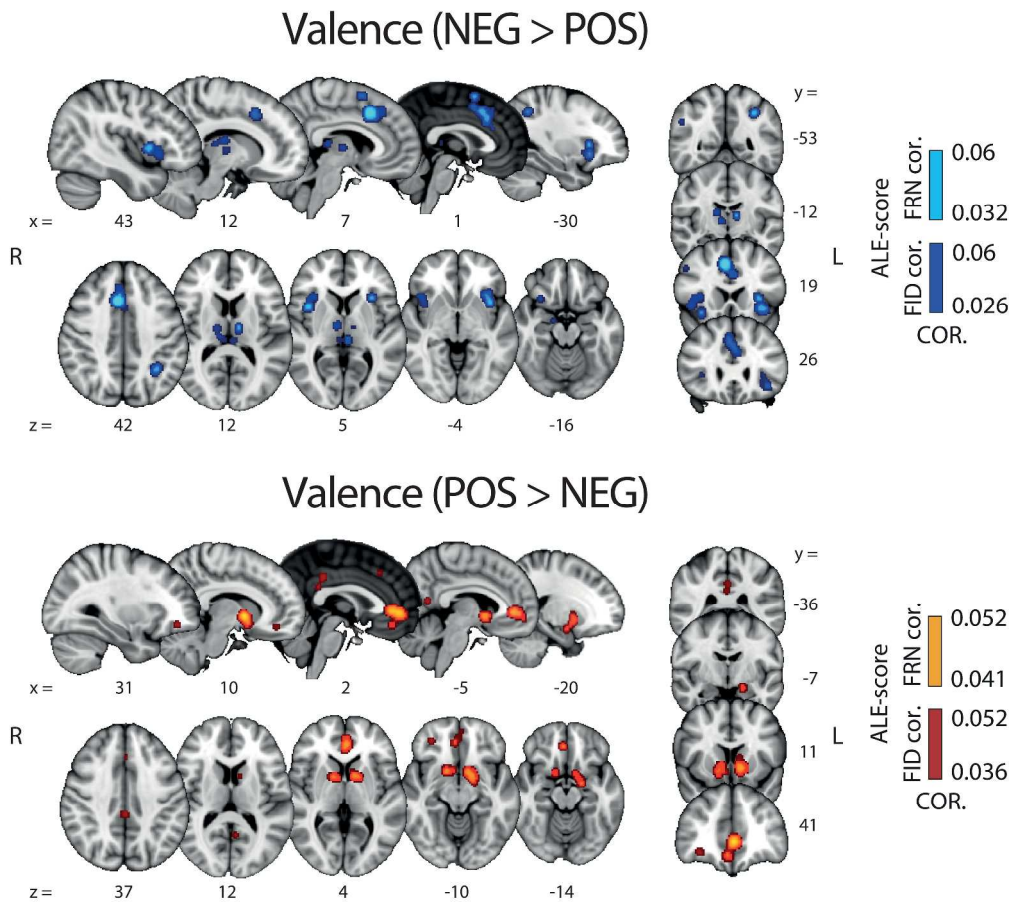
1223

fig.1

91x81mm (300 x 300 DPI)

Valence (NEG > POS)

Valence (POS > NEG)

fig.2

378x339mm (300 x 300 DPI)

fig.4

203x100mm (300 x 300 DPI)

Conjunction Surprise & Valence

fig.4

285x172mm (300 x 300 DPI)

Valence - Surprise

x = -10   -5   41

y = 43
7
-14

z = -14   -3   11

R

■ Valence (pos+neg) > Surprise
■ Positive > Surprise
■ Negative > Surprise

ALE score
FRN corrected

Surprise - Valence

x = -1   39   49

y = 22
8
-45

z = 2   34   41

L

■ Surprise > Valence (pos+neg)
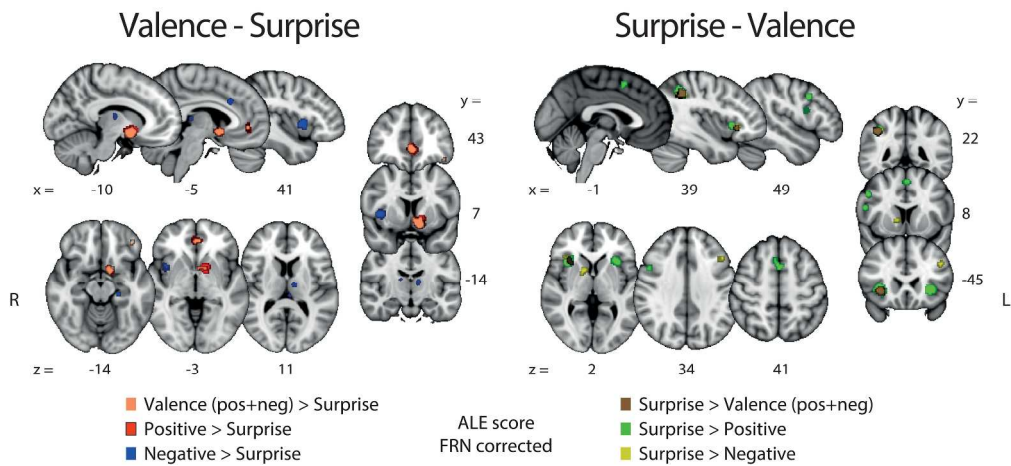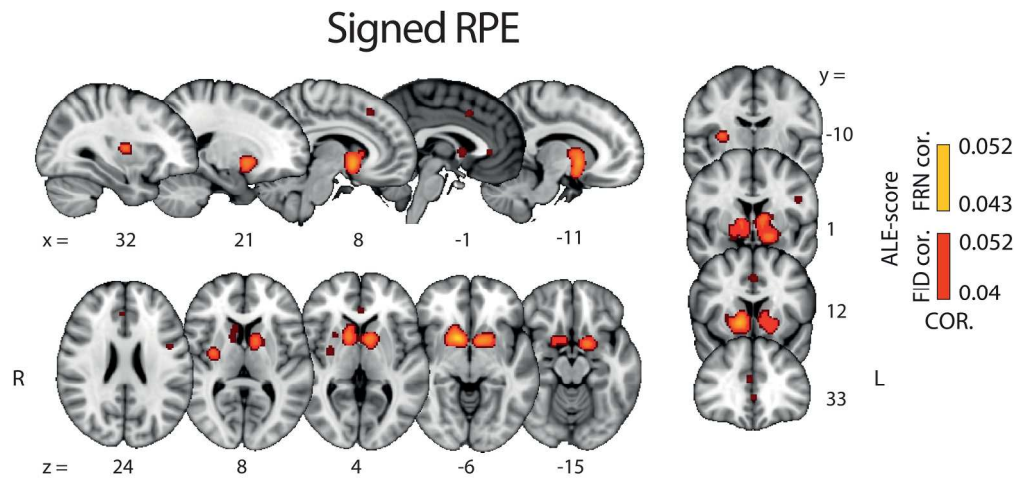■ Surprise > Positive
■ Surprise > Negative

fig.5

252x115mm (300 x 300 DPI)

fig.6

207x98mm (300 x 300 DPI)

Conjunction Surprise & Valence & Signed PE

fig.7

164x66mm (300 x 300 DPI)

Positive - signed RPE

vmPFC

■ Positive > Signed RPE
■ Signed RPE > Positive (non sig.)

Negative - signed RPE

dMCC

STR

THAL

mINS

■ Negative > Signed RPE
■ Signed RPE > Negative

all valence - signed RPE

dMCC

STR

vmPFC

mINS

■ Valence > Signed RPE
■ Signed RPE > Valence

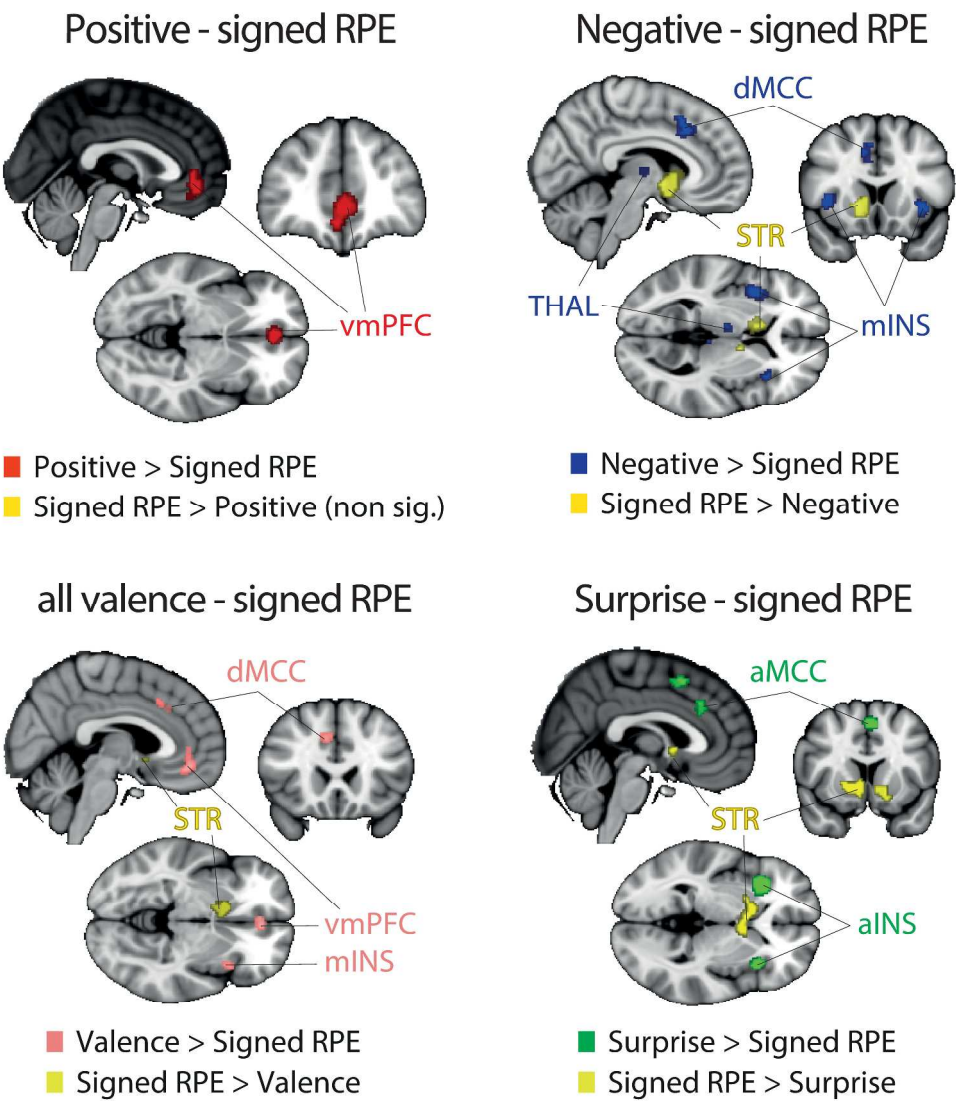## Surprise - signed RPE

aMCC

STR

aINS

■ Surprise > Signed RPE
■ Signed RPE > Surprise

fig.8

366x400mm (300 x 300 DPI)