

1 **Dorsomedial prefrontal cortex activity during learning discriminates**
2 **response to Cognitive Behavioural Therapy in depression.**

3

4 Filippo Queirazza^{1,*}, Elsa Fouragnan^{1,2}, J. Douglas Steele³, Jonathan Cavanagh^{4,5} and
5 Marios G. Philiastides^{1,5,*}

6

7 **Affiliations:**

8 ¹Institute of Neuroscience & Psychology, University of Glasgow, Glasgow, UK

9 ²Department of Experimental Psychology, University of Oxford, Oxford, UK

10 ³Division of Neuroscience, University of Dundee, Dundee, UK

11 ⁴Sackler Centre for Psychobiological Research, Institute of Health & Wellbeing,
12 University of Glasgow, Glasgow, UK

13 ⁵Co-senior author

14

15 ***Correspondence:** f.queirazza.1@research.gla.ac.uk; Marios.Philiastides@gla.ac.uk

16 Institute of Neuroscience & Psychology, 58 Hillhead street, University of Glasgow,
17 Glasgow, UK, G12 8QB

18

19 Number of words (Abstract): 244

20 Number of words (Main text): 4000

21 Number of figures: 7

22 Number of supplemental information: 1

23

24 **Keywords:** depression; self-help CBT; fMRI; computational psychiatry;
25 reinforcement learning.

26

27 **Short title:**

28 Predictors of CBT response in depression

29

30 **ABSTRACT**

31 **BACKGROUND:** Cognitive behavioural therapy (CBT) is an effective evidence
32 based treatment for depression. At present there is no reliable predictor of CBT in
33 depression. Although the key to successful CBT in depression lies in altering
34 maladaptive information processing, no previous imaging study has probed predictors
35 of CBT response using pre-treatment neural encoding of information processing.

36 **METHODS:** Using functional magnetic resonance imaging we scanned 37
37 unmedicated depressed subjects before and after completing computerised CBT
38 (cCBT). We model the trial-by-trial appraisal of feedback information during a
39 probabilistic learning task by means of a dynamic learning rate. To discriminate
40 response to cCBT we capitalise on the pre-treatment blood oxygen level dependent
41 (BOLD) activity encoding the dynamic learning rate as a function of feedback
42 congruence and valence. Additionally, we probe between-group differences in the
43 learning style encoded in the model's parameters.

44 **RESULTS:** We show BOLD activity in the dorsomedial prefrontal cortex (dmPFC) to
45 be encoding the dynamic learning rate. Crucially, responders exhibit greater BOLD
46 activity in the dmPFC during incongruent negative trials but lower BOLD activity
47 during congruent negative trials than non-responders. Additionally, on between-group
48 comparisons of model's parameter estimates we show responders take relatively
49 greater account of previous feedback history and make comparatively smaller
50 adjustments to the learning rate as a result of outcome surprisingness.

51 **CONCLUSIONS:** Our findings provide novel and important insights into the
52 cognitive mechanisms underpinning response to cCBT and lend support to the
53 feasibility and validity of neurocomputational approaches to treatment prediction
54 research in psychiatry.

55

56 **Keywords:** depression; self-help CBT; fMRI; computational psychiatry;

57 reinforcement learning; dynamic learning rate.

58

59 INTRODUCTION

60 The cognitive model of depression posits that biased acquisition and processing of
61 feedback information gives rise to and perpetuates depressive symptoms (1, 2). Based
62 on this theoretical formulation, Aaron T. Beck developed Cognitive Behavioural
63 therapy (CBT) as a treatment for depression (3). CBT is an effective evidence-based
64 intervention for depressive disorder (4, 5). In the UK computerised CBT (cCBT) (that
65 is, self-help internet-delivered CBT) is recommended as a treatment option for mild to
66 moderate depression in the National Institute for Health and Care Excellence
67 guidelines.

68

69 Despite recent multidisciplinary efforts (6-12), there is no reliable predictor of clinical
70 response to CBT in depression (13, 14). Neural predictors using functional magnetic
71 resonance imaging (fMRI) provide the additional advantage of illuminating the
72 functional neuroanatomy that underpins response to CBT in depression (15).

73

74 Remarkably, although the clinical practice of CBT in depression primarily involves
75 evaluating and ultimately correcting negatively biased inferences drawn from
76 probabilistic information, only emotion eliciting (6, 16, 17) or task-free resting-state
77 paradigms (8, 18, 19) have so far been employed to probe pre-treatment fMRI
78 predictors of CBT response.

79

80 In contrast, the computational framework of value-based learning (also known as
81 reinforcement learning) paradigms affords a rigorous and quantitative account of the
82 cognitive mechanisms implicated in drawing inferences from probabilistic feedback.
83 Crucially, within this framework, a weighting factor known as the learning rate

84 controls how the value of unexpected feedback information (also know as the
85 prediction error) is appraised. (20). In experimental settings mimicking real-life,
86 volatile environments, adaptive value-based learning is implemented via a dynamic
87 (that is, time varying) rather than constant learning rate (21-25). Abnormal tuning of
88 the dynamic learning rate distorts processing of probabilistic feedback and impairs
89 learning (26). Prior behavioural evidence suggests that tuning of the dynamic learning
90 rate is disrupted in depression as a function of feedback congruence and valence (27,
91 28).

92

93 In this work, we capitalise on the neural encoding of the dynamic learning rate during
94 probabilistic learning to discriminate response to cCBT. Since successful CBT
95 involves cognitive restructuring of negatively biased thinking we predicted processing
96 of probabilistic negative feedback would uncover between-group differential pre-
97 treatment BOLD activity. Additionally, we found that a learning style that facilitates
98 reframing of negative cognitive biases was associated with response to cCBT.

99

100

101 **METHODS AND MATERIALS**

102 *Sample*

103 All 37 participants (18 women) were recruited via self-referral through local
104 newspaper advertisement. Eligibility criteria were a primary diagnosis of depressive
105 disorder as operationalised by ICD-10 diagnostic criteria and a score ≥ 14 on the
106 Beck's Depression Inventory-II (BDI-II). To avoid any potential confound associated
107 with psychotropic medications, we only recruited unmedicated depressed subjects.
108 Exclusion criteria included current involvement with other CBT-based interventions
109 or psychological therapies, a comorbid diagnosis of other major mental disorder, CBT
110 treatment within past 3 years, a diagnosis of psychoactive substance dependence,
111 previous history of brain injury.

112

113 Eleven subjects (~29.7%) did not attend the post-treatment assessment. Of these
114 subjects, 6 deteriorated and required treatment with antidepressant medications and 5
115 did not complete cCBT due to lack of efficacy. High dropout rates are common in
116 studies examining Internet delivered psychotherapies (12). Twenty-six subjects
117 (~70.3%) attended the post-treatment appointment and of these only one was unable
118 to undergo scanning.

119

120 In total 19 subjects were classified as responders and 18 subjects were classified as
121 non-responders. The overall cCBT response rate was 51.3%. The average post-
122 treatment improvement in BDI-II score was around 62% ($\pm 40\%$).

123

124 Between-group comparisons revealed no significant differences in age ($t_{35}=0.08$;
125 $p=0.93$) and sex ($\chi^2_1=0.02$; $p=0.86$). Non-responders had a significantly higher pre-

126 treatment BDI-II score than responders ($t_{35}=2.86$; $p=0.006$). On further sensitivity
127 analyses we found this significant difference to be mainly determined by those non-
128 responders who did not complete cCBT ($t_{27.3}=4.74$; $p<0.001$) rather than by those who
129 did ($t_{24}=0.62$; $p=0.53$). In spite of these between-group differences in pre-treatment
130 BDI-II scores, we wanted to probe differences in the neurocomputational mechanisms
131 underlying probabilistic learning in order to expose the internal cognitive mechanisms
132 implicated in response to cCBT.

133

134 All participants provided written, informed consent. The study protocol was approved
135 by the West of Scotland Ethics Committee (10/S0703/71).

136

137 *Experimental design*

138 This study adopted a naturalistic longitudinal design. Participants attended two
139 appointments (pre- and two months post-treatment with cCBT) (see Supplemental
140 Materials for further details on cCBT and study design). Each appointment included a
141 clinical evaluation by a qualified psychiatrist followed by an fMRI scan. A clinical
142 diagnosis of depression was corroborated using the Clinical Interview Schedule –
143 Revised (CIS-R) (29). To measure depression severity we used the BDI-II, which is a
144 clinically-validated tool to assess intensity of depression (30).

145

146 We regarded non-completion of cCBT as an index of treatment failure and classified
147 all non-completers as non-responders. For the remaining subjects response to cCBT
148 was defined as a 50% or greater reduction in the pre-treatment BDI-II score. To
149 account for response bias we performed sensitivity analyses excluding non-
150 completers from the non-responders group.

151

152 *fMRI experimental paradigm*

153 To probe the neural correlates of probabilistic reinforcement learning we employed a
154 probabilistic reversal-learning task during fMRI (Figure 1). This task involved
155 learning which of two stimuli yielded the highest payoff rate and included 180 trials
156 lasting approximately 20 minutes. Participants were required to choose between two
157 abstract visual stimuli both yielding either positive (+10) or negative (-10) outcome
158 devoid of monetary value. Stimulus-outcome contingencies were probabilistic and
159 asymmetrically skewed (70-30%) (Figure 1).

160

161 Furthermore, to create a volatile environment stimulus-outcome contingencies were
162 reversed in the course of the experiment. Notably, reversals were triggered when
163 participants chose the high probability stimulus (that is, the stimulus with a greater
164 chance of yielding a positive outcome) five times over the last six trials. As a result
165 participants experienced a different number of reversals.

166

167 Participants were advised of the probabilistic nature of the task and that stimulus-
168 outcome contingencies might reverse based upon their performance. Moreover, to
169 ensure participants understood the nature of the task they underwent a 5 minutes
170 practice session prior to the fMRI scan. The task was programmed using
171 Presentation® (Neurobehavioural Systems) stimulus delivery software.

172

173 The probabilistic reversal-learning task is ideally suited to exposing information
174 processing biases during probabilistic learning. Ultimately optimal performance rests
175 upon attaching a greater weight (via a greater learning rate) to more informative

176 *congruent* feedback (that is, feedback that is consistent with the most likely outcome
177 associated with the chosen stimulus) than to less informative *incongruent* feedback
178 (that is, feedback that is not consistent with the most likely outcome associated with
179 the chosen stimulus). To do so subjects need to be able to infer whether fluctuations
180 in observed stimulus-outcome associations reflect either noise (given underlying
181 stochastic contingencies) or sudden environmental changes (that is, reversals).

182

183 In depression there is substantial behavioural evidence that both congruence (27, 28)
184 and the affective quality (i.e. valence) (2) of probabilistic feedback significantly
185 contribute to biasing information processing. Accordingly, in our analyses we
186 leverage the combined effect of congruence and valence on the dynamic learning rate
187 to probe behavioural and neural differences between responders and non-responders.

188

189 ***Behavioural statistical analysis***

190 To verify that individual performance was better than chance we used a binomial test
191 to compare the number of correct choices with chance level.

192

193 To examine for between-group differences in task performance we regressed subject-
194 wise percentage of high probability stimulus choices on clinical outcome after
195 adjusting for pre-treatment BDI-II score. Additionally we tested for any between-
196 group difference in the effect of the high probability stimulus on choice behaviour
197 (the higher the effect of the high probability stimulus the more optimal the choice
198 behaviour and thus learning) using a generalised mixed-effects linear model (see
199 Supplemental Materials for further details on this analysis).

200

201 Finally we tested for any between-group difference in the model-derived dynamic
202 learning rate estimates as a function of congruence and valence (see Supplemental
203 Materials for further details on this analysis).

204

205 *Computational modelling of behavioural data*

206 We fitted 5 different models (see Supplementary Materials for further details) that use
207 trial-wise scaling of the prediction error but make different assumptions on the
208 computational mechanisms supporting this dynamic tuning of learning rate.

209

210 In the winning model the learning rate scales with the slope of the smoothed unsigned
211 prediction error (i.e. absolute value of prediction error) (24). This model reprises
212 Pearce-Hall's theory that surprise (formalised as the unsigned prediction error) drives
213 the acquisition of stochastic stimulus-outcome contingencies (31) but with some
214 important refinements. Indeed, compared to the Pearce-Hall's model, the smoothing
215 of the unsigned prediction error (the degree of which is regulated by a free
216 parameter ρ) should render the inference process about whether a change has
217 occurred in the environment more robust to the inherent task stochasticity. Moreover,
218 an additional free parameter γ controls the extent to which the dynamic updating of
219 the learning rate is influenced by the slope. For example, whilst lower values of γ
220 yield substantial trial-by-trial changes of the dynamic learning rate even in the
221 presence of small slope estimates (that is, low surprise), higher values of γ result in a
222 more stable learning rate even in the presence of significant slope estimates (that is,
223 high surprise). Hence, this model also allows for the possibility that subjects might be
224 employing a relatively fixed learning rate. The decision function for all models was a

225 standard sigmoid function parameterised by the inverse of the temperature parameter,
226 β .

227

228 ***Model fitting and model comparison***

229 To optimise each model's free parameters we implemented the hierarchical type II
230 maximum likelihood fitting procedure described in (32) (see Supplemental Materials
231 for further details). The fits of all models were compared using the Integrated
232 Bayesian Information Criterion (BIC_{int}) (32) (see Supplemental Materials for further
233 details).

234

235 To ascertain any between-group differences in the learning style encoded in the
236 model's fixed parameters we regressed subject-wise parameter estimates against
237 clinical outcome (response vs. non-response) after adjusting for pre-treatment BDI-II
238 score. To account for the response bias we performed additional sensitivity analyses.

239

240 Finally we ran sanity checks on the winning model's goodness of fit. To verify
241 accuracy of model's fit we first binned predicted choice propensities according to
242 their quintiles and subsequently measured the strength of their linear association with
243 corresponding observed choice probabilities using Pearson's correlation coefficient.
244 Additionally, we employed a binomial test to test whether the number of choices
245 correctly predicted by the model exceeded that expected by chance (33).

246

247 ***fMRI data acquisition***

248 We used a 3T GE system with an 8-channel parallel imaging head coil. We acquired a
249 high-resolution T_1 - weighted structural image (0.5 x 0.5 x 1 mm voxels, 320 x 320

250 matrix, 160 axial slices, TI = 500 ms, TR = 7700 ms, TE = 1.5 ms, flip angle = 12°)
251 using an optimized Inversion Recovery Fast SPOiled GRAdient echo sequence (IR-
252 FSPGR) and a functional echo planar imaging (EPI) scan (3 mm isotropic voxels, 64
253 x 64 matrix, 608 axial slices, TR = 2000 ms, TE = 30 ms, flip angle = 80°). Slice
254 orientation was tilted -20° from the AC-PC plane to alleviate signal drop out in the
255 orbitofrontal cortex (34). The first four volumes of the functional scan were discarded
256 in order to allow for the magnetic field to reach the steady state.

257

258 *fMRI data preprocessing and statistical analysis*

259 Pre-treatment and post-treatment fMRI data preprocessing and statistical analyses
260 were performed using FSL software (35). Preprocessing pipeline involved intra-
261 modal motion correction using MCFLIRT (36), slice timing correction, spatial
262 smoothing with an isotropic 5 mm FWHM Gaussian kernel, high-pass temporal
263 filtering with 110 sec. cut-off frequency and grand-mean intensity normalisation of
264 each entire 4D dataset. Functional scans were subsequently co-registered with skull-
265 stripped structural images using boundary-based registration (FLIRT) (37, 38) and
266 spatially normalised into MNI152 space using FNIRT non-linear registration.

267

268 Whole brain statistical analyses of pre-treatment fMRI data were performed using a
269 multilevel mixed-effects approach as implemented in FLAME1 (FSL) (39). At the
270 first-level the subject-specific general linear model included the model-derived
271 dynamic learning rate as the regressor of interest plus a number of additional nuisance
272 covariates (see Supplementary Materials for further details). Most importantly, we
273 included a nuisance regressor accounting for surprise (here formalised as unsigned

274 prediction error) to retrieve the unique effect of the dynamic learning rate on BOLD
275 activity.

276

277 To improve efficiency of our fMRI statistical analysis all model-derived regressors
278 were obtained by generating subject-wise model fits using the population-level
279 parameter means (40). All regressors were convolved with a hemodynamic response
280 function (double gamma function).

281

282 We estimated the subject-wise linear contrasts of parameter estimates and
283 subsequently entered these contrast images into a second-level mixed-effects analysis
284 where we tested both mean group effect and between-group (responders vs. non-
285 responders) differences. We thresholded the resulting Z statistic images using cluster-
286 defining threshold of $Z > 3.1$ and a FWE-corrected significance threshold of $p = 0.05$.

287

288 Furthermore, we examined for any between-group difference in the pre-treatment
289 BOLD activity encoding the dynamic learning rate. To this end we first retrieved
290 BOLD percent signal change time-locked to the onset of the outcome phase from a
291 cluster broadly corresponding to the dmPFC (max $Z = 7.25$; $-4 -2 56$). We chose this
292 cluster since it is consistent with previous reports on the fMRI correlates of the
293 dynamic learning rate (24, 41). Subsequently we performed between-group
294 comparisons as a function of feedback congruence and valence using mixed-effects
295 linear models (see Supplemental Materials for further details on this analysis).

296

297 We also assessed for a robust linear association between subject-wise (that is,
298 aggregated up to the subject-level) pre-treatment event-locked BOLD percent signal
299 change and post-treatment symptomatic improvement.

300

301 Having ascertained between-group differences in pre-treatment BOLD activity we
302 sought to determine whether such BOLD activity was either a moderator or mediator
303 of cCBT response. (see Supplemental Materials for further details on this analysis).

304

305

306 **RESULTS**

307 *Clinical outcome*

308 Between-group comparisons revealed no significant differences in age ($t_{35}=0.08$;
309 $p=0.93$) and sex ($\chi^2_1=0.02$; $p=0.86$) between responders and non-responders. Non-
310 responders had a significantly higher pre-treatment BDI-II score than responders
311 ($t_{35}=2.86$; $p=0.006$). On further sensitivity analyses we found this significant
312 difference to be mainly determined by those non-responders who did not complete
313 cCBT ($t_{27,3}=4.74$; $p<0.001$) rather than by those who did ($t_{24}=0.62$; $p=0.53$). In spite
314 of these between-group differences in pre-treatment BDI-II scores, we wanted to
315 probe differences in the neurocomputational mechanisms underlying probabilistic
316 learning in order to expose the internal cognitive mechanisms implicated in response
317 to cCBT.

318

319 *Task performance and behavioural estimates of the dynamic learning rate do* 320 *not discriminate response to cCBT*

321 All participants did not choose randomly (binomial test, $\mu=0.5$, $p<0.05$) during the
322 task. We did not find any significant between-group difference in the number of
323 buffer trials ($t_{35}=-0.41$; $p=0.68$) during the pre-treatment experiment. Task
324 performance was similar across responders and non-responders (see Supplemental
325 Results for further details).

326

327 Pre-treatment behavioural estimates of the dynamic learning rate as a function of
328 congruence and valence did not discriminate between responders and non-responders
329 (congruent positive: $t_{38,77}=-0.24$, $p=0.80$; incongruent positive: $t_{37,2}=0.45$, $p=0.65$;

330 congruent negative: $t_{38.07}=0.30$, $p=0.76$; incongruent negative: $t_{159.97}=-0.41$, $p=0.67$)
331 (Figure 2).

332

333 ***Model comparison***

334 On formal Bayesian model comparison we found that the best fitting model was that
335 described in (24) with a BIC_{int} of ~ 287 (32). (Krugel et al. model with additional
336 parameter α^1 : $BIC_{int}=540.82$; Hierarchical Gaussian Filter: $BIC_{int}=828.68$; Pearce-
337 Hall: $BIC_{int}=332.72$; Kalman filter K1 variant: $BIC_{int}=708.70$) (Figure 3A).

338

339 We verified the model's goodness of fit using a binomial test and found that under the
340 null hypothesis that on each trial the model was choosing at chance level, the
341 probability of model's correctly predicted n choices was <0.05 across all subjects.
342 Additionally we found that observed and model's predicted choice probabilities were
343 significantly correlated ($r = 0.88$; $p < 0.001$), further endorsing the quality of the model
344 fits (Figure 3B).

345

346 ***Parameters reflecting specific information processing style are associated*** 347 ***with differential response to cCBT***

348 After adjusting for pre-treatment BDI-II score, we found a statistically significant
349 between-group difference in the estimates of the model's parameter ρ ($\text{logit}(\rho)$: $t_{34}=-$
350 2.15 ; $p=0.038$) and γ ($\text{log}(\gamma)$: $t_{34}=2.11$; $p=0.041$) (Figure 4) but no significant
351 difference for β ($\text{log}(\beta)$: $t_{34}=0.61$; $p=0.54$). Additional sensitivity analyses revealed
352 these significant differences to be predominantly driven by non-responders who did
353 not complete cCBT ($\text{logit}(\rho)$: $t_{27}=-3.12$; $p=0.004$; $\text{log}(\gamma)$: $t_{27}=2.99$; $p=0.005$) rather
354 than by those who did ($\text{logit}(\rho)$: $t_{23}=-0.72$; $p=0.47$; $\text{log}(\gamma)$: $t_{23}=0.65$; $p=0.51$).

355

356 In sum, responders were on average more prone (that is, lower mean estimate of ρ) to
357 smoothing over previous unsigned prediction errors than non-responders. This implies
358 that responders took greater account of previous feedback history than non-
359 responders. Additionally, on average responders had a tendency to make relatively
360 smaller trial-wise adjustments of the learning rate (that is, greater mean estimate of γ)
361 as a result of outcome surprisingness than non-responders.

362

363 *Neural signature of dynamic learning rate*

364 Whole-brain model-based fMRI analysis of pre-treatment fMRI data revealed
365 significant bilateral activations correlating with the dynamic learning rate in a set of
366 brain regions broadly located in the dorsomedial prefrontal cortex (dmPFC),
367 dorsolateral prefrontal cortex (including precentral gyrus, postcentral gyrus) and
368 occipital cortex (fusiform gyrus, lingual gyrus) (Supplemental Table S1). Significant
369 activations survived family-wise error (FWE) correction for multiple comparisons (p
370 < 0.05) at the cluster level with a cluster-defining threshold of $p < 0.001$.

371

372 A number of previous studies have suggested that the dmPFC in particular might be
373 involved in the dynamic online adjusting of the learning rate (22, 41, 42), although
374 they employed different samples, experimental paradigms and modelling
375 assumptions. Moreover, previous work reported abnormal activation of this region in
376 response to probabilistic feedback in a sample of unmedicated depressed subjects
377 compared to healthy controls (28). We therefore used BOLD activity in the bilateral
378 dmPFC shown in Figure 5 to discriminate treatment response (see Materials and
379 Methods for further details).

380

381 *Pre-treatment BOLD activity encoding dynamic learning rate discriminates*
382 *between responders and non-responders*

383 A two-sample unpaired t-test on the dynamic learning rate contrast images did not
384 identify any significant between group differences. However, on between-group
385 comparisons of the outcome-locked BOLD activity in the dmPFC we found that
386 responders exhibited significantly greater BOLD percent signal change during
387 incongruent negative trials ($F_{1,34.07}=8.87$, $p=0.005$) and significantly lower BOLD
388 percent signal change during congruent negative trials ($F_{1,33.01}=5.15$, $p=0.029$)
389 compared to non-responders (Figure 6). No significant between-group differences
390 were found following positive feedback (congruent: $F_{1,34.72}=1.45$, $p=0.23$;
391 incongruent: $F_{1,33.79}=0.33$, $p=0.56$) (Figure 6).

392

393 Without non-completers between-group differences remained significant only during
394 incongruent negative trials ($F_{1,23.18}=4.99$, $p=0.03$). No significant difference was
395 detected regarding congruent negative feedback ($F_{1,23.02}=1.23$, $p=0.27$) (Supplemental
396 Figure S1). Moreover, we found a statistically significant negative correlation
397 between treatment response magnitude and BOLD activity in incongruent negative
398 trials ($t_{24}=-3.87$, $p<0.001$) but not in congruent negative trials ($t_{24}=0.36$, $p=0.72$)
399 (Figure 7).

400

401

402 **DISCUSSION**

403 In the present study, to discriminate response to cCBT in depression we examined
404 pre-treatment BOLD activity correlating with dynamic appraisal of probabilistic
405 feedback during reinforcement learning. We found that greater BOLD activity in the
406 dmPFC during incongruent negative trials and lower BOLD activity during congruent
407 negative trials were associated with response to cCBT. Furthermore, we provided
408 preliminary evidence that this pattern of BOLD activity may be moderating, rather
409 than mediating, the effect of cCBT on clinical outcome (see Supplemental Results).

410

411 Crucially, although the focus of CBT in depression is centred in fostering adaptive
412 reappraisal strategies (43), to the best of our knowledge no previous study has
413 examined processing of probabilistic feedback as a function of response to CBT.
414 Moreover, in spite of converging evidence that behavioural and neural responses to
415 probabilistic feedback are abnormal in depression, the cognitive mechanisms
416 underlying such impairment are still unknown and relatively unexplored. In this study
417 we have addressed this knowledge gap by explicitly modelling the
418 neurocomputational mechanisms of inference implicated in probabilistic
419 reinforcement learning as a function of treatment response.

420

421 At a computational level we have capitalised on the notion that the model's optimised
422 fixed parameters capture a person's typical mode of appraising incoming probabilistic
423 information. We found that responders took greater account of previous feedback
424 history than non-responders by means of greater smoothing over previous unsigned
425 prediction errors. This operation helps averaging out noisy feedback from the trial-by-
426 trial online computation of surprise and therefore makes inference more robust to

427 random noisy fluctuations in the statistics of the environment. Somewhat
428 complementary to this finding we found that responders were inclined to make
429 smaller adjustments to the learning rate as a result of surprising outcomes. Again, this
430 feature confers a relatively greater degree of robustness against noise to the inference
431 process as it makes the temporal trajectory of the learning rate more stable.

432

433 On the whole, this specific information processing style fosters comparatively more
434 farsighted and judicious decision-making in the responders group. CBT helps patients
435 reframe extreme and dysfunctional negative thoughts in a more considered and
436 balanced fashion (44). One fundamental requirement for CBT to be effective is the
437 patient's ability to actively engage in such work of cognitive restructuring. Depressed
438 patients who are more adept at thoughtfully sieving through the barrage of noisy
439 feedback information surrounding them exhibit a greater predisposition to critical
440 thinking. This may translate in a greater ability to challenge maladaptive thinking
441 patterns. Consistent with this line of reasoning is the prior finding that pre-treatment
442 clinical ratings indicative of lower dysfunctional attitudes (as a result of less rigid and
443 extreme thinking) predict better response to cCBT (12).

444

445 At the behavioural and neural level we have probed between-group differences in the
446 online inference process underlying probabilistic learning, here quantified by a
447 dynamic learning rate. Behaviourally model's derived estimates of the dynamic
448 learning rate did not discriminate response to cCBT. In contrast, BOLD activity did.
449 Significant imaging but not behavioural results are not uncommon in cognitive
450 neuroscience (45). One possible explanation is that underlying between-group
451 differences are better captured by measures of neural activity. Additionally, lack of

452 behavioural corroboration in our study may be due to the specific modelling approach
453 employed here or simply to inadequate power. Crucially, our finding that neural, and
454 not behavioural, measures of information processing discriminate response to cCBT
455 underscores the translational potential of neuroimaging biomarkers in treatment
456 prediction research in psychiatry.

457

458 We have shown that BOLD activity in the dmPFC underpins online appraisal of
459 probabilistic feedback. Our cluster in the dmPFC largely overlaps with the anterior
460 mid-cingulate cortex (aMCC). In a recent review of the functional roles ascribed to
461 the aMCC Shackman et al. proposed that the core function of this region is to
462 determine optimal choice behaviour under conditions of uncertainty (46). During
463 probabilistic learning the differential evaluation of noisy and stochastic feedback is
464 the main mechanism by which an agent tackles environmental uncertainty in the
465 service of adaptive instrumental behaviour. In the computational framework
466 employed here the dynamic learning rate encodes this flexible trial-wise weighting of
467 information. Thus, our finding fits in with current theoretical accounts regarding the
468 involvement of the aMCC in uncertainty resolution (46, 47).

469

470 In particular, we have shown that BOLD activity in the dmPFC denotes a tendency to
471 assign comparatively smaller weight to negative, although still informative, feedback
472 amongst responders. This may be symptomatic of a comparatively greater ability to
473 maintain a positive outlook in spite of adverse events as the expected value of adverse
474 events is more likely to be underestimated (48). Ultimately, a relatively greater
475 capacity to lessen the detrimental impact that day-to-day setbacks can have on one's

476 own view of himself and the surrounding world may facilitate reframing of unhelpful
477 thoughts and help boost response to CBT.

478

479 In contrast, during incongruent negative trials, BOLD activity in the dmPFC speaks to
480 a relatively greater propensity to encode negative but less accurate feedback in the
481 responders group and is thus expected to bias behavioural responses away from
482 advantageous choices. Notably, this finding was robust to response bias and was
483 further corroborated by a significant negative correlation between BOLD activity and
484 post-treatment symptomatic improvement. Interestingly, our results concord with a
485 previous report by Taylor Tavares et al. that lower BOLD activity in the dmPFC was
486 associated with failure to disregard incongruent negative feedback (28).

487

488 At first glance, it would seem counterintuitive for such a pattern of pre-treatment
489 BOLD activity to be associated with clinical improvement. Relatively greater
490 underestimation of the value of generally advantageous choices may cause lack of
491 engagement with positive states and ultimately engender an unduly negative view of
492 the self and the world. One obvious interpretation is that this pattern of BOLD activity
493 is the therapeutic target of CBT. However, this hypothesis is called into question by
494 the observation that cross-session BOLD activity does not change within the
495 responders group (see Supplemental Results). An alternative explanation is that a
496 comparatively greater underestimation of advantageous choices serves the purpose of
497 dampening down expectations on the amount of positive feedback that can be derived
498 from those choices. This may render responders relatively more immune to the
499 disappointment of high expectations not being met and thus buffer the deleterious

500 consequences that negative, although somewhat misleading, events may have on
501 mood.

502

503 In conclusion, in this study we have provided evidence supporting utility and
504 feasibility of a neurocomputational approach to treatment response prediction in
505 depression and, more in general, in mental health research. In particular, we have
506 shown that computational and neural correlates of probabilistic reinforcement
507 learning enable early discrimination of treatment response to cCBT in depression.

508

509

510 **ACKNOWLEDGEMENTS**

511 The authors are grateful to Dr Rajeev Krishnadas for helpful discussion on the study,
512 Dr John McLean for expert technical support and Prof Chris Williams for his help
513 with the CBT-based intervention (<http://lltff.com>).

514 This work was supported by a Chief Scientist Office grant (PN09CP214) and the Dr
515 Mortimer and Theresa Sackler Foundation. F.Q. is supported by an MRC- and MRF-
516 funded clinical research training fellowship (PsySTAR). E.F. and M.G.P were
517 supported by the Biotechnology and Biological Sciences Research Council (BBSRC;
518 grant BB/J015393/2 to M.G.P.) and the Economic and Social Research Council
519 (ESRC; grant ES/L012995/1 to M.G.P.).

520

521 **FINANCIAL DISCLOSURE**

522 The authors have neither conflicts of interest nor financial interests to declare.

523

524 REFERENCES

- 525 1. Beck AT (1972): *Depression; causes and treatment*. Philadelphia,: University
526 of Pennsylvania Press.
- 527 2. Eshel N, Roiser JP (2010): Reward and punishment processing in depression.
528 *Biological psychiatry*. 68:118-124.
- 529 3. Beck AT (1976): *Cognitive therapy and the emotional disorders*. New York:
530 International Universities Press.
- 531 4. Cuijpers P, Berking M, Andersson G, Quigley L, Kleiboer A, Dobson KS
532 (2013): A meta-analysis of cognitive-behavioural therapy for adult depression, alone
533 and in comparison with other treatments. *Can J Psychiatry*. 58:376-385.
- 534 5. Cuijpers P, Donker T, Johansson R, Mohr DC, van Straten A, Andersson G
535 (2011): Self-guided psychological treatment for depressive symptoms: a meta-
536 analysis. *PloS one*. 6:e21274.
- 537 6. Siegle GJ, Thompson WK, Collier A, Berman SR, Feldmiller J, Thase ME, et
538 al. (2012): Toward clinically useful neuroimaging in depression treatment: prognostic
539 utility of subgenual cingulate activity for determining depression outcome in
540 cognitive therapy across studies, scanners, and patient characteristics. *Arch Gen*
541 *Psychiatry*. 69:913-924.
- 542 7. Siegle GJ, Carter CS, Thase ME (2006): Use of fMRI to predict recovery
543 from unipolar depression with cognitive behavior therapy. *The American journal of*
544 *psychiatry*. 163:735-738.
- 545 8. McGrath CL, Kelley ME, Holtzheimer PE, Dunlop BW, Craighead WE,
546 Franco AR, et al. (2013): Toward a neuroimaging treatment selection biomarker for
547 major depressive disorder. *JAMA Psychiatry*. 70:821-829.

- 548 9. Button KS, Wiles NJ, Lewis G, Peters TJ, Kessler D (2012): Factors
549 associated with differential response to online cognitive behavioural therapy. *Soc*
550 *Psychiatry Psychiatr Epidemiol.* 47:827-833.
- 551 10. Spek V, Nyklicek I, Cuijpers P, Pop V (2008): Predictors of outcome of group
552 and internet-based cognitive behavior therapy. *J Affect Disord.* 105:137-145.
- 553 11. de Graaf LE, Hollon SD, Huibers MJ (2010): Predicting outcome in
554 computerized cognitive behavioral therapy for depression in primary care: A
555 randomized trial. *J Consult Clin Psychol.* 78:184-189.
- 556 12. Donker T, Batterham PJ, Warmerdam L, Bennett K, Bennett A, Cuijpers P, et
557 al. (2013): Predictors and moderators of response to internet-delivered Interpersonal
558 Psychotherapy and Cognitive Behavior Therapy for depression. *J Affect Disord.*
559 151:343-351.
- 560 13. Thase ME (2014): Using biomarkers to predict treatment response in major
561 depressive disorder: evidence from past and present studies. *Dialogues in clinical*
562 *neuroscience.* 16:539-544.
- 563 14. Kapur S, Phillips AG, Insel TR (2012): Why has it taken so long for biological
564 psychiatry to develop clinical tests and what to do about it? *Molecular psychiatry.*
565 17:1174-1179.
- 566 15. DeRubeis RJ, Siegle GJ, Hollon SD (2008): Cognitive therapy versus
567 medication for depression: treatment outcomes and neural mechanisms. *Nat Rev*
568 *Neurosci.* 9:788-796.
- 569 16. Ritchey M, Dolcos F, Eddington KM, Strauman TJ, Cabeza R (2011): Neural
570 correlates of emotional processing in depression: changes with cognitive behavioral
571 therapy and predictors of treatment response. *J Psychiatr Res.* 45:577-587.

- 572 17. Siegle GJ, Thompson W, Carter CS, Steinhauer SR, Thase ME (2007):
573 Increased amygdala and decreased dorsolateral prefrontal BOLD responses in
574 unipolar depression: related and independent features. *Biological psychiatry*. 61:198-
575 209.
- 576 18. Goldapple K, Segal Z, Garson C, Lau M, Bieling P, Kennedy S, et al. (2004):
577 Modulation of cortical-limbic pathways in major depression: treatment-specific
578 effects of cognitive behavior therapy. *Arch Gen Psychiatry*. 61:34-41.
- 579 19. McGrath CL, Kelley ME, Dunlop BW, Holtzheimer PE, 3rd, Craighead WE,
580 Mayberg HS (2014): Pretreatment brain states identify likely nonresponse to standard
581 treatments for depression. *Biological psychiatry*. 76:527-535.
- 582 20. Maia TV (2009): Reinforcement learning, conditioning, and the brain:
583 Successes and challenges. *Cogn Affect Behav Neurosci*. 9:343-364.
- 584 21. Payzan-LeNestour E, Bossaerts P (2011): Risk, unexpected uncertainty, and
585 estimation uncertainty: Bayesian learning in unstable settings. *PLoS Comput Biol*.
586 7:e1001048.
- 587 22. Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007): Learning the
588 value of information in an uncertain world. *Nat Neurosci*. 10:1214-1221.
- 589 23. Nassar MR, Wilson RC, Heasly B, Gold JJ (2010): An approximately
590 Bayesian delta-rule model explains the dynamics of belief updating in a changing
591 environment. *The Journal of neuroscience : the official journal of the Society for*
592 *Neuroscience*. 30:12366-12378.
- 593 24. Krugel LK, Biele G, Mohr PN, Li SC, Heekeren HR (2009): Genetic variation
594 in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt
595 decisions. *Proceedings of the National Academy of Sciences of the United States of*
596 *America*. 106:17951-17956.

- 597 25. Mathys C, Daunizeau J, Friston KJ, Stephan KE (2011): A bayesian
598 foundation for individual learning under uncertainty. *Front Hum Neurosci.* 5:39.
- 599 26. Browning M, Behrens TE, Jocham G, O'Reilly JX, Bishop SJ (2015): Anxious
600 individuals have difficulty learning the causal statistics of aversive environments. *Nat*
601 *Neurosci.* 18:590-596.
- 602 27. Murphy FC, Michael A, Robbins TW, Sahakian BJ (2003):
603 Neuropsychological impairment in patients with major depressive disorder: the effects
604 of feedback on task performance. *Psychol Med.* 33:455-467.
- 605 28. Taylor Tavares JV, Clark L, Furey ML, Williams GB, Sahakian BJ, Drevets
606 WC (2008): Neural basis of abnormal response to negative feedback in unmedicated
607 mood disorders. *NeuroImage.* 42:1118-1126.
- 608 29. Lewis G, Pelosi AJ, Araya R, Dunn G (1992): Measuring psychiatric disorder
609 in the community: a standardized assessment for use by lay interviewers. *Psychol*
610 *Med.* 22:465-486.
- 611 30. Beck AT, Steer RA, Ball R, Ranieri W (1996): Comparison of Beck
612 Depression Inventories -IA and -II in psychiatric outpatients. *J Pers Assess.* 67:588-
613 597.
- 614 31. Pearce JM, Hall G (1980): A model for Pavlovian learning: variations in the
615 effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev.* 87:532-
616 552.
- 617 32. Huys QJ, Eshel N, O'Nions E, Sheridan L, Dayan P, Roiser JP (2012): Bonsai
618 trees in your head: how the pavlovian system sculpts goal-directed choices by pruning
619 decision trees. *PLoS Comput Biol.* 8:e1002410.

- 620 33. Huys QJ, Cools R, Golzer M, Friedel E, Heinz A, Dolan RJ, et al. (2011):
621 Disentangling the roles of approach, activation and valence in instrumental and
622 pavlovian responding. *PLoS Comput Biol.* 7:e1002028.
- 623 34. Deichmann R, Gottfried JA, Hutton C, Turner R (2003): Optimized EPI for
624 fMRI studies of the orbitofrontal cortex. *NeuroImage.* 19:430-441.
- 625 35. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE,
626 Johansen-Berg H, et al. (2004): Advances in functional and structural MR image
627 analysis and implementation as FSL. *NeuroImage.* 23 Suppl 1:S208-219.
- 628 36. Jenkinson M, Bannister P, Brady M, Smith S (2002): Improved optimization
629 for the robust and accurate linear registration and motion correction of brain images.
630 *NeuroImage.* 17:825-841.
- 631 37. Jenkinson M, Smith S (2001): A global optimisation method for robust affine
632 registration of brain images. *Medical image analysis.* 5:143-156.
- 633 38. Greve DN, Fischl B (2009): Accurate and robust brain image alignment using
634 boundary-based registration. *NeuroImage.* 48:63-72.
- 635 39. Beckmann CF, Jenkinson M, Smith SM (2003): General multilevel linear
636 modeling for group analysis in FMRI. *NeuroImage.* 20:1052-1063.
- 637 40. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006): Cortical
638 substrates for exploratory decisions in humans. *Nature.* 441:876-879.
- 639 41. Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HE,
640 et al. (2013): Hierarchical prediction errors in midbrain and basal forebrain during
641 sensory learning. *Neuron.* 80:519-530.
- 642 42. McGuire JT, Nassar MR, Gold JI, Kable JW (2014): Functionally dissociable
643 influences on learning rate in a dynamic environment. *Neuron.* 84:870-881.
- 644 43. Beck AT (1979): *Cognitive therapy of depression.* New York: Guilford Press.

- 645 44. Wright JH, Beck AT (1983): Cognitive Therapy of Depression - Theory and
646 Practice. *Hosp Community Psych.* 34:1119-1127.
- 647 45. Wilkinson D, Halligan P (2004): The relevance of behavioural measures for
648 functional-imaging studies of cognition. *Nat Rev Neurosci.* 5:67-73.
- 649 46. Shackman AJ, Salomons TV, Slagter HA, Fox AS, Winter JJ, Davidson RJ
650 (2011): The integration of negative affect, pain and cognitive control in the cingulate
651 cortex. *Nat Rev Neurosci.* 12:154-167.
- 652 47. Rushworth MF, Behrens TE (2008): Choice, uncertainty and value in
653 prefrontal and cingulate cortex. *Nat Neurosci.* 11:389-397.
- 654 48. Sharot T (2011): The optimism bias. *Time.* 177:40-44, 46.
- 655 49. Wager TD, Nichols TE (2003): Optimization of experimental design in fMRI:
656 a general framework using a genetic algorithm. *NeuroImage.* 18:293-309.
- 657
- 658

659 **FIGURE LEGENDS**

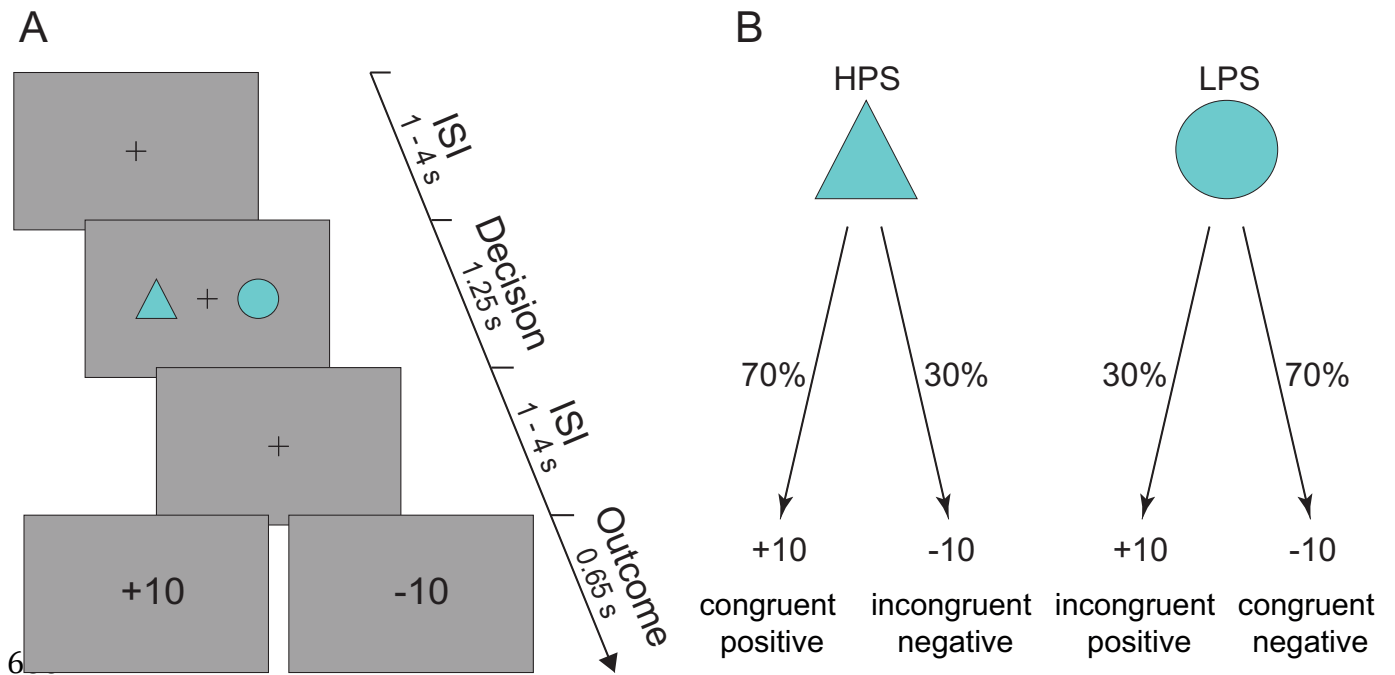
660 **Figure 1. Probabilistic reversal-learning task.**

661 (A) Each trial commenced with a jittered interstimulus interval (1-4s) displaying a
662 fixation cross. Subsequent to this two abstract visual stimuli appeared randomly
663 on either side of the screen for 1.25s. For each participant the two stimuli were
664 randomly chosen from a pool of 18 different geometrical shapes. Participants
665 were given 1s to choose a stimulus via a button press. Following a second jittered
666 interstimulus interval (1-4s) participants were presented with the outcome of their
667 decision for 0.65s. Outcome was either positive (+10) or negative (-10). To
668 maximize design efficiency the duration of jittered interstimulus intervals was
669 optimized implementing a genetic algorithm described in (49). ISI: Interstimulus
670 Interval.

671 (B) Classification of probabilistic feedback as a function of congruence and valence.
672 Stimulus-outcome contingencies were asymmetrically skewed (70-30%) so that
673 the expected value of the two stimuli was of the same magnitude but of opposite
674 sign. This meant that whilst one stimulus (here referred to as the high probability
675 stimulus) was associated with a greater likelihood of positive outcome, the other
676 stimulus (here referred to as the low probability stimulus) was associated with a
677 greater likelihood of negative outcome. Reversals were self-paced and occurred
678 when participants chose the high probability stimulus five times over the last six
679 trials. To prevent participants from figuring out the underlying reversal rule we
680 ran a randomly generated number of buffer trials from a zero-truncated Poisson
681 distribution before reversing stimulus-outcome contingencies. The stimulus-
682 outcome association strength was chosen to enable detection of reversals.
683 Participants were only advised of the probabilistic nature of the task and that

684 stimulus-outcome contingencies might reverse based upon their performance.

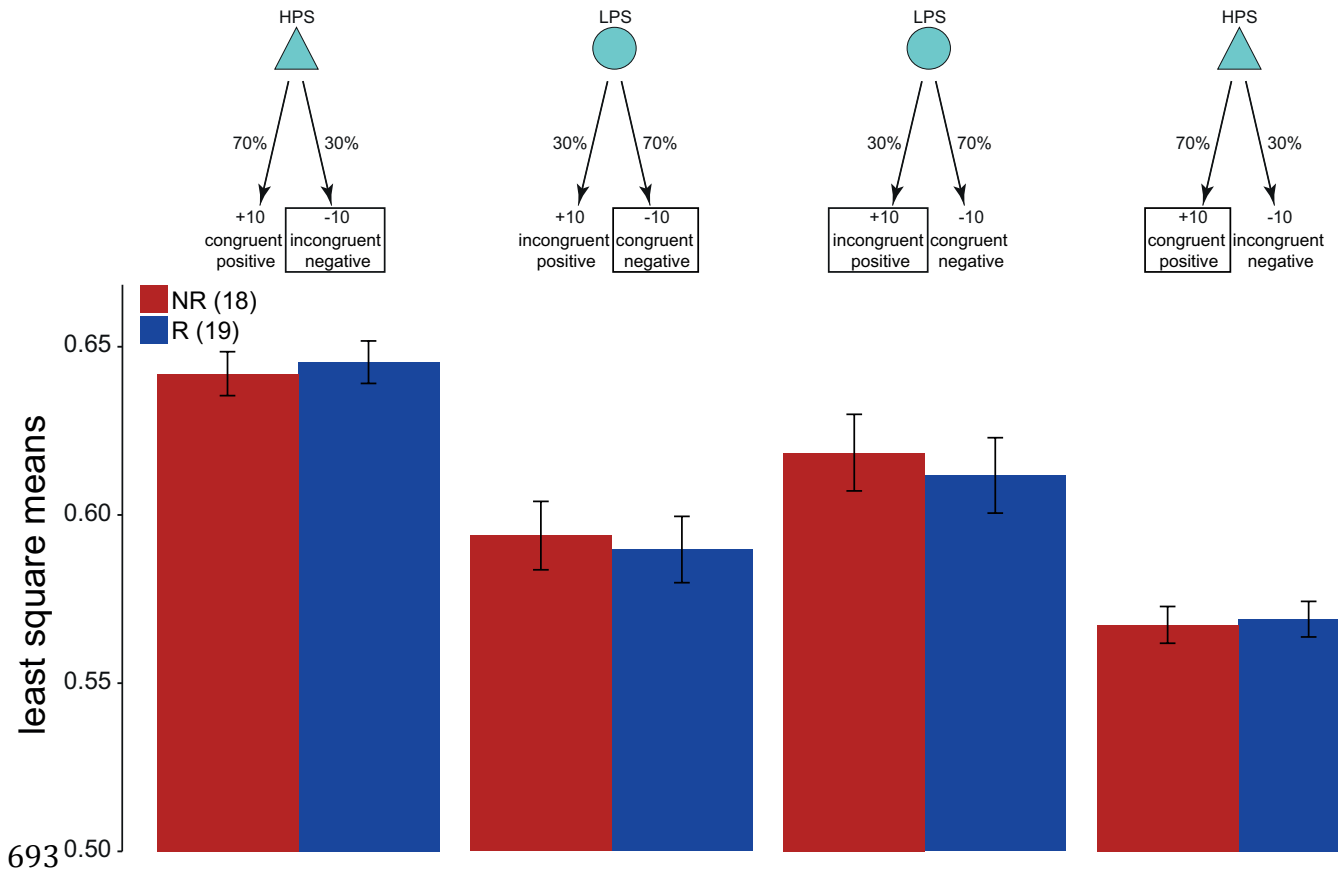
685 HPS: High Probability Stimulus. LPS: Low Probability Stimulus.



687

688 **Figure 2. Analysis of behavioural estimates of dynamic learning rate.**

689 Least-square mean estimates \pm SEM (error bars) of dynamic learning rate in the
690 responders (blue; n=19) and non-responders (red; n=18) groups are plotted as a
691 function of feedback congruence and valence. None of the post-hoc between-group
692 comparisons were statistically significant.

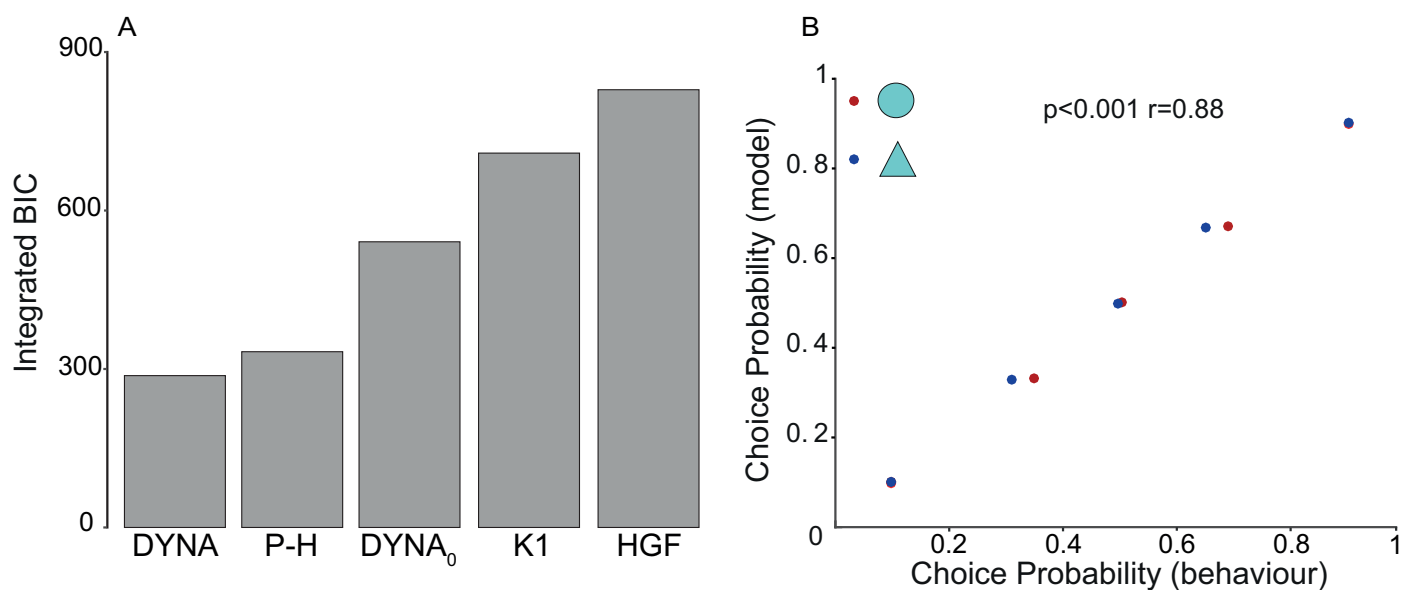


694

695 **Figure 3. Computational model comparison and behavioural fit.**

696 (A) BIC_{Int} scores for all models. Lower scores indicate better fit. DYNA is the
697 winning model. DYNA: Krugel et al.'s model. DYNA₀: Krugel et al.'s model
698 with additional parameter α^1 . HGF: Hierarchical Gaussian Filter. P-H: Pearce-
699 Hall. K1: Kalman filter K1 variant.

700 (B) Scatterplot shows linear relationship between empirical and predicted choice
701 probabilities. r : Pearson's correlation coefficient.

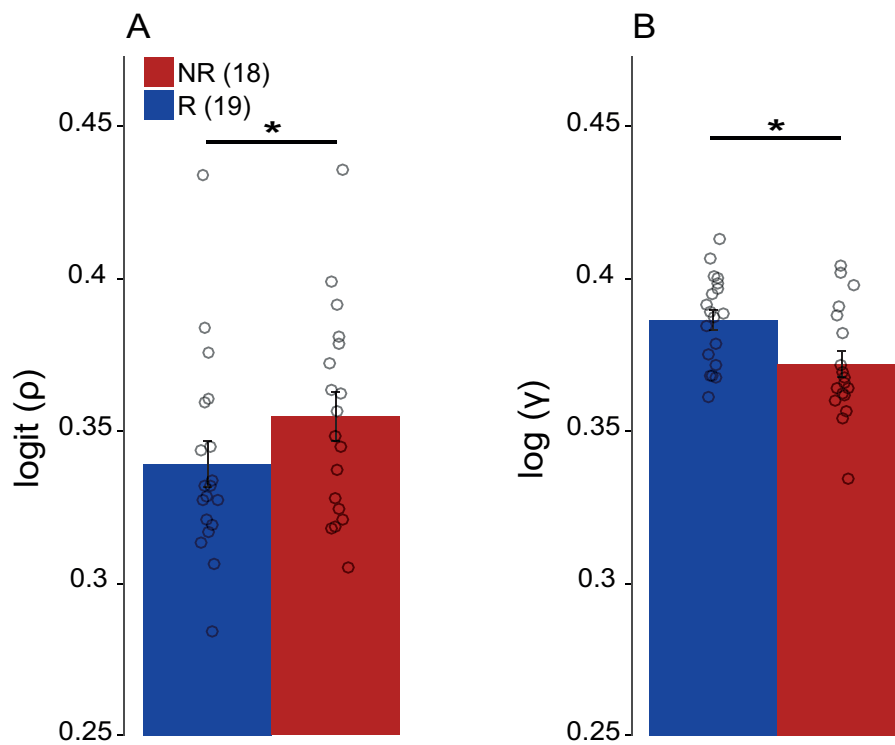


702

703

704 **Figure 4. Computational parameters.**

705 Mean estimates \pm SEM (error bars) of ρ (left) and γ (right) in the responders (blue)
706 and non-responders (red) groups. Parameter estimates are shown in their native space
707 (logit for ρ and log for γ). After adjusting for pre-treatment BDI-II score, on average
708 responders exhibited a greater tendency to smooth over previous unsigned prediction
709 errors (lower ρ mean estimate) and to make smaller adjustments to the dynamic
710 learning rate (greater γ mean estimate) compared to non-responders. Black circles
711 represent individual subjects. * $p < 0.05$.

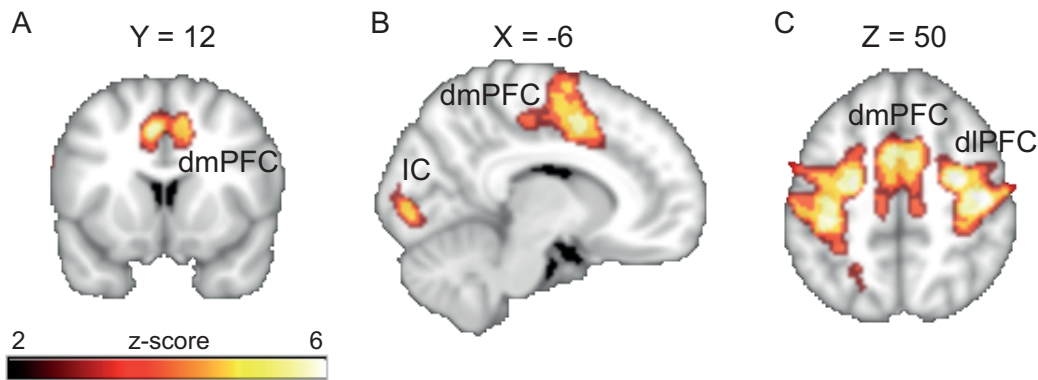


712

713

714 **Figure 5. Pre-treatment whole-brain results for dynamic learning rate**
715 **contrast.**

716 Activations shown here survived a cluster-defining threshold of $Z > 3.1$ and FWE-
717 corrected significance threshold of $p = 0.05$. (A) dmPFC: dorsomedial prefrontal
718 cortex. (B) IC: intracalcarine cortex (C) dlPFC: dorsolateral prefrontal cortex.
719 Coordinates are given in the MNI space. We used the cluster in the dmPFC to extract
720 trial-wise feedback-locked BOLD percent signal change for all participants and used
721 them to discriminate treatment response.

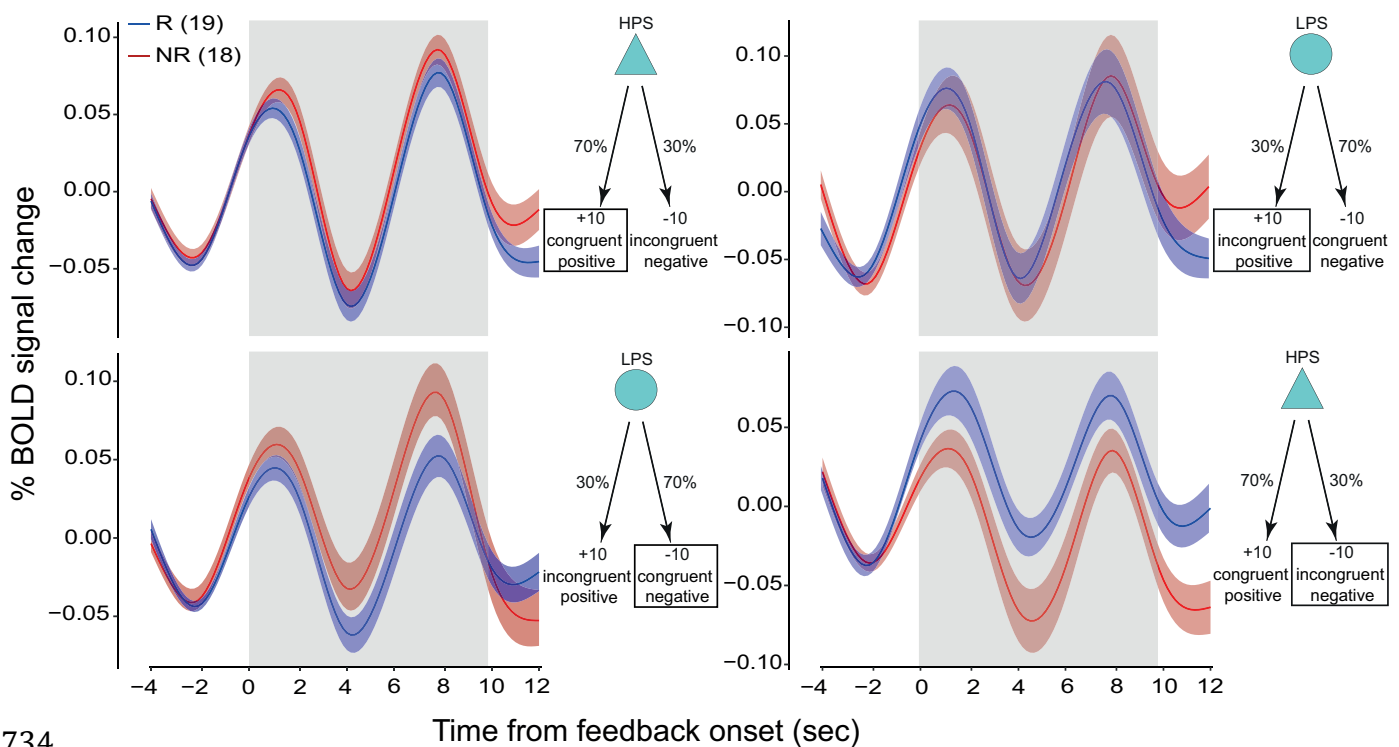


722

723

724 **Figure 6. Pre-treatment time course of BOLD percent signal change in the**
725 **dmPFC.**

726 Average feedback-locked BOLD percent signal change \pm SEM (blue and red
727 shading) in responders (blue; n=19) and non-responders (red; n=18) groups is plotted
728 as a function of feedback valence and congruence. Averages are adjusted for pre-
729 treatment BDI-II score. We performed between-group comparisons over the temporal
730 window shown by the grey shaded box. Responders exhibited comparatively lower
731 BOLD activity during congruent negative trials (bottom left; $F(1,33.01)=5.15$,
732 $p=0.029$) but comparatively higher BOLD activity during incongruent negative trials
733 (bottom right; $F(1,34.07)=8.87$, $p=0.005$).

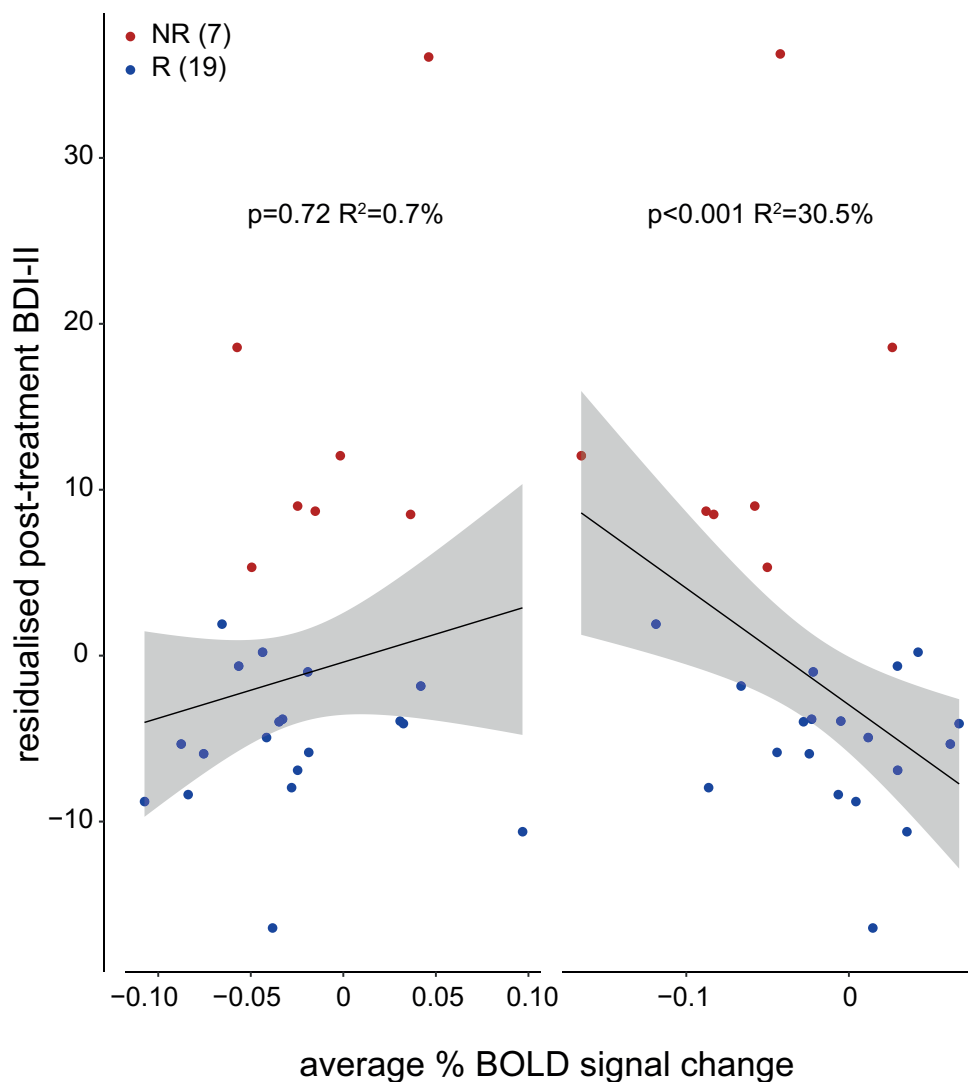


734

735

736 **Figure 7. Relationship between BOLD percent signal change in the**
737 **dmPFC and treatment response.**

738 Scatterplots show robust linear fit \pm 95% CI (grey shading) of adjusted post-treatment
739 BDI-II score against pre-treatment BOLD percent signal change in the dmPFC. Left:
740 Congruent negative trials. Right: Incongruent negative trials. Post-treatment BDI-II
741 score is adjusted for pre-treatment BDI-II score. Subject-wise pre-treatment BOLD
742 percent signal change is averaged over a time window extending from 4 to 6 seconds
743 after the onset of the outcome phase. Blue dots indicate responders (n=19) and red
744 dots indicate non-responders (n=7).



745

746

747 **Supplemental Figure S1. Pre-treatment time course of BOLD percent**
748 **signal change in the dmPFC without non-completers.**

749 Average feedback-locked BOLD percent signal change \pm SEM (blue and red
750 shading) in responders (blue; n=19) and completers non-responders (red; n=7) groups
751 is plotted as a function of feedback valence and congruence. Averages are adjusted
752 for pre-treatment BDI-II score. We performed between-group comparisons over the
753 temporal window shown by the grey shaded box. Between-group differences were
754 robust to response bias during incongruent negative trials (bottom right;
755 $F(1,23.2)=4.99$, $p=0.03$) but not during congruent negative trials (bottom left;
756 $F(1,22.6)=1.23$, $p=0.27$).

