

Neural Arbitration between Social and Individual Learning Systems

Andreea O. Diaconescu^{1,2,3,4¶}, Madeline Stecy^{1,2,5¶}, Lars Kasper^{1,2,6}, Christopher J. Burke², Zoltan Nagy², Christoph Mathys^{1,7,8}, Philippe N. Tobler²

¹Translational Neuromodeling Unit, Institute for Biomedical Engineering, University of Zurich & ETH Zurich, Switzerland;

²Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Switzerland;

³University of Basel, Department of Psychiatry (UPK), Basel, Switzerland;

⁴Krembil Centre for Neuroinformatics, Centre for Addiction and Mental Health (CAMH), University of Toronto, Canada

⁵Rutgers Robert Wood Johnson Medical School, New Jersey, United States;

⁶Institute for Biomedical Engineering, MRI Technology Group, ETH Zürich & University of Zurich, Switzerland;

⁷Interacting Minds Centre, Aarhus University, Aarhus, Denmark;

⁸Scuola Internazionale Superiore di Studi Avanzati (SISSA), Trieste, Italy

¶The authors contributed equally to this work

Correspondence should be addressed to:
Andreea O. Diaconescu, PhD
Krembil Centre for Neuroinformatics (CAMH)
250 College St. M5T 1R8
Primary email: andreea.diaconescu@camh.ca
Secondary email: andreea.diaconescu@utoronto.ca

Keywords: hierarchical Bayesian inference, observational learning, reinforcement learning, precision, uncertainty, fMRI, dopamine

34 **Acknowledgments:**

35 We are grateful for support by the Swiss National Science Foundation (Ambizione
36 grant PZooP3_167952 to AOD; PPooP1_150739, 100014_165884, and 100019_176016 to
37 PNT) and the Krembil Foundation to AOD. We are also grateful to Klaas Enno
38 Stephan for providing guidance and funding for the study.

39

40 **Disclosure statement:**

41 The authors have no disclosures or conflict of interest.

42

43 **Code Availability:**

44 The routines for all analyses are available as Matlab code:
45 <https://github.com/andreadiaconescu/arbitration>. The instructions for running the
46 code in order to reproduce the results can be found in the ReadMe file.

47

48 **Abstract**

49 Decision making requires integrating self-gathered information with advice from
50 others. However, the arbitration process by which one source of information is
51 selected over the other has not been fully elucidated. In this study, we formalised
52 arbitration as the relative precision of predictions, afforded by each learning system,
53 using hierarchical Bayesian modelling. In a probabilistic learning task, participants
54 predicted the outcome of a lottery using recommendations from a more informed
55 advisor and/or self-sampled outcomes. Decision confidence, as measured by the
56 number of points participants wagered on their predictions, varied with our relative
57 precision definition of arbitration. Functional neuroimaging demonstrated arbitration
58 signals that were independent of decision confidence and involved modality-specific
59 brain regions. Arbitrating in favour of self-gathered information activated the
60 dorsolateral prefrontal cortex and the midbrain, whereas arbitrating in favour of social
61 information engaged the ventromedial prefrontal cortex and the amygdala. These
62 findings indicate that relative precision captures arbitration between social and
63 individual learning systems at both behavioural and neural levels.

64

65

Introduction

66 As social primates navigating an uncertain world, humans use multiple information
67 sources to guide their decisions (Charness et al., 2013). For example, in investment
68 decisions investors may either choose to follow a financial expert's advice about a
69 particular stock or base the decision on their own previous experience with that stock.
70 When information from personal experience and social advice conflict, one source
71 must be favoured over the other to guide decision making. We conceptualize the
72 process of selecting between information sources as arbitration. Arbitration is
73 particularly important in uncertain situations when different source of information
74 have different reliability. While stock performance may fluctuate, the advisor could
75 pursue selfish interests. It is challenging to infer the intentions of the advisor because
76 they are concealed or expressed indirectly, requiring inference from observations of
77 ambiguous behaviour. Optimal arbitration should therefore consider the relative
78 uncertainty associated with each source of information.

79 Arbitration between different types of reward predictions based on experiential
80 learning acquired by an individual has been associated with the prefrontal cortex.
81 Specifically, the dorsolateral prefrontal cortex (DLPFC) and the frontopolar cortex
82 have been shown to arbitrate between habitual (model-free) and planned (model-
83 based) learning systems (Lee et al., 2014). By contrast, comparatively little is known
84 about how humans weigh self-gathered (individual) reward information against
85 observed (social) information. To fill this gap, we considered two hypotheses: First,
86 arbitration involving social information could rely on theory of mind (ToM) processes,
87 i.e., inference about others' mental states (Frith and Frith, 2005; Schaafsma et al., 2015)
88 and higher-level social representations (Frith, 2012; Devaine et al., 2014a). Accordingly,
89 arbitration involving the intentions of others may rely on activity in classical ToM
90 regions, such as the temporoparietal junction (TPJ) and dorsomedial prefrontal cortex
91 (Carrington and Bailey, 2009; Frith and Frith, 2010; Baker et al., 2011; Schurz et al.,
92 2014). Alternatively, arbitration between individual and social information may involve
93 similar neural networks as those orchestrating between model-free and model-based
94 learning (Lee et al., 2014), and thus engage lateral prefrontal and frontopolar regions.

95 It is also worth noting that arbitration depends on both experienced and inferred
96 value learning. Similar to directly experienced reward learning, inferring on others'
97 intentions engages the striatum, potentially signalling the value associated with social
98 feedback during probabilistic reward learning tasks. For example, parts of the striatum
99 including the caudate show stronger activations in response to reciprocated compared
100 to unreciprocated cooperation during iterative trust games (Delgado et al., 2005a;
101 King-Casas et al., 2008; Fareri et al., 2015), and represent social prediction errors
102 signalling a change in fidelity (Delgado et al., 2005b; Biele et al., 2009; Klucharev et al.,
103 2009; Campbell-Meiklejohn et al., 2010; Braams et al., 2014; Diaconescu et al., 2017).

104 Moreover, with respect to tracking higher-level, contextual change about both reward
105 contingencies and intentionality one may expect the involvement of the anterior
106 cingulate cortex (ACC). In addition to being associated with volatility tracking in a
107 probabilistic reward learning task (Behrens et al., 2007), the ACC was shown to
108 represent volatility precision-weighted PEs during social learning (Diaconescu et al.,
109 2017).

110 An additional intriguing question is which neuromodulatory system supports the
111 arbitration process. Since arbitration is dependent on the uncertainty of predictions
112 afforded by each learning system, several neuromodulatory systems are good
113 candidates. For non-social forms of learning, previous studies have implicated
114 dopaminergic, cholinergic, and noradrenergic systems in signalling uncertainty, the
115 inverse of precision (Yu and Dayan, 2005; Iglesias et al., 2013; Payzan-LeNestour et al.,
116 2013; Schwartenbeck et al., 2014; Marshall et al., 2016). Here, we examined the unique
117 contribution of arbitration to activity across dopaminergic, cholinergic, and
118 noradrenergic neuromodulatory systems.

119 To investigate arbitration between individual and social learning systems, we
120 simulated the aforementioned stock investment scenario in the laboratory. In other
121 words, we examined how people arbitrate between individual reward information and
122 social advice regarding a probabilistic lottery where contingencies changed over time.
123 Participants learned to predict the colour of a binary card draw using uncertain advice
124 from a more informed advisor and uncertain information inferred from individually
125 observed card outcomes (*Figure 1*).

126 We separately manipulated the degree of uncertainty (or its inverse, precision)
127 associated with each of the two information sources by independently varying the rate
128 of change with which each information source predicted the drawn card colour (i.e.,
129 volatility; Behrens et al., 2007). The advisor was motivated to give correct or incorrect
130 advice depending on the phase of the task, resulting in variations in the reliability of
131 social information. Performing well in the task therefore required keeping track of the
132 probabilities of the two sources of information and arbitrating between them. We
133 assumed that participants weigh the predictions afforded by each information source
134 as a function of their precision. Thus, we expected participants to rely more on the
135 advice when the advisor's intentions were perceived as stable, and on their personal
136 experience when the advice was perceived to be volatile.

Arbitrating between social and individual learning systems

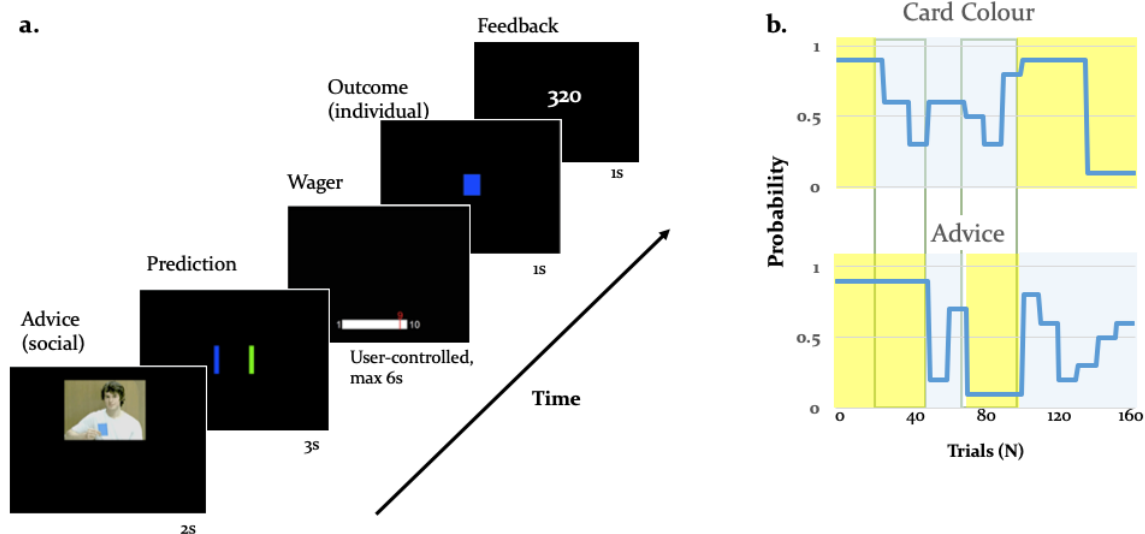


Figure 1

Figure 1| Experimental Paradigm: (a) Binary lottery game requiring arbitration between individual experience and social information. Volunteers predicted the outcome of a binary lottery, i.e., whether a blue or green card would be drawn. They could base the prediction on two sources of information: advice from a gender-matched advisor (video, presented for 2s) who was better informed about the colour of the drawn card, and on an estimate about the statistical likelihood of the cards being one or the other colour that the participant had to infer from own experience (outcome, 1s). After predicting the colour of the rewarded lottery card (user-controlled, maximum 3s), participants also wagered one to ten points (user-controlled, maximum 6s), which they would win or lose depending on whether the prediction was right or wrong. After the outcome, participants viewed their cumulative score on the feedback screen (1s). (b) Contingencies of individual reward and social advice information. Card colour probability corresponds to the likelihood of a given colour (e.g., blue) being rewarded. The probabilities were matched on average for the two information sources (55% for the card colour information and 56% for the advice information). The two sources of information were uncorrelated as illustrated by phases of low (yellow) and high (light grey) volatility, enabling a factorial analysis of information source and volatility.

138 **Results**

139 To examine the neural mechanisms underlying arbitration, we recruited 48 volunteers
140 (mean age 23.6 ± 1.4 , 32 females) to perform a binary lottery task requiring arbitration
141 between individual experienced card outcomes and expert advice. We combined fMRI
142 with a computational modelling approach using the hierarchical Gaussian filter (HGF)
143 (Mathys et al., 2011, 2014). This hierarchical Bayesian model is ideally suited to address
144 our question as it examines multi-level inference and provides trial-wise estimates of
145 estimated precision of predictions about each information source. This framework
146 operationalises arbitration as a precision ratio, corresponding to the relative perceived
147 precision of each information source (*Figure 2*). Thus, arbitration is a function of the
148 relative stability of the advice or the card colour probabilities. In our paradigm, this
149 quantity increased when the precision of the predictions about one of the two sources
150 of information was high and decreased when both sources were either stable or
151 volatile (see *Figure 3* for the arbitration signal averaged across participants).

152 ***Behaviour: Accuracy of lottery outcome prediction and wager amount***

153 Using the factorial structure of the task, we tested the impact of volatility on
154 performance with a two-factor repeated measures ANOVA, where the two factors
155 were information source (card versus advice) and phase (stable versus volatile). Across
156 all behavioural metrics, we observed an effect of phase, indicating a reduction in
157 performance in volatile compared to stable phases, and a phase \times information
158 interaction, indicating that the effect was larger for the social than the individual
159 source of information. First, for the accuracy with which participants predicted lottery
160 outcome, we found a main effect of phase ($df = (1,36)$, $F = 187.94$, $p = 7.7e-16$) and an
161 information source-by-phase interaction ($df = (1,36)$, $F = 11.13$, $p = 0.0020$) (see *Figure 1 -*
162 *figure supplement 1a*). Thus, in-keeping with the rationale that arbitration relates to
163 relative information quality, the degree to which participants relied on each
164 information source was a function of precision as manipulated by the volatility
165 structure of the task. Participants performed significantly better in stable compared to
166 volatile periods of the task. These effects were not modulated by fatigue, as we found
167 no significant differences between early and late phases of the task.

168 Second, also advice-taking behaviour differed as a function of volatility and
169 information source: For the percentage of trials in which participants followed a given
170 source of information, we detected a main effect of phase ($df = (1,36)$, $F = 56.26$, $p =$
171 $7.3073e-09$) and an information source-by-phase interaction ($df = (1,36)$, $F = 25.86$, $p =$
172 $1.1561e-05$) (*Figure 1 - figure supplement 1b*). Thus, participants took advice less often
173 particularly when it was volatile rather than stable.

174 Third, the amount of points wagered also depended on the task volatility and the
175 information source. We observed a main effect of phase ($df = (1,36)$, $F = 28.78$, $p =$

176 4.54e-06) and an information source-by-phase interaction ($df = (1,36)$, $F = 16.75$, $p =$
177 2.21e-04 (Figure 1 - figure supplement 1c). Participants wagered fewer points particularly
178 when advice was volatile. Moreover, the number of points wagered correlated
179 significantly with the total score in stable phases ($r = 0.37$, $p = 0.02$), but not in volatile
180 phases ($r = 0.30$, $p = 0.06$). Simulations using a 2-level HGF (with low and fixed
181 volatility) suggested that tracking volatility is beneficial for task performance: a
182 hypothetical person who did not take the volatility of the task phases into account
183 gained on average 21.6 points less than an agent tracking volatility. In line with
184 previous evidence (Behrens et al., 2008), these results emphasize the impact of
185 volatility on the willingness to invest and investment success as measured here by
186 total score.

187 *Participants Process Social Information and Distinguish it from* 188 *Individual Information*

189 **Advisor Ratings:** Participants were asked to rate the advisor (i.e., helpful, misleading,
190 or neutral with regard to suggesting the correct outcome) in a multiple-choice
191 question presented 5 times during the experiment. The time points were associated
192 with different social and individual information (initial/prior: 1st trial; stable advice,
193 stable card phase = 14th trial); stable advice, volatile card phase (49th trial); volatile
194 advice, volatile card phase (73rd trial); volatile advice, stable card phase = 115th trial).
195 On average, participants rated the advice as 75.0% \pm 4.6% (mean \pm standard deviation)
196 helpful in the stable advice phase. The corresponding values were 50% \pm 3.4% in the
197 volatile advice phase, 63.8% \pm 4.4% in the stable card phase, and 61.2% \pm 3.8% in the
198 volatile card phase.

199
200 We examined the extent to which participants' ratings changed as a function of the
201 task phases, and found a significant main effect of phase ($df = (1,36)$, $F = 15.67$, $p = 3.3e-$
202 04) and a significant information source \times phase interaction ($df = (1,36)$, $F = 8.42$, $p =$
203 0.0062). This suggests that advice ratings decreased during volatile compared to stable
204 phases, and this effect was more strongly related to the advice compared to the card
205 information.

206

207 **Debriefing Questionnaire:** After completing the task, participants filled out a task-
208 specific debriefing questionnaire, assessing their perception of the advisor and how
209 they integrated the social information during the task. The questions were originally
210 presented to participants in their native German, and are translated here into English.

211 First, participants were asked to describe the strategy the advisor used in the game
212 (debriefing question 3: "Did the advisor intentionally use a strategy during the task? If
213 yes, describe what strategy that was"). Thirty out of 38 participants answered "Yes" to
214 this question, and described (in their own words) the advisor's strategy. We repeated

215 our analyses including only these 30 participants and found that all conclusions
216 remained statistically the same. Second, participants were asked to rate the advice on
217 a 6-point Likert scale ranging from unhelpful to very helpful (debriefing question 4:
218 “*How helpful did you perceive the advice you received?*”). In general, participants rated
219 the advisors’ recommendations as helpful (mean ratings 4.2 ± 1.0 , ranging from 2 to 6).
220 Finally, we also asked participants to rate, in terms of percentages, how often they
221 followed the advice (debriefing question 5: “*How often did you follow the*
222 *recommendations of the advisor?*”). On average, participants reported that they
223 followed the advice 60% of the time (mean ratings 60 ± 12), which significantly
224 differed from chance ($t(37) = 5.02$, $p = 1.29e-05$). Thus, participants experienced
225 advisors as intentional and helpful, which are core characteristics of social agents.

226

227 ***Model-based results***

228 We used computational modelling with hierarchical Gaussian Filters (HGF; *Figure 2*)
229 to explain participants’ responses on every trial. To contrast competing mechanisms
230 underlying learning and arbitration, our model space included a total of 9 models
231 (*Figure 3a*). Non-normative perceptual models varied in complexity of volatility
232 processing (3-level full HGF vs. 2-level no-volatility HGF), normative perceptual
233 models assumed optimal Bayesian inference (normative HGF), and response models
234 varied in the extent of arbitration (arbitration; no arbitration: advice only; no
235 arbitration: card information only). Bayesian model selection (Stephan et al., 2009)
236 served to compare models (see Methods and *Figure 2* for details). For model
237 comparison, we used the log model evidence (LME), which represents a trade-off
238 between model complexity and model fit.

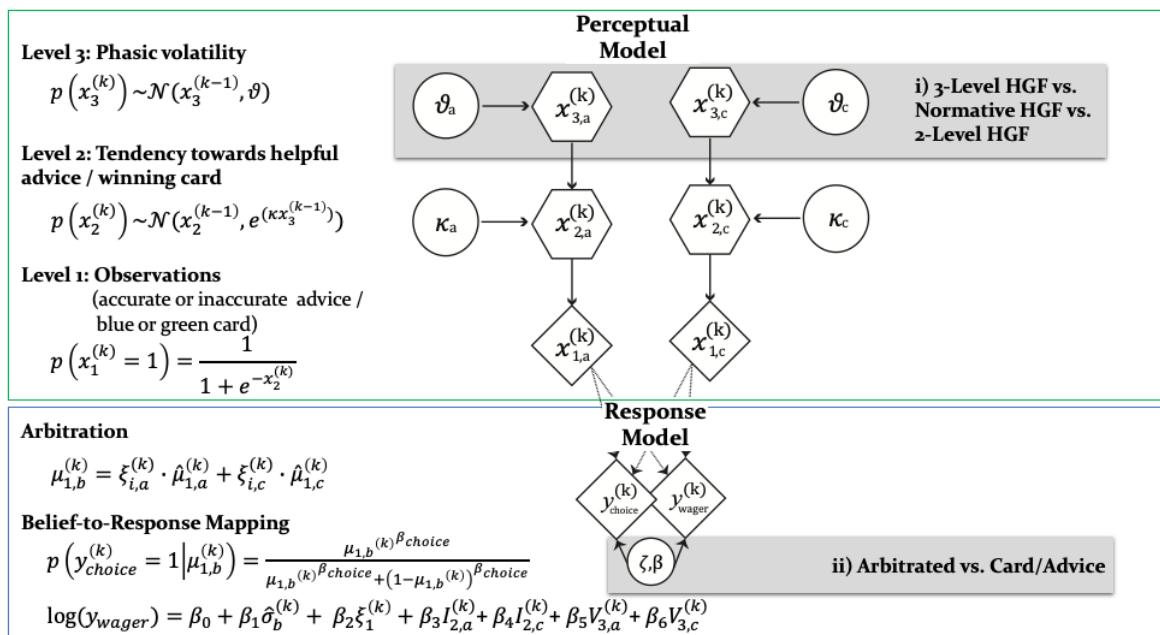


Figure 2

Figure 2| **Computational learning and arbitration model:** In this graphical notation, circles represent constants whereas hexagons and diamonds represent quantities that change in time (i.e., that carry a time/trial index). Hexagons in contrast to diamonds additionally depend on the previous state in time in a Markovian fashion. The two-branch HGF describes the generative model for advice and card probability: x_1 represents the accuracy of the current advice/card colour probability, x_2 the tendency of the advisor to offer helpful advice tendency of card colour to be rewarded, and x_3 the current volatility of the advisor's intentions/card colour probabilities. Learning parameters describe how the states evolve in time. Parameter κ determines how strongly x_2 and x_3 are coupled, and ϑ represents the meta-volatility of x_3 . The response model maps the predicted colour probabilities to choices. The response model also assumes that trial-wise wagers and predictions arise from a linear combination of arbitration, informational uncertainty (advice and card), and volatility (advice and card). For model selection we combined three perception with three response models (see Figure 3). All the models considered can be grouped according to common features and divided into model families: (i) the Perceptual model families distinguish between more (non-normative and normative 3-level) and less (2-level) complex types of HGFs. More specifically, the distinction between 3-level and 2-level HGFs refers to estimating or fixing the volatility of the third level; normative but not non-normative HGFs assume optimal Bayesian inference. (ii) Response model families distinguish between arbitrated and single-information source – advice or card only – models, which correspond to estimating parameter ζ or fixing it to reduce arbitration to either the advice prediction or the card colour prediction.

239

240 *Do participants arbitrate between advice and individually sampled card outcomes?*

241 The winning model was the 3-level HGF with arbitration ($\phi_p = 0.999$; Bayes Omnibus
 242 Risk = 4.26e-11; Figure 3b; Table 2a). This model captured arbitration with the ratio of
 243 precisions: the precision of the prediction about advice accuracy and colour
 244 probability, divided by total precision. Moreover, the model included a social bias
 245 parameter reflecting the degree to which participants followed the advisor irrespective

Arbitrating between social and individual learning systems

246 of task information. The model family that included volatility of both information
 247 sources outperformed models without volatility, in-keeping with the model-
 248 independent finding that perceived volatility of both information sources affected
 249 behaviour.

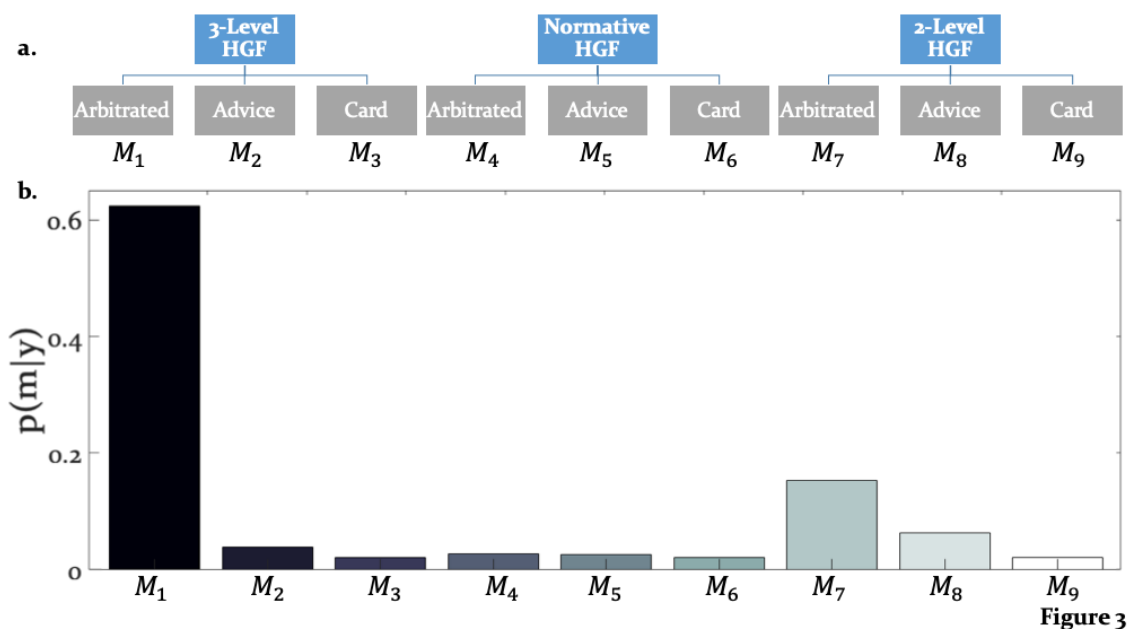


Figure 3| **Hierarchical structure of the model space and model selection results:** (a) The learning and arbitration models considered in this study have a 3 x 3 factorial structure and can be displayed as a tree. The nodes at the top level represent the perceptual model families (3-level HGF, normative HGF, 2-level non-volatility HGF). The leaves at the bottom represent response models which integrate and arbitrate between social and individual sources of information (“Arbitrated”) or exclusively consider social (“Advice”) or individual (“Card”) information. (b) Random effects Bayesian model selection revealed one winning model, the Arbitrated 3-level HGF. Posterior model probabilities or $p(m|y)$ indicated that this model best explained participants’ behaviour in the majority of the cases.

250

251 *Is the parameter estimation robust?*

252 The winning 3-level full HGF model includes multiple parameters that need to be
 253 estimated. A general question is whether these parameters are “practically
 254 identifiable”, i.e., whether their values can be recovered accurately given the actual
 255 experimental design. To examine this question, we simulated responses based on all
 256 participants’ maximum-a-posteriori estimates of the parameters, and then fitted the
 257 model to those simulated responses in order to test whether we could recover the
 258 same parameter estimates.

259 To assess and compare degrees of parameter recovery, we categorized it in terms of
 260 effect sizes, i.e., whether the relationship between the original and the recovered
 261 values indicates small, medium, or large effect sizes as quantified by Cohen’s f . For a
 262 multiple regression analysis, a Cohen’s f above 0.4 is conventionally regarded as a

263 large effect size. Based on this criterion, we could recover all parameters well, as all
 264 Cohen's f values equaled or exceeded 0.4 (see *Figure 2 - figure supplement 1*).

265 ***Do participants differ in how they learn from advice and use it to predict lottery***
 266 ***outcomes?***

267 Three parameters modulated the arbitration signal of the winning model. These
 268 included: (i) κ or the coupling between the two hierarchical levels that determined the
 269 impact of volatility on the inferred predictions of each information source (Eq. 6), (ii)
 270 ϑ , determining the variance of the volatility (Eq. 12), and (iii) ζ , the social bias which
 271 reflected the reliance on the advice independent of its reliability (Eq. 19). Both
 272 coupling κ and volatility parameter ϑ did not differ significantly between learning
 273 from individual and social information ($t(36)= 0.28$, $p=0.77$ for κ and $t(36)= -1.59$,
 274 $p=0.12$ for ϑ ; *Figure 4a-b*). In fact, they were highly correlated: $r_1=0.55$, $p_1=0.003$ for κ
 275 and $r_2=0.64$, $p_2=0.001$ for ϑ . This result suggests that participants learned similar
 276 amounts from individual (volatile card probabilities) and social (advisor fidelity)
 277 information.

278 The reliability-independent social bias parameter ζ differed significantly from zero
 279 ($t(36)= 5.09$, $p = 1.07e-05$). Importantly, since the social bias parameter ζ is coded in
 280 log-space, the prior value of zero refers to a uniform weighting of the two cues in
 281 linear parameter space. Thus, on average, participants relied more on the advisor's
 282 recommendations compared to their own sampling of the card outcomes (*Figure 4c*).

283 ***Do the response model parameter estimates explain wager behaviour?***

284 Decisions of how many points participants were willing to wager on a given trial (a
 285 measure of confidence) were related to several model-based quantities, including
 286 (irreducible) uncertainty of the agent's beliefs about the decision, arbitration, and the
 287 estimated volatility of the advisor's intentions (belief uncertainty: $t(37)= -10.37$,
 288 $p_{bonf}=1.0e-11$; arbitration: $t(37)=5.16$, $p_{bonf}=5e-05$; and estimated advisor volatility:
 289 $t(37)=-7.41$ $p_{bonf}= 4.75e-08$) (*Figure 5*). The stronger the bias to arbitrate in favour of
 290 social information, the more points participants wagered. Conversely, estimated
 291 advisor volatility was negatively associated with the amount wagered: the higher the
 292 estimated advisor volatility, the fewer points participants were willing to wager on a
 293 given trial (see Table 1 for the priors over the parameters, Table 2b for all parameter
 294 estimates, and *Figure 5* for the trial-wise influence of the average computational
 295 quantities on wager amount).

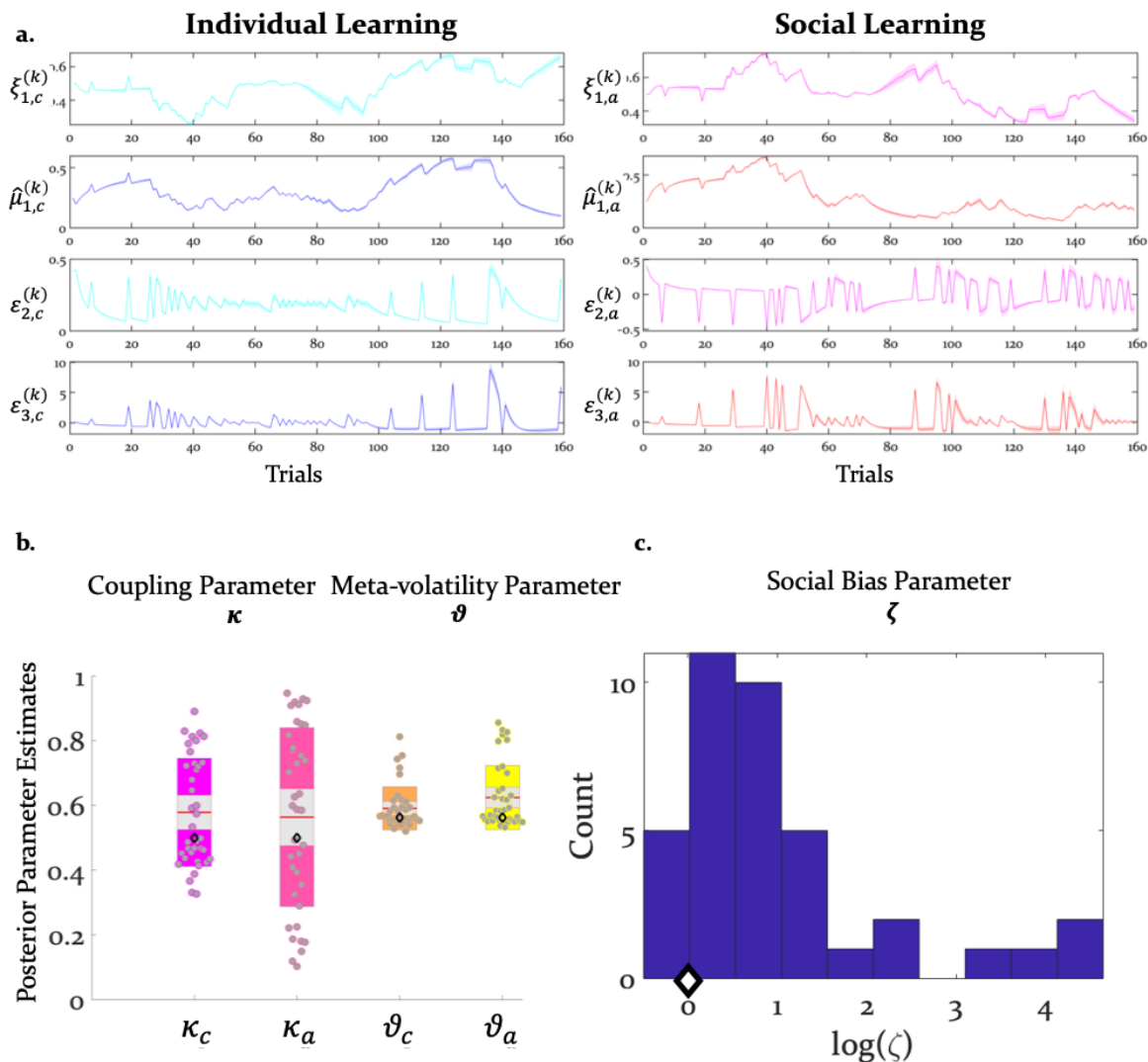


Figure 4

Figure 4| **Inference and arbitration of individual and social learning:** (a) Average trajectories for arbitration and hierarchical precision-weighted PEs for individual and social learning (see Methods for the exact equations): ξ_a = arbitration in favour of the advice (Eq. 19); ξ_c = arbitration in favour of individually estimated card colour probability (Eq. 20). $\hat{\mu}_{1,a}$ = estimated advice accuracy (Eq. 4); $\hat{\mu}_{1,c}$ = individually estimated card colour probability (Eq. 18). $\varepsilon_{2,a}$ = precision-weighted prediction error (PE) of advisor fidelity (Eq. 8); $\varepsilon_{2,c}$ = unsigned (absolute) precision-weighted PE of card outcome (absolute value of Eq. 14). $\varepsilon_{3,a}$ = precision-weighted advice volatility PE (Eq. 13); $\varepsilon_{3,c}$ = precision-weighted card colour volatility PE (Eq. 15). Line plots were generated by averaging the computational trajectories of the winning (Arbitrated 3-HGF: Figure 2) model across all participants for each of the 160 trials. The shaded area around each line depicts +/- standard error of the mean over participants. (b) Group means, standard deviations and prior values for the perceptual model parameters determining dynamics of computational trajectories in (a). Jittered participant-specific estimates are plotted for each perceptual model parameter, red lines indicate the group mean, grey areas reflect 1 SD of the mean, and coloured areas the 95% confidence intervals of the mean. (c) Distribution of $\log(\zeta)$ values. In (b) and (c), black diamonds denote the priors of each parameter (for details, see Table 1).

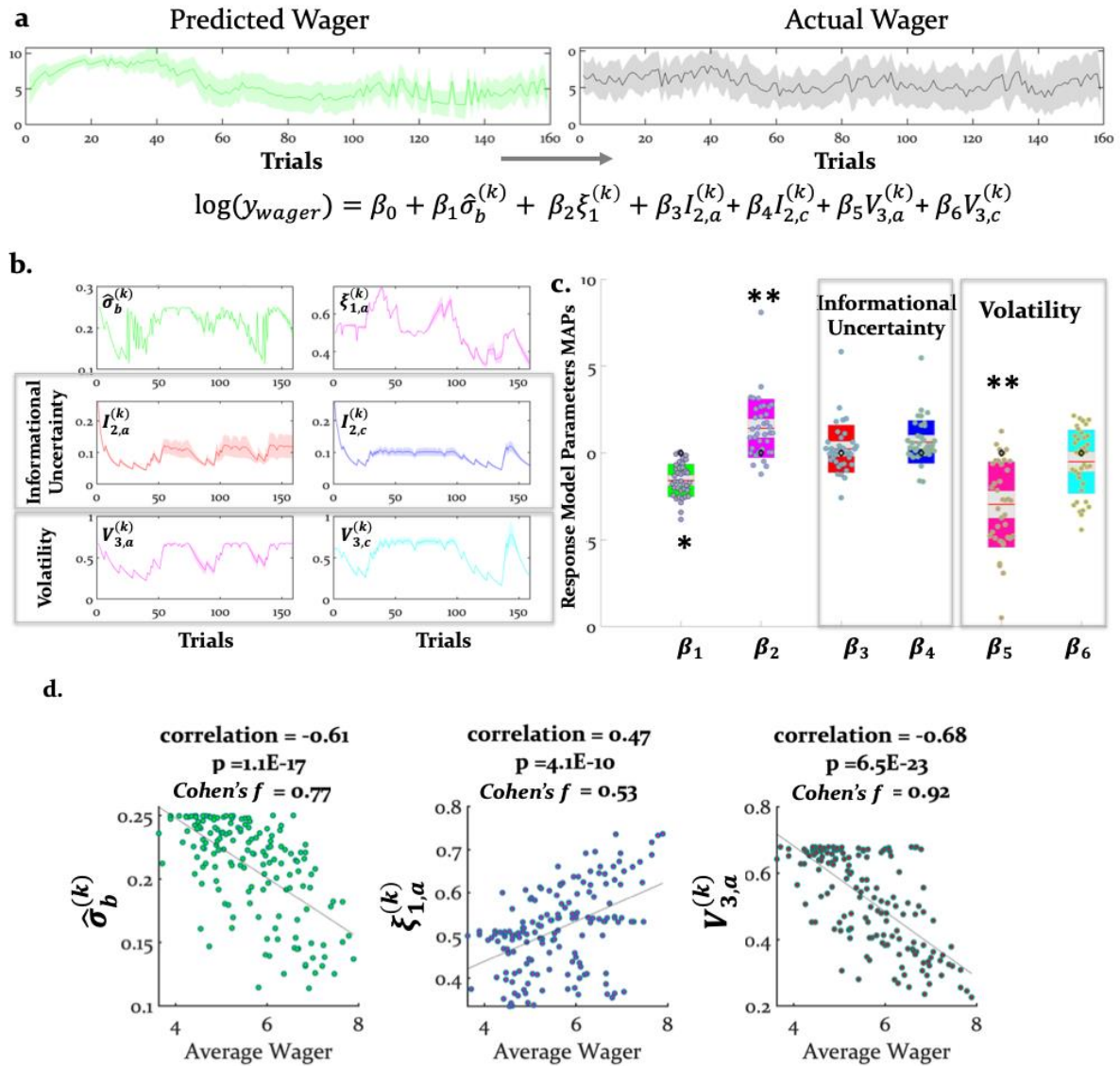


Figure 5

Figure 5| **Computational quantities and model parameters explaining wager amount:** (a) With our response model, we predicted that the actual trial-wise wager (right) could be explained (left and bottom) by the six key trajectories (see Eq. 21) given in (b). These include (i) (irreducible) belief uncertainty (based on the integrated belief of individual and advice predictions; Eq. 24); (ii) arbitration in favour of advice (Eq. 19); (iii) informational uncertainty (Eq. 24) and volatility of the advice (Eq. 25) and (iv) informational uncertainty (Eq. 26) and volatility of the card (Eq. 28). (a) and (b) show group averages (see Methods for the exact equations). For the model-based parameters, the line plots were generated by averaging the computational trajectories of the winning (Arbitrated 3-HGF: Figure 2) model across all participants for each of the 160 trials. The shaded areas depict +/- standard error of this mean over participants. (c) Group means, standard deviations and prior values for the response model parameters determining the impact of those trajectories (i.e., uncertainties and arbitration) on trial-wise wager amount. Jittered raw data are plotted for each parameter. Red lines indicate the mean, grey areas reflect 1 SD from the mean, and the coloured areas the 95% confidence intervals of the mean. The black diamonds denote the prior of the parameters, which in this case is zero. * $p < 0.05$, ** $p < 0.001$. (d) Scatter plots with average actual wager on the x-axis and average of the computational variables assumed to impact the trial-wise wager: belief uncertainty, arbitration in favour of advice, and volatility of advice on the y-axes, respectively. The correlation coefficients (with corresponding p values), regression slopes, and effect sizes (Cohen's f) are included to quantify the relationship between the

actual wager and the computational quantities that showed a significant relation to wagers.

297

298 *Do the model parameter estimates explain perceived advice accuracy and wager*
 299 *amount?*

300 We aimed to examine at the behavioural level whether the model predictions were
 301 consistent with participants' perceptions of the advice accuracy during the
 302 experiment. Participants judged advice accuracy (i.e., helpful, misleading, or neutral
 303 with regard to predicting actual card colour) in a multiple-choice question presented 5
 304 times during the experiment (initial/prior: 1st trial; stable advice, stable card phase =
 305 (14th trial); stable advice, volatile card phase (49th trial); volatile advice, volatile card
 306 phase (73rd trial); volatile advice, stable card phase = 115th trial). We first tested
 307 whether the responses to these questions positively related to estimates of advice
 308 accuracy ($\hat{\mu}_{1,a}^{(k)}$) that were extracted from the winning model. A linear regression
 309 analysis demonstrated that the inferred advice accuracy or $\hat{\mu}_{1,a}^{(k)}$ measured at the time
 310 of the multiple-choice question, predicted participants' selections. Specifically, the
 311 estimated beta parameter estimate across all task phases was significantly different
 312 from zero ($t(36) = 4.71, p = 3e-05$). These findings suggest that the model predicted
 313 independently (and discretely) measured perception of advice accuracy, in-keeping
 314 with the internal validity of the model.

315 Next, we tested whether the wager amounts predicted by the model correlated with
 316 participants' actual wagers. In all four conditions of the task, the predicted wager
 317 significantly correlated with the number of points participants actually wagered: (i)
 318 advice stable phase $r_1 = 0.62, p_1 = 3e-05$; (ii) advice volatile phase $r_2 = 0.63, p_2 = 2e-05$;
 319 (iii) card stable phase $r_3 = 0.81, p_4 = 9e-10$; and (iv) card volatile phase $r_4 = 0.80, p_4 = 1e-$
 320 09 ; *Figure 5 - figure supplement 1*). These findings suggest that the winning model
 321 explained variation in (the continuously measured) actual wager amount.

322 *Do the model parameter estimates explain participants' self-reports?*

323 We used classical multiple regression and post-hoc tests to examine whether the
 324 model parameter estimates extracted from the winning model (M_1) explained
 325 participants' advisor ratings, as measured by debriefing questions after the main
 326 experiment outside the scanner. Participants who reported that the advisor
 327 intentionally tried to help or mislead at different phases of the task showed a trend
 328 towards a larger estimate of the social weighting parameter ζ ($df = (1,36), F = 3.49, p =$
 329 0.06). Moreover, advice helpfulness ratings were explained by model parameter
 330 estimates ($R^2 = 32.2\%, F = 2.46, p=0.04$). This effect was primarily driven by parameter
 331 κ_a ($r(37)=0.47, p=0.0026$), indicating that participants who rated the advice as being
 332 helpful showed stronger coupling between two levels of the hierarchical model. More
 333 specifically, participants who rated the advice as more helpful displayed higher κ_a

334 values, i.e., increased sensitivity to the changing phases of advice validity, adjusting
335 their wagering behaviour more strongly to the advisor's strategy. Thus, not only did the
336 participants perceive the advice in our task as intentional and helpful, our model also
337 explained some of these impressions.

338 *Neural signatures of arbitration*

339 Using behaviourally fitted computational trajectories to generate participant-specific
340 GLMs for model-based fMRI analysis, we examined how the brain arbitrates between
341 social and individual learning systems. We conceptualised the learning and arbitration
342 process as hierarchical Bayesian inference, and fitted the participant-specific
343 trajectories that reflect arbitration (Eq. 20) to fMRI data.

344 Hierarchical precision-weighted PE signals were replicated in the same dopaminergic
345 and frontoparietal regions as in previous studies using other sensory and social
346 learning domains (see Iglesias et al., 2013; Diaconescu et al., 2017), indicating that the
347 modifications in the experimental paradigm did not affect basic learning processes
348 (see *Figure 6 - figure supplement 1-2*).

349 Undirected tests for arbitration activity identified ventral prefrontal regions, such as
350 the right ventromedial PFC (peak at [0, 46, -8]) and the right orbitofrontal cortex
351 (OFC) [32, 30, -14]. Interestingly, frontal activations also included the right frontopolar
352 cortex [4, 54, 30] and ventrolateral prefrontal cortex (VLPFC) [50, 36, 0], regions
353 previously associated with arbitration between model-based and model-free forms of
354 individual learning (Lee et al., 2014) (*Figure 6*). The right VLPFC showing arbitration-
355 related effects at [48, 35, -2] significantly overlapped with the arbitration-related
356 reliability activations detected by Lee and colleagues, supporting the notion that
357 arbitration is to some extent domain-independent.

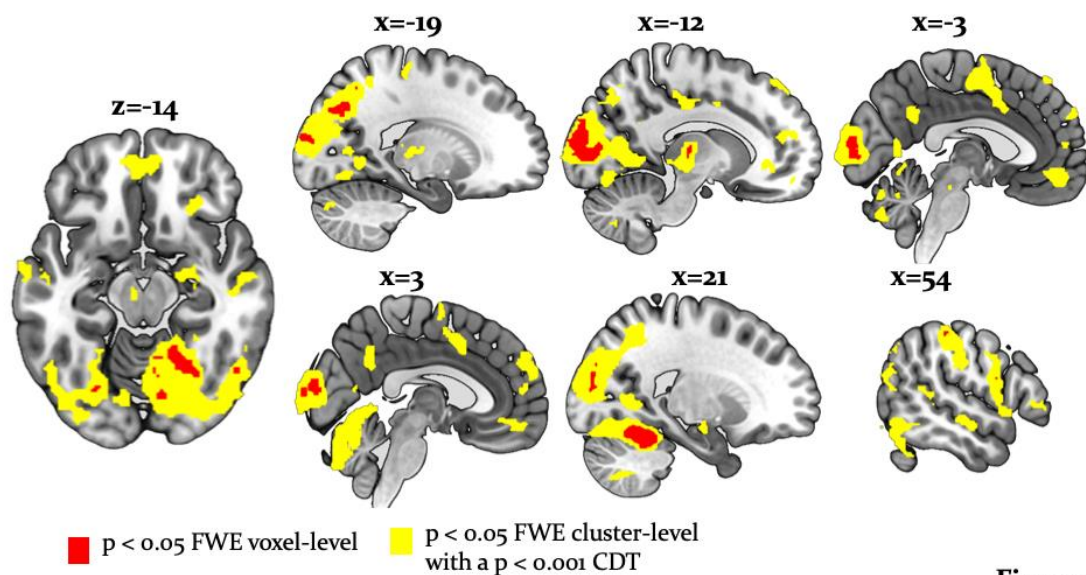


Figure 6

Figure 6| **Whole-brain undirected arbitration signals:** Effects of arbitration in favour of one or the other source of information were detected in ventromedial PFC, orbitofrontal cortex, right frontopolar cortex, VLPFC, the left midbrain, bilateral fusiform gyrus, lateral occipital gyrus, lingual gyrus, anterior insula, right amygdala, left thalamus, right cerebellum, bilateral middle cingulate sulcus and SMA. The figure shows whole-brain FWE-corrected voxel (red) - and cluster-level-corrected (yellow) results of an undirected *F*-test, $p < 0.05$ (CDT = cluster defining voxel-level threshold).

358 In addition, we found that a wide network of cortical and subcortical regions
 359 contributes to arbitration that included occipital areas, the anterior insula, left
 360 thalamus, left putamen, bilateral middle cingulate sulcus, supplementary motor area
 361 (SMA) [-4, 0, 56], left dorsal middle cingulate gyrus [-10, -26, 44], the right amygdala
 362 [18, -10, -16] and the left midbrain [-6, -20, -10] (Table 3, Figure 6). Thus, a network of
 363 cortical and subcortical regions contributed to arbitration.

364 Directed tests for arbitration in favour of individual over social information identified
 365 activity increases in the right dorsolateral PFC [36, 46, 30], left SMA/anterior cingulate
 366 sulcus [-2, -8, 52] and the midbrain [-6, -18, -12] (Figure 7a). The BOLD signal change in
 367 these regions peaked during the time window of the wager decision. In summary,
 368 primarily dorsal regions of PFC were modulated by arbitration in favour of
 369 individually estimated card probability.

370 Conversely, activity in the right amygdala, VLPFC, orbitofrontal and ventromedial PFC
 371 was modulated by arbitration in favour of the advisor's suggestions (Figure 7b).
 372 Outside PFC, the right anterior TPJ [56, -52, 24], right superior temporal gyrus [52, -18,
 373 -8], and right precuneus [6, -52, 32] showed similar effects (Tables 4 and 5 for the
 374 entire list of brain regions). Thus, primarily ventral regions of PFC together with
 375 temporal and parietal regions were more active during arbitration in favour of social
 376 information.

Arbitrating between social and individual learning systems

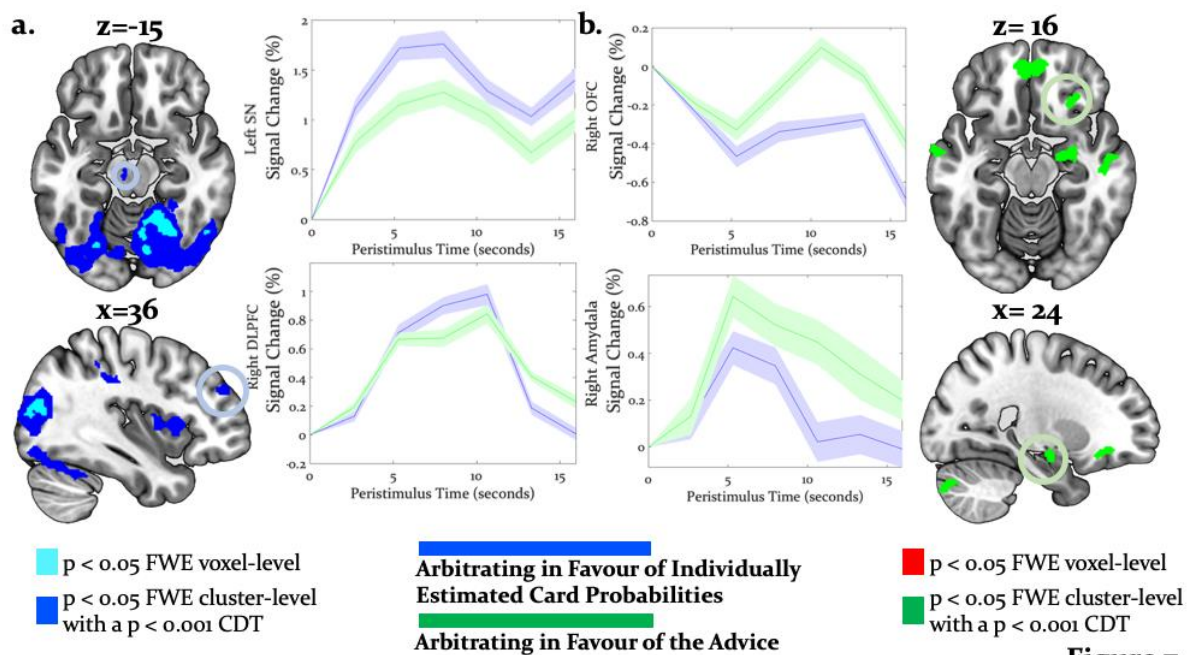


Figure 7

Figure 7| **Neural arbitration directed to specific source of information:** (a) Activity in the left midbrain (substantia nigra (SN)) [-6, -20, -10] (top) and the right DLPFC [36, 46, 30] (bottom) during the prediction of card colour increased more when participants arbitrated in favour of individually estimated card colour probability as compared to the advisor's suggestions (whole-brain FWE cluster-level corrected, $p < 0.05$). (b) Activity in right (OFC [32, 30, -14] (top) and in right amygdala [18, -10, -16] (bottom) increased more when participants arbitrated in favour of the advisor's suggestion than when they arbitrated in favour of the individually learned estimates of card probability (whole-brain FWE cluster-level corrected, $p < 0.05$). The line plots reflect the average BOLD signal activity in the respective significantly activated cluster aligned to the onset of advice presentation relative to pre-advice baseline averaged across trials for one representative participant in midbrain and DLPFC (a) or OFC and amygdala (b). The shaded areas depict \pm standard error of this mean. In this figure the scales reflect t -values.

377 To examine effects of arbitration in dopaminergic, cholinergic, and noradrenergic
 378 regions we also performed region-of-interest (ROI) analyses using a combined
 379 anatomical mask of dopaminergic, cholinergic, and noradrenergic nuclei. A single
 380 cluster in the right substantia nigra survived small-volume correction ($p < 0.05$ FWE
 381 voxel-level corrected for the entire ROI; peak at [-6, -18, -10]; *Figure 8*). Activity in this
 382 region increased with arbitration in favour of individual estimates of card probabilities
 383 rather than advice.

Arbitrating between social and individual learning systems

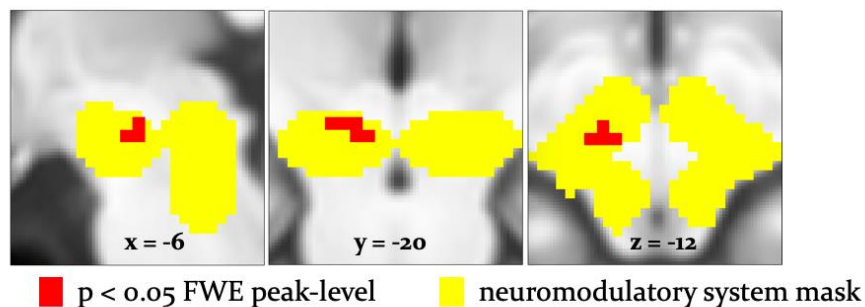


Figure 8

Figure 8| **Arbitration signals in neuromodulatory ROI:** Activation of the dopaminergic midbrain was associated with arbitrating in favour of individually learned information. Activation (red) is shown at $p < 0.05$ FWE corrected for the full anatomical ROI comprising dopaminergic, cholinergic, and noradrenergic nuclei (yellow).

384 It is important to note that these regions showed significantly larger effects of
 385 arbitration than of the amount of points wagered. Responses reflecting arbitration
 386 dominated over responses reflecting wager amount in cerebellar, midbrain, occipital,
 387 parietal, medial prefrontal, and temporal regions including the amygdala. Activity in
 388 precuneus and ventromedial prefrontal cortex in turn correlated with wager amount
 389 (Figure 9). As wager amount can be taken as a proxy for decision value or confidence
 390 (Lebreton et al., 2015), these data suggest that arbitration signals arise on top of
 391 decision value and confidence. Moreover, we captured arbitration as a model-derived,
 392 continuous, and time-resolved variable. Thus, our findings elucidate the process
 393 rather than the result of arbitration.

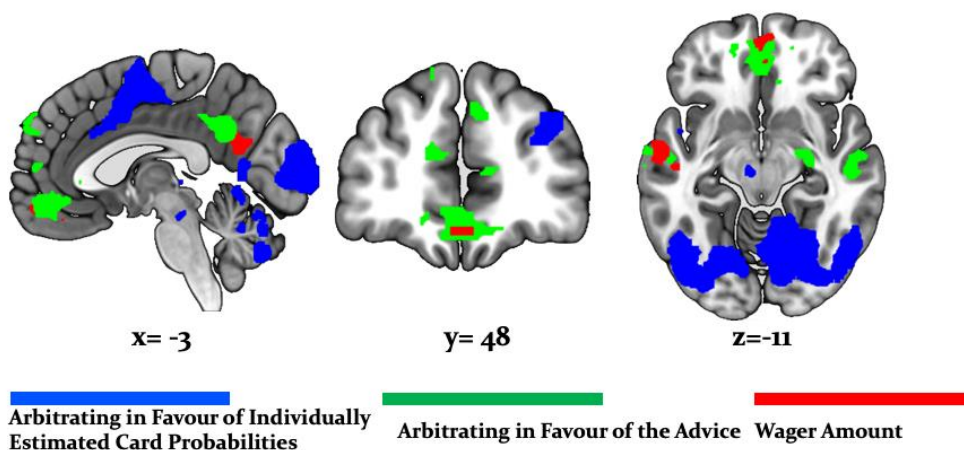


Figure 9

Figure 9| **Arbitration vs. Wager Amount:** Effects of arbitration (individual) (blue) were significantly larger in cortical and subcortical brain regions when compared to wager amount. Effects of arbitration in favour of social information were also significantly larger in ventromedial PFC and amygdala when

compared to wager amount (green). Activity in precuneus and ventromedial PFC regions increased with increases in wager amount (magenta) (whole-brain FWE cluster-level corrected, $p < 0.05$).

394 ***Main effect of stability and interaction with source of information***

395 To examine arbitration from a different angle, we also conducted a factorial analysis.
396 This was possible because we employed a 2 x 2 factorial design – i.e., two sources of
397 information (individual versus social) in two different states (stable versus volatile)
398 (*Figure 10a*). Specifically, we contrasted volatile with stable phases across both
399 information modalities. Volatility is closely tied to arbitration because it potentiates
400 the perceived uncertainty associated with a given information source, and thereby the
401 need to arbitrate. We assumed that arbitration increased when one of the two
402 information sources was perceived as being more stable than the other. In all
403 comparisons, we controlled for decision value and confidence by using the trial-wise
404 wager amount as a parametric modulator in the analysis of brain data. We found two
405 significant results (*Figure 10b*): (i) a main effect of task phase (i.e., stability/volatility),
406 and (ii) a significant interaction of task phase with source of information.

407 By contrasting stable against volatile phases, irrespective of information source, we
408 found that the right supramarginal gyrus, bilateral inferior occipital gyri,
409 postcentral/precentral gyri, and the right anterior insula were more active for stable
410 compared to volatile periods. Furthermore, an interaction between task phase and
411 information source showed preferential activity for stable card information in the
412 midbrain [-4, -22, -8]. Additional activations were detected in the right OFC, VLPFC,
413 dorsomedial cingulate gyrus, and anterior cingulate sulcus/SMA (*Figure 10*, Tables 6
414 and 7). These regions processed stability (vs. volatility) more strongly for card than
415 advice information.

Arbitrating between social and individual learning systems

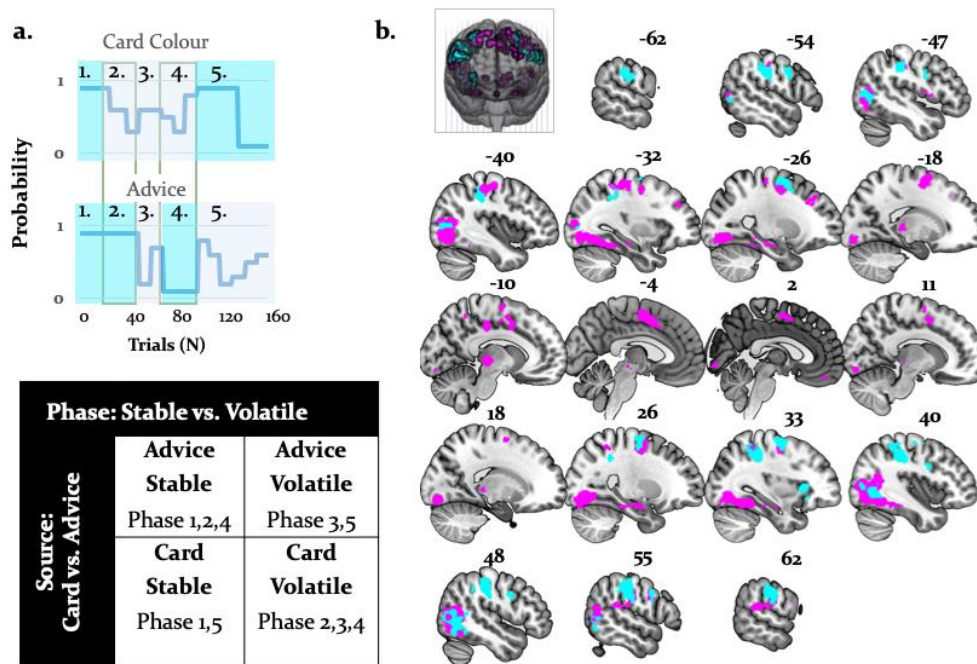


Figure 10

Figure 10 | **Activations related to task phase and interaction with source of information:** (a) The task mapped onto a factorial structure with four conditions: (i) stable card and stable advisor, (ii) stable card and volatile advisor, (iii) volatile card and stable advisor, and (iv) volatile card and volatile advisor, as reflected by the shaded areas: blue (stable), grey (volatile). (b) The main effect of stability irrespective of source of information activated primarily parietal regions and the anterior insula (cyan, whole-brain FWE cluster-level corrected, $p < 0.05$). Moreover, the interaction between task phase and source of information was localized to left midbrain, occipital regions, anterior insula, thalamus, middle cingulate sulcus, SMA, OFC, and VLPFC (magenta, whole-brain FWE cluster-level corrected, $p < 0.05$).

416 Importantly, the regions processing stability (vs. volatility) more strongly for advice
 417 than card information also overlapped with the arbitration signal, and included the
 418 amygdala, the superior temporal sulcus, and the ventromedial PFC (Figure 11). Thus,
 419 model-dependent and model-independent analyses agree in localizing arbitration to
 420 frontoparietal regions in the individual domain and to ventromedial prefrontal and
 421 amygdala regions in the social domain.

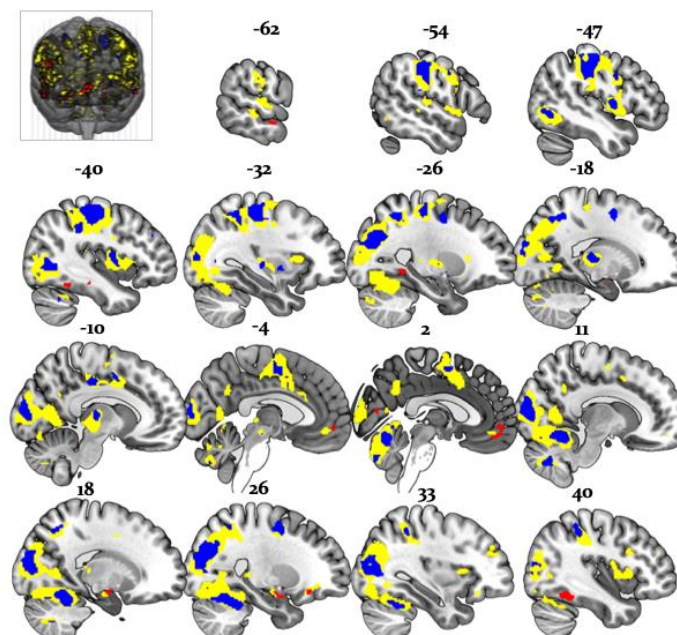


Figure 11

Figure 11 | **Overlap between model-dependent and model-independent results:** Arbitration signal (Eq. 19) (yellow) overlapped with the regions showing an enhanced effect of stability for individual compared to social learning systems (blue) and regions showing enhanced effects of stability in the social compared to individual learning systems (red) (whole-brain FWE peak-level corrected, $p < 0.05$).

422

423 ***Are there neural differences in the representation of social versus non-***
 424 ***social information?***

425 To address the question of distinct representation of social compared to non-social
 426 signatures of learning, we investigated precision-weighted predictions of social and
 427 non-social outcomes. The precision-weighted predictions consist of the two factors
 428 that enter the computation of integrated beliefs (Eqn. 21) about the outcome. The first
 429 reflects the individual card colour estimates weighted by arbitration in favour of the
 430 individually sampled card probabilities (non-social weighting), whereas the second
 431 reflects the predictions of advice accuracy weighted by arbitration in favour of the
 432 advisor (social weighting). Increased effects of non-social compared to social
 433 weighting were detected in bilateral cerebellum, occipital cortices (lingual gyrus,
 434 superior occipital cortex), left anterior cingulate sulcus, right supramarginal gyrus,
 435 and left postcentral gyrus. Conversely, we found increased representations of social
 436 compared to non-social weighting in the left subgenual ACC with a maximum at $[-7,$
 437 $36, -11]$ (Figure 7 – figure supplement 1).

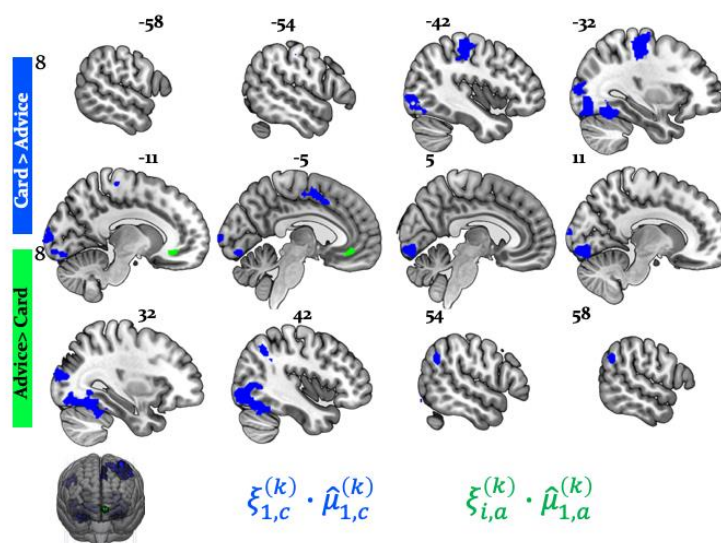


Figure 7 - figure supplement 1 | **Social versus non-social weighting** (equation 21). Whole-brain activations by non-social weighting (one's individual predictions about the card colour outcome) compared to social weighting were detected in bilateral cerebellum, occipital cortices (lingual gyrus, superior occipital cortex), left anterior cingulate sulcus, right supramarginal gyrus, and left postcentral gyrus (blue). Conversely, activation by social weighting was significantly larger in the subgenual ACC (green) (whole-brain FWE cluster-level corrected, $p < 0.05$).

438

439 **Replication of Hierarchical Precision-weighted PE Effects across**
 440 **Learning Domains**

441 To test whether the task used in this study replicates previous findings on the
 442 representation of hierarchical precision-weighted PEs (Diaconescu et al., 2017; Iglesias
 443 et al., 2013), we performed the same model-based analysis using Bayesian surprise
 444 (equivalent to an unsigned precision-weighted outcome PE; the absolute value of
 445 equation 14). Replicating the previous study (Iglesias et al., 2013), we found that the
 446 outcome-related BOLD activity of the substantia nigra positively correlated with the
 447 unsigned precision-weighted outcome PE, as did the bilateral inferior/middle occipital
 448 gyri, anterior insula, (ventro)lateral PFC, and the intraparietal sulcus (*Figure 6 - figure*
 449 *supplement 1a* and Supplementary file 1A). In the previous study, participants
 450 predicted a visual outcome using an auditory cue (Iglesias et al., 2013). Thus, the PE
 451 coding of these regions seems to be sensory modality-independent.

452 With respect to the signed precision-weighted advice PE (equation 8), we also
 453 replicated results from another recent study (Diaconescu et al., 2017) that employed a
 454 different advice-taking paradigm, where participants learned about advice and
 455 integrated it along with unambiguous individual information to predict the outcome
 456 of a binary lottery. Effects of signed precision-weighted advice PE were detected in

457 right VTA/substantia nigra, the right insula, left middle temporal cortex, right
 458 dorsolateral, left dorsomedial and middle frontal cortex (Figure 6 – figure supplement
 459 1b and Supplementary file 1B).

460 Please note that we used the unsigned (absolute) precision-weighted PEs for the card
 461 outcomes, but the signed precision-weighted PEs for the advice. In the case of the
 462 card, the sign of this PE depends on an arbitrarily chosen coding of the colour and the
 463 sign is meaningless (see Iglesias et al., 2013). In contrast, for the advice the sign refers
 464 to the valence and instances of surprise where the advisor was more helpful than
 465 predicted may have a different meaning than instances of surprise where the advisor
 466 was more misleading than predicted (see Diaconescu et al., 2017). For completeness,
 467 we also investigated the neural correlates of the signed reward precision-weighted PE
 468 and noted a similar network of posterior parietal and dorsolateral prefrontal regions.

469

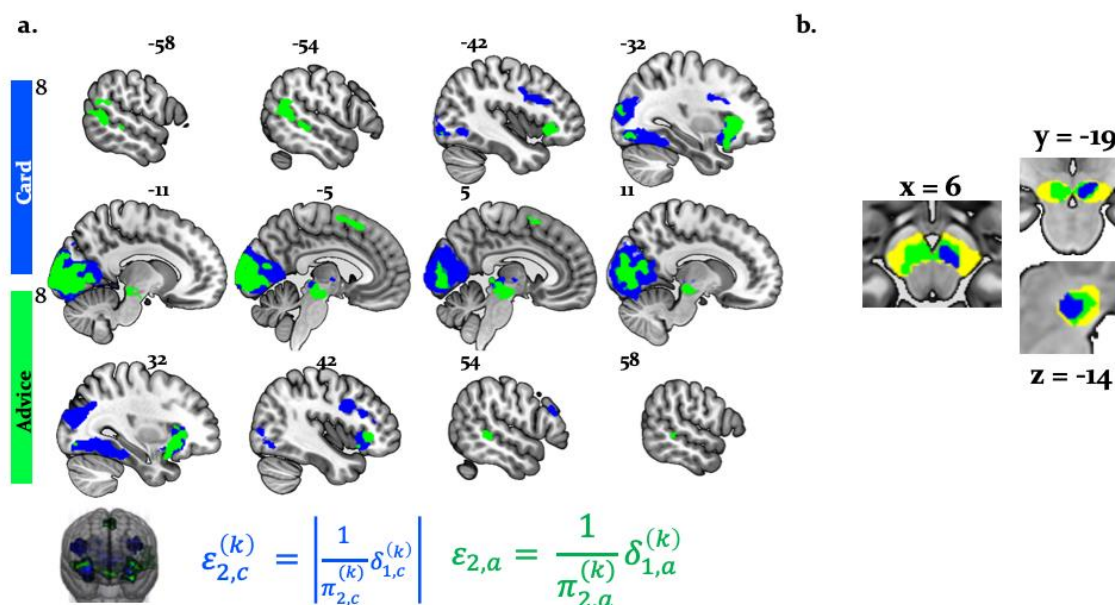


Figure 6 – figure supplement 1 | **Main effects of precision-weighted PEs about card and advice outcomes** (equations 8 and 14). (a) Whole-brain activation by ϵ_2 : Activations by unsigned precision-weighted PE about the card probabilities (blue) were detected in the bilateral inferior/middle occipital gyri, anterior insula, bilateral inferior, medial and middle frontal gyri, and the bilateral intraparietal sulcus (whole-brain FWE peak- and cluster-level corrected, $p < 0.05$). Activations by signed precision-weighted PE about the adviser fidelity (green) were observed in the bilateral fusiform gyrus, lingual gyrus, anterior insula, bilateral supplementary motor area, left middle temporal cortex, right posterior superior temporal sulcus, temporal-parietal junction, bilateral dorsolateral and left dorsomedial prefrontal cortex (whole-brain FWE peak- and cluster-level corrected, $p < 0.05$). (b) Activation of the right VTA was associated with the unsigned precision-weighted PE about the card probabilities (blue) and activation of bilateral VTA/SN associated with the signed precision-weighted prediction error about the adviser fidelity (green). This activation is shown at $p < 0.05$ FWE corrected for the volume of

Arbitrating between social and individual learning systems

our anatomical mask comprising dopaminergic nuclei (**yellow**).

470 Effects of precision-weighted volatility PEs for card outcomes were represented in the
 471 right superior temporal gyrus, supramarginal gyrus, and posterior insula (*Figure 6 –*
 472 *figure supplement 2a*) while the effects of precision-weighted volatility PEs for the
 473 adviser fidelity were encoded in the right anterior supplementary motor area (SMA)
 474 and anterior insula.

475 Finally, we also replicated the finding that higher-level, volatility PEs (equations 13 and
 476 15) were represented in cholinergic regions. This time, however, we observed effects of
 477 advice volatility precision-weighted PEs in the cholinergic nuclei in the tegmentum of
 478 the brainstem, i.e., the pedunculopontine tegmental (PPT) and laterodorsal tegmental
 479 (LDT) nuclei ($p < 0.05$ FWE voxel-level within an anatomical mask including all
 480 cholinergic nuclei) (*Figure 6 – figure supplement 2b*).

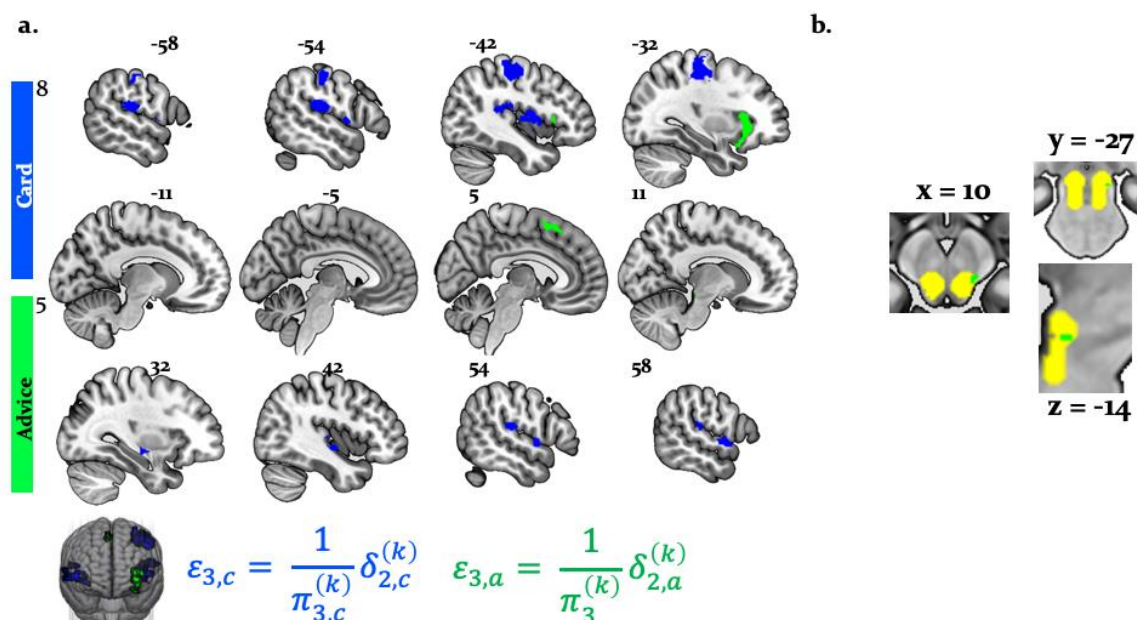


Figure 6 – figure supplement 2 | **Main effects of precision-weighted PEs about card and advice volatility.** (a) Whole-brain activation by ϵ_3 : Whole-brain activations by signed precision-weighted volatility PEs about the card probabilities (**blue**) were detected in the right superior temporal gyrus, supramarginal gyrus, and posterior insula. Whole-brain activations by signed precision-weighted volatility PEs about the adviser fidelity (**green**) were detected in the right anterior SMA and anterior insula (whole-brain FWE cluster-level corrected, $p < 0.05$). (b) Whole-brain activation by ϵ_3 in the PPT/LDT nuclei: Activation of the right cholinergic PPT/LDT associated with the signed precision-weighted volatility prediction error about the adviser fidelity is shown at $p < 0.05$ FWE corrected for the volume of our anatomical mask comprising cholinergic nuclei (**yellow**).

481
 482
 483

484 Discussion

485 Our study shows how healthy participants arbitrate between uncertain social and
486 individual information under varying conditions of stability during a binary lottery
487 task. (*Figure 1*). Participants arbitrated between the two information sources by taking
488 into account their relative precision. The more precise one information source was
489 over the other and the more stable the advisor was perceived to be, the more points
490 participants were willing to wager.

491 By showing that participants tracked the volatility of both the advice and the card
492 colour probabilities (*Figure 3*), our study underscores the importance of volatility in
493 arbitrating between social advice and individual reward-relevant information. At the
494 behavioural level, trial-by-trial accuracy of participant predictions, frequency of taking
495 advice into account, and amount of points wagered on each trial (*Figure 5 - figure
496 supplement 1*) were all reduced by volatility. Thus, in stable compared to volatile
497 environments, the propensity for arbitration in favour of the more precise information
498 source increases. Numerous studies have demonstrated an important role of volatility
499 in higher-level learning (Behrens et al., 2007, 2008; Nassar et al., 2010; Iglesias et al.,
500 2013; Vossel et al., 2013; Diaconescu et al., 2017; Pulcu and Browning, 2017), in-keeping
501 with the present findings.

502 *Evidence for domain-generalty of arbitration in lateral prefrontal cortex*

503 Using both model-based and model-independent (factorial) fMRI analysis, we found
504 that the arbitration signal correlated with activity in dorsolateral and ventrolateral
505 PFC, frontopolar, and orbitofrontal cortex (*Figures 6 and 11*). These findings
506 corroborate previous insights on arbitration between different forms of individual
507 information also pointing to lateral prefrontal cortex (Lee et al., 2014), in line with
508 domain generality for arbitrating. Note though that arbitration activity in the
509 prefrontal cortex followed a self-versus-other axis: dorsal prefrontal activity increased
510 the more strongly participants weighed their own predictions of reward probabilities
511 over the perceived reliability of the advisor. Conversely, activity in the ventromedial
512 PFC and orbitofrontal cortex showed the opposite pattern and increased in activity as
513 participants relied more heavily on their own reward probability estimates relative to
514 the advice (*Figure 7*). Together, arbitration appears to be sensitive to the source of
515 information entering the arbitration process, contrary to an entirely domain-general
516 process.

517 *Arbitration in the dopaminergic system*

518 The results of both model-based and factorial analyses suggest a key role of the
519 midbrain in arbitrating for individual estimates about card colour over advice (*Figure
520 8*). Primate studies found that sustained dopamine neuron activity signalled expected
521 uncertainty (Fiorillo, 2003; Schultz, 2010; Schultz et al., 2008). This was further

522 supported by human pharmacological studies (Burke et al., 2018; Ojala et al., 2018) as
523 well as fMRI research showing possible involvement of dopamine in risk taking and of
524 dopaminergic regions, such as the caudate, anterior insula, ACC and the medial
525 PFC in uncertainty coding (e.g. (Dreher et al., 2006; Preuschoff et al., 2008; Tobler et
526 al., 2009)). In particular, studies employing hierarchical Bayesian models have
527 identified ventral tegmental area/substantia nigra activation correlated to precision of
528 predictions about desired outcomes (Friston et al., 2014; Schwartenbeck et al., 2015).

529 These findings may also underscore the role of dopamine in modulating participants'
530 ability to optimise learning to suit ongoing estimates of environmental volatility.
531 Potential neurobiological mechanisms include meta-learning models, which propose
532 an important role of phasic dopamine signals in training the dynamics of the
533 prefrontal system, to infer on the structure of the environment (Collins and Frank,
534 2016; Wang et al., 2018). Such models imply that improved learning of the structure of
535 the environment, e.g. current levels of volatility, results in more appropriate
536 arbitration adjustment.

537 *Arbitrating in favour of the advisor activates the amygdala and* 538 *orbitofrontal cortex*

539 The amygdala processed perceived reliability of social information, reflected in activity
540 decreasing the more participants discounted their own estimates of rewarded card
541 colour probabilities relative to advisor information. The amygdala has been implicated
542 in processing facial expressions related to affective ToM (Schmitgen et al., 2016) and
543 more generally, processing affective value and motivational significance of various
544 stimuli, including other people (Güroğlu et al., 2008; Zink et al., 2008; Zerubavel et al.,
545 2015). Together, the amygdala may represent the uncertainty of socially-relevant
546 stimuli, inferred from intentions of others.

547 Similar to the amygdala, the orbitofrontal cortex showed a significant interaction
548 between task phase and information source, indicative of arbitrating in favour of social
549 information. This finding is consistent with the hypothesis that the orbitofrontal
550 cortex and other areas of the social brain evolved to enable primates and particularly
551 humans to successfully navigate complex social situations (Dunbar, 2009). This notion
552 received support from strong positive correlations between orbitofrontal cortex grey
553 matter volume and social network size (Powell et al., 2012), as well as sociocognitive
554 abilities (Powell et al., 2010; Scheuerecker et al., 2010). Furthermore, in-keeping with a
555 role of orbitofrontal cortex in mental state attribution for ambiguous social stimuli
556 (Deuse et al., 2016), our findings suggest that this region reduces the uncertainty of
557 social cues that signal changes of intentionality.

558 With respect to social learning signatures, we observed that the sulcus of the ACC
559 represents predictions related to one's own estimates of the card colour outcomes,
560 whereas the subgenual ACC represents predictions about the advisor's fidelity. This is

561 consistent with previous findings that the sulcus of ACC dorsal to the gyrus plays a
562 domain-general role in motivation (Rushworth et al., 2007; Rushworth and Behrens,
563 2008; Apps et al., 2016), whereas the gyrus of the ACC signals information related to
564 other people (Behrens et al., 2008; Apps et al., 2013, 2016; Lockwood, 2016).

565

566 *Implications for mentalizing disorders*

567 An intriguing extension of the current study concerns the question of whether
568 arbitration occurs differently in patients with psychiatric and neurodevelopmental
569 disorders involving ToM processes. If so, how do these processing differences affect
570 behaviour? For example, individuals with autism spectrum disorder may preferentially
571 rely on their own experiences rather than on the recommendations of others. Indeed,
572 they appear to represent social prediction errors less strongly than individuals without
573 autism (Balsters et al., 2017). Accordingly, they may be able to better infer the volatility
574 of the card colour probability compared to the advice in our task. In contrast, patients
575 with schizophrenia may be overly confident about their ability to judge advice validity
576 due to fixed beliefs about the advisor's intentions (Freeman and Garety, 2014) or show
577 an over-reliance on social information in line with accounts of over-mentalization in
578 this disorder (Montag et al., 2011; Andreou et al., 2015). Future work may test these
579 intriguing possibilities.

580 *Limitations*

581 One limitation of our study is that it did not include reciprocal social interactions, but
582 rather used pre-recorded videos of human partners. ToM processes may be more
583 prominent in interactive paradigms (Diaconescu et al., 2014) or interactions that
584 involve higher levels of recursive thinking (Devaine et al., 2014a, 2014b). By extension,
585 our study may have limited generalizability to real-world social interactions. However,
586 assessing arbitration between social and individual information necessitated the
587 standardization of the advice given to each participant. To make the task as close as
588 possible to a realistic social exchange, the videos of the advisor were extracted from
589 trials when they truly intended to help or truly intended to mislead. More
590 importantly, to adequately compare learning from social and individual information in
591 stable and volatile phases, we needed to ensure that the two information types were
592 orthogonal to each other and balanced in terms of volatility.

593 Second, we did not include a non-social control task. Thus, it is unclear how “social”
594 the presently investigated form of learning about the advisor's fidelity and volatility
595 actually is. The differences in activated regions at least suggest that our participants
596 processed the two sources of information differently. However, whether the process
597 we identified is specifically social in nature or rather reflects learning from an indirect

598 information source needs to be examined in future studies by including an additional
599 control condition.

600 In order to distinguish general inference processes under volatility from inference
601 specific to intentionality, we previously included a control task (Diaconescu et al.,
602 2014), in which the advisor was blindfolded and provided advice with cards from
603 predefined decks that were probabilistically congruent to the actual card colour. This
604 control task closely resembled the main task, with the exception of the role of
605 intentionality. Model selection results suggested that participants in the control task
606 did not incorporate time-varying estimates of volatility about the advisor into their
607 decisions. In the current study, we tested this by including models without volatility,
608 but found that they performed substantially worse than models with volatility (see
609 *Figure 2* and Table 2a for details). Thus, our participants appeared to process advisor
610 intentionality.

611 Conclusions

612 Our study indicates that arbitrating between social and individual sources of
613 information corresponds to weighing the relative reliability of each source. This
614 process appears to engage different brain regions for social and individual
615 information, in-keeping with domain specificity. However, the lateral prefrontal
616 cortex appears to adjudicate between several different types of learning, in-keeping
617 with domain generality. These findings contribute to our understanding of arbitration
618 in neurotypical individuals, which may provide a knowledge basis for future insight
619 into disorders with impaired arbitration.

620

621 Materials and Methods

622 Ethics Statement

623 The study was approved by the Ethics Committee of the Canton of Zürich (KEK-ZH
624 2010-0327). All participants gave written informed consent before taking part.

625 Participants

626 We recruited 48 volunteers (mean age 23.6 ± 1.4 , 32 females) who were non-smokers,
627 right-handed, and had normal or corrected-to-normal vision. Participants had no
628 history of neurological or psychiatric illness, or of drug abuse. Psychology students
629 were excluded from participation because of previous exposure to similar advice-
630 taking paradigms in their courses. Participants were asked to abstain from alcohol 24
631 hours prior to the study and from medication, including aspirin, 3 days prior to the
632 study. We did not analyse the data of ten participants: two pilot participants; one
633 participant who stopped the experiment midway due to head pain; one participant
634 who fell asleep; and six participants where stimulus presentation malfunctioned
635 during the experiment. Altogether, 38 participants (mean age 24.2 ± 1.3 ; 26 females)
636 entered the final analysis.

637 Stimuli and task

638 We modified the deception-free binary lottery game of Diaconescu and colleagues
639 (2014). In each trial, the participant had to predict the colour of a card draw – blue or
640 green. Participants could base their predictions on social information and/or on
641 individually experienced recent outcome history (see below). They received social
642 information from the “advisor”, who held up a card in one of the two colours before
643 every draw, recommending to the participant which option to choose. The advisor
644 based his or her suggestion on information that was true with a probability of 80%,
645 although the participants were not informed of this fact. Furthermore, the advisor
646 received monetary incentives to change his or her strategy and thus provide either
647 helpful or misleading advice at different stages of the game (*Figure 1b*) with the
648 average probability of advice being correct in 56% of trials. To compare participants in
649 terms of their learning and decision-making parameters, we needed to standardize the
650 advice. This means that each participant received the same input sequence, i.e., order
651 and type of videos.

652 To display social information in a standardized fashion and gender-match advisors
653 and participants, we created videos from two male and two female advisors, who
654 changed their advice as a function of the incentives in a previously recorded face-to-
655 face session (see Diaconescu et al, 2014). Their advice on each trial was recorded for an
656 entire experimental session and the full-length videos were edited into 2-sec

657 segments, focusing on the advice period. We received informed consent from all
658 advisors in the initial (face-to-face) behavioural study to record and use the advice-
659 giving videos in subsequent studies. All video clips were matched in terms of their
660 luminance, contrast, and colour balance using Adobe Photoshop Premiere CS6.

661 To standardize the advice, avoid implicit cues of deception, and make the task as close
662 as possible to a social exchange in real time, the videos of the advisor were extracted
663 from trials when they truly intended to help or truly intended to mislead. Although
664 each participant received the same advice sequence throughout the task, the advisors
665 displayed in the videos varied between participants, in order to ensure that physical
666 appearance and gender did not impact on their decisions to take advice into account.
667 Advisor-to-participant assignment was randomized (within the gender-matching
668 constraint) and balanced. We found no differences in performance and degree of
669 reliance on advice between the four advisors: $F(1,36) = 1.82, p = 0.16$.

670 In contrast to previous studies (Diaconescu et al., 2014, 2017), participants had to infer
671 card colour probabilities (blue versus green) from individually experienced outcomes
672 of previous trials rather than being provided with (changing) pie charts explicitly
673 stating the probabilities. In each trial, they had to arbitrate between following either
674 social information (previous advice, inferring on intention) or individual information
675 (previous cards, inferring on probability). Moreover, also in contrast to previous
676 studies, for each lottery prediction, participants wagered between one and ten points
677 to indicate how confident they were about their predictions. The tick mark on the
678 wager bar was randomly positioned in each trial to avoid providing a reference point
679 (a regression analysis confirmed that the starting position of the wager indeed failed
680 to explain each participant's trial-wise wager selection, $t(37) = -0.89, p = 0.31$).
681 Depending on the correctness of the prediction, the wager was added to or subtracted
682 from the cumulative score and thereby affected the participant's payment at the end of
683 the experiment (see below).

684 Each trial (*Figure 1a*) began with a video of the advisor holding up a card, followed by a
685 decision screen in which participants selected the blue or green card. At the next
686 screen, they were asked to provide the wager. The subsequent outcome screen
687 revealed the drawn card. Finally, the updated cumulative score appeared. The colour-
688 to-button assignment used to convey the lottery prediction (blue or green) and the
689 orientation of the wager bar were randomized between participants to prevent
690 confounding with visuomotor processes.

691 Across trials, the colour-reward probabilities and the advisor intentions varied
692 independently of each other. In other words, the probability distributions of the two
693 information sources – card colour and advice – were designed to be statistically
694 independent. This allowed for a 2 x 2 factorial design structure, where trials could be
695 divided into four conditions: (i) stable card and stable advisor, (ii) stable card and

696 volatile advisor, (iii) volatile card and stable advisor, and (iv) volatile card and volatile
697 advisor in a total of 160 trials (*Figure 1b*). Based on this factorial structure, we
698 predicted that arbitration signals would vary as a function of the stability of each
699 information source.

700 Procedure

701 We explained the deception-free task to participants and ensured their
702 comprehension with a written questionnaire, which required them to describe the
703 instructions in their own words. The task instructions, which were originally
704 presented to participants in their native German, were translated into English for the
705 purpose of this paper. Pronouns were adapted to the advisor's gender: "The advisor
706 has generally more information than you about the outcome on each trial. The
707 objective of the advisor is to use this information to guide your choices and reach
708 his/her own goals. Note that the advisor does not have 100% accurate information
709 about which colour "wins" and he/she might be incorrect. Nevertheless, he/she will on
710 average have better information than you and his/her advice may be valuable to you."
711 The actual experiment was divided into two sessions, with a two-minute break in the
712 middle when participants could close their eyes and rest. The first session included 70
713 trials and the second session 90 trials.

714 To test the construct validity of our computational model and verify whether
715 participants inferred on the advisor's fidelity, we asked them to rate the usefulness of
716 the advisor's card recommendation based on a multiple choice question (including,
717 "helpful," "misleading," or "neutral"). This question was presented six times
718 throughout the task and responses allowed us to assess whether at any point in time,
719 the model could significantly predict participants' responses.

720 Participants could earn a bonus of 10 Swiss Francs for a cumulative score of at least
721 380 points, and a bonus of 20 Swiss Francs for winning more than 600 points.
722 Importantly, participants were not given any information about the bonus thresholds
723 in order to prevent induction of local risk-seeking or risk-averse wagering behaviour
724 (reference point effects) when participants were close to a threshold. Participants on
725 average reached the first reward bonus and were paid 82.3 ± 8.4 Swiss Francs
726 (including the performance-dependent bonus) at the end of the study. After the task,
727 participants completed a debriefing questionnaire, and we revealed to them the
728 general trajectory of the advisor's intentions.

729 Data Acquisition and Preprocessing

730 We acquired functional magnetic resonance images (fMRI) from a Philips Achieva 3T
731 whole-body scanner with an 8-channel SENSE head coil (Philips Medical Systems,
732 Best, The Netherlands) at the Laboratory for Social Neural Systems Research at the
733 University Hospital Zurich. The task was presented on a display at the back of the

734 scanner, which participants viewed using a mirror placed on top of the head coil. The
735 first five volumes of each session were discarded to allow for magnetic saturation.

736 During the task, we acquired gradient echo T₂*-weighted echo-planar imaging (EPI)
737 data with blood-oxygen-level dependent (BOLD) contrast (slices/volume = 33; TR =
738 2665 ms; voxel volume = 2 x 2 x 3 mm³; interslice gap = 0.6 mm; field of view (FOV) =
739 192 x 192 x 120 mm; echo time (TE) = 35 ms; flip angle = 90°). The images were oblique,
740 slices with -20° right-left angulation from a transverse orientation. The entire
741 experiment comprised 1300 volumes, with 600 volumes in the first session and 700 in
742 the second. Heart rate and breathing of the participants were recorded for
743 physiological noise correction purposes using ECG and a pneumatic belt, respectively.

744 We also measured the homogeneity of the magnetic field with a T₁-weighted 3-
745 dimensional (3-D) fast gradient echo sequence (FOV = 192 x 192 x 135 mm³; voxel
746 volume = 2 x 2 x 3 mm³; flip angle = 6°; TR = 8.3 ms; TE₁ = 2 ms; TE₂ = 4.3 ms). After
747 the experiment, we acquired T₁-weighted structural scans from each participant using
748 an inversion-recovery sagittal 3-D fast gradient echo sequence (FOV = 256 x 256 x 181
749 mm³; voxel volume = 1 x 1 x 1 mm³; TR = 8.3 ms; TE = 3.9 ms; flip angle = 8°).

750 The software package SPM12 version 6470 (Wellcome Trust Centre for Neuroimaging,
751 London, UK; <http://www.fil.ion.ucl.ac.uk/spm>) was used to analyse the fMRI data.
752 Temporal and spatial preprocessing included slice-timing correction, realignment to
753 the mean image, and co-registration to the participant's own structural scan. The
754 structural image underwent a unified segmentation procedure combining
755 segmentation, bias correction, and spatial normalization (Ashburner and Friston,
756 2005); the same normalization parameters were then applied to the EPI images. As a
757 final step, EPI images were smoothed with an isotropic Gaussian kernel of 6mm full-
758 width half-maximum.

759 BOLD signal fluctuations due to physiological noise were modelled with the PhysIO
760 toolbox (<http://www.translationalneuromodeling.org/tapas>) (Kasper et al., 2017) using
761 Fourier expansions of different order for the estimated phases of cardiac pulsation (3rd
762 order), respiration (4th order) and cardio-respiratory interactions (1st order; (Glover et
763 al., 2000)). The 18 modelled physiological regressors entering the subject-level GLM
764 along with the six rigid-body realignment parameters and regressors of interest were
765 used to account for BOLD signal fluctuations induced by cardiac pulsation,
766 respiration, and the interaction between the two.

767 Computational Modelling

768 We formalised arbitration in terms of hierarchical Bayesian inference as the relative
769 perceived reliability of each information source. In other words, arbitration was
770 defined as a ratio of precisions: the precision of the prediction about advice accuracy
771 and colour probability, divided by the total precision. The precisions of the predictions

772 afforded by each learning system are obtained by applying a two-branch hierarchical
 773 Gaussian filter (Mathys et al., 2011, 2014) along with a response model (see below) to
 774 participants' trial--wise behaviour (i.e., choices and wagers).

775 ***Learning Model: Hierarchical Gaussian Filter***

776 The HGF is a model of hierarchical Bayesian inference widely used for computational
 777 analyses of behaviour (e.g., (Iglesias et al., 2013; Vossel et al., 2013; Hauser et al., 2014;
 778 de Berker et al., 2016; Marshall et al., 2016). To apply it to our task, we assumed that
 779 the rewarded card colour (individual learning) and the advice accuracy (social
 780 learning) varied as a function of hierarchically coupled hidden states: $x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}$.
 781 They evolved in time by performing Gaussian random walks. At every level, the step
 782 size was controlled by the state of the next-higher level (*Figure 2a*).

783 Starting from the bottom of the hierarchy, states $x_{1,a}$ and $x_{1,c}$ represented binary
 784 variables, namely the advice accuracy (1 for accurate, 0 for inaccurate) and the
 785 rewarded card colour (1 for blue, 0 for green). All states higher than x_1 were
 786 continuous. They denoted (i) the advisor fidelity and tendency for a given card colour
 787 to be rewarded, and (ii) the rate of change of the advisor's intentions and card colour
 788 contingencies, respectively. Four learning parameters, namely, κ_a , κ_c , ϑ_a and
 789 ϑ_c determined how quickly the hidden states evolved in time. Parameter κ represented
 790 the degree of coupling between the second and the third levels in the hierarchy,
 791 whereas ϑ determined the variability of the volatility over time (meta-volatility). This
 792 constitutes the *generative model* of the process producing the outcomes observed by
 793 participants. The overall model and the formal equations describing these relations in
 794 a social learning context are detailed in Diaconescu et al., 2014.

795 ***Model Inversion: Agent-specific arbitration***

796 In accordance with Bayes' rule, we assumed that participants who make inferences on
 797 advice and card colours form posterior beliefs over the hidden states (i.e., congruency
 798 of advice with actual card colour; rewarded card colour) based on the outcomes they
 799 observe. Model inversion is the application of Bayes' rule to a generative model such as
 800 the one described above. This leads to a *recognition* or *perceptual model*, which
 801 describes participants' beliefs about hidden states. Assuming Gaussian distributions,
 802 these agent-specific beliefs are denoted by their summary statistics, i.e., μ (mean) and
 803 σ (variance/uncertainty) or the inverse of the variance $\pi = 1/\sigma$ (precision/certainty).

804 Using variational Bayes under the mean-field approximation, simple analytical trial-
 805 by-trial update equations can be derived. The posterior means $\mu_i^{(k)}$ or predictions on
 806 each trial k at each level of the hierarchy i change as a function of precision-weighted
 807 prediction errors (PEs):

$$\Delta\mu_i^{(k)} \propto \frac{\hat{\pi}_{i-1}^{(k)}}{\pi_i^{(k)}} \delta_{i-1}^{(k)} \quad (1)$$

808

809 Throughout, predictions or prior beliefs about the hidden states (before observing the
810 outcome) are denoted with a hat symbol. States $\hat{\pi}_{i-1}^{(k)}$ and $\pi_i^{(k)}$ represent the estimated
811 precisions about (i) the input from the level below (i.e., precision of the data – advice
812 congruency or rewarded card colour) and (ii) the belief at the current level,
813 respectively.

814 The updates about the advisor’s fidelity are:

$$\Delta\mu_{2,a}^{(k)} = \frac{1}{\pi_{2,a}^{(k)}} \delta_{1,a}^{(k)} \quad (2)$$

where

$$\delta_{1,a}^{(k)} = u^{(k)} - \hat{\mu}_{1,a}^{(k)}. \quad (3)$$

815 Variable $u^{(k)}$ is the sensory input at trial k , where given advice is either accurate
816 ($u^{(k)} = 1$) or inaccurate ($u^{(k)} = 0$). Furthermore, $\hat{\mu}_{1,a}^{(k)}$ corresponds to the logistic
817 sigmoid of the current expectation of the advisor fidelity:

$$\hat{\mu}_{1,a}^{(k)} = s(\mu_{2,a}^{(k-1)}) = \frac{1}{1 + \exp(-\mu_{2,a}^{(k-1)})} . \quad (4)$$

818 The current belief precision is equivalent to:

819

$$\pi_{2,a}^{(k)} = \hat{\pi}_{2,a}^{(k)} + \frac{1}{\hat{\pi}_{1,a}^{(k)}} \quad (5)$$

with the predicted (i) belief precision $\hat{\pi}_{2,a}^{(k)}$ and (ii) the sensory, lower-
level precision about the advice $\hat{\pi}_{1,a}^{(k)}$ computed as:

$$\hat{\pi}_{2,a}^{(k)} = \frac{1}{\frac{1}{\pi_{2,a}^{(k-1)}} + \exp(\kappa\mu_{3,a}^{(k-1)} + \omega)} \quad (6)$$

$$\hat{\pi}_{1,a}^{(k)} = \frac{1}{\hat{\mu}_{1,a}^{(k)}(1 - \hat{\mu}_{1,a}^{(k)})}. \quad (7)$$

820 Thus, the advice belief precision depends on (i) the predicted sensory precision of the
821 input, $\hat{\pi}_1^{(k)}$ and (ii) the predicted volatility, $\mu_{3,a}^{(k-1)}$ from the level above via equation 6.

822 The precision-weighted PE about the advice, which is used to update the belief about
823 fidelity is equivalent to:

$$\varepsilon_{2,a} = \frac{1}{\pi_{2,a}^{(k)}} \delta_{1,a}^{(k)} \quad (8)$$

824 Going up the hierarchy, the updates of advice volatility are proportional to precision-
825 weighted PEs:

$$\Delta\mu_3^{(k)} \propto \frac{1}{\pi_3^{(k)}} \delta_{2,a}^{(k)}. \quad (9)$$

826 They depend on the higher-level volatility PE $\delta_{2,a}$:

$$\delta_{2,a}^{(k)} = \frac{\hat{\pi}_{2,a}^{(k)}}{\pi_{2,a}^{(k)}} + (\pi_{2,a}^{(k)})^2 \hat{\pi}_{2,a}^{(k)} (\Delta\mu_{2,a}^{(k)})^2 - 1, \quad (10)$$

827 and the higher-level volatility precision π_3 :

$$\pi_{3,a}^{(k)} = \hat{\pi}_{3,a}^{(k)} + \frac{1}{2} (\gamma_{2,a}^{(k)})^2 + (\gamma_{2,a}^{(k)})^2 \delta_{2,a}^{(k)} - \frac{1}{2} \gamma_{2,a}^{(k)} \delta_{2,a}^{(k)}, \quad (11)$$

828 with the precision of the prediction about volatility given by

$$\hat{\pi}_{3,a}^{(k)} = \frac{1}{\frac{1}{\pi_{3,a}^{(k-1)}} + \vartheta_a}. \quad (12)$$

829 The third level, the precision-weighted volatility PE is equivalent to:

$$\varepsilon_{3,a} = \frac{1}{\pi_3^{(k)}} \delta_{2,a}^{(k)}. \quad (13)$$

830

831 The same form of update equations (and precision-weighted PEs) can be derived for
 832 the individual information source, updating beliefs about the rewarded card colour,
 833 i.e.:

$$\varepsilon_{2,c} = \frac{1}{\pi_{2,c}^{(k)}} \delta_{1,c}^{(k)} \quad (14)$$

834 and

$$\varepsilon_{3,c} = \frac{1}{\pi_{3,c}^{(k)}} \delta_{2,c}^{(k)}. \quad (15)$$

835 The prediction errors exhibit a similar form as for the advice, with

$$\delta_{1,c}^{(k)} = u^{(k)} - \hat{\mu}_{1,c}^{(k)} \quad (16)$$

836 for the outcome PE and

$$\delta_{2,c}^{(k)} = \frac{\hat{\pi}_{2,c}^{(k)}}{\pi_{2,c}^{(k)}} + (\pi_{2,c}^{(k)})^2 \hat{\pi}_{2,c}^{(k)} \left(\Delta \mu_{2,c}^{(k)} \right)^2 - 1 \quad (17)$$

837 for the card volatility PE. The individually estimated card colour probability is
 838 equivalent to the logistic sigmoid of the current expectation of the rewarding card
 839 colour:

$$\hat{\mu}_{1,c}^{(k)} = s\left(\mu_{2,c}^{(k-1)}\right) = \frac{1}{1 + \exp\left(-\mu_{2,c}^{(k-1)}\right)}. \quad (18)$$

840

841 In this context, Bayes-optimality is individualized with respect to the values of the
 842 learning parameters, which were allowed to differ across participants.

843 **Arbitration Signal**

844 Within this computational framework, we defined arbitration as the relative perceived
 845 precision associated with each information source, which is equivalent to the
 846 precision of the prediction of each information channel (advice or card; i.e., $\hat{\pi}$)
 847 divided by the total precision. Arbitration is consistent with Bayes' rule representing
 848 the optimal integration of the two inferred states by their precisions.

849 Arbitration towards advice – i.e., the perceived reliability of the social information
 850 source is equivalent to:

$$\xi_{i,a}^{(k)} = \frac{\zeta \hat{\pi}_{i,a}^{(k)}}{\zeta \hat{\pi}_{i,a}^{(k)} + \hat{\pi}_{i,c}^{(k)}} \quad (19)$$

851 on each trial k at each level of the hierarchy i with ζ as the social bias or the additional
 852 bias towards the advice.

853 At the first level and at $i = 1$, the participant relies preferentially on the social input
 854 during action selection when $\xi_{1,a}^{(k)}$ exceeds 0.5. Conversely, when $\xi_{1,a}^{(k)}$ is below 0.5 (see
 855 Eq. 9), the participant relies more on individual (estimates of) card colour
 856 probabilities:

$$\xi_{1,c}^{(k)} = \frac{\hat{\pi}_{1,c}^{(k)}}{\zeta \hat{\pi}_{1,a}^{(k)} + \hat{\pi}_{1,c}^{(k)}} = 1 - \xi_{1,a}^{(k)} \quad (20)$$

857

858 **Response Model**

859 To map beliefs to decisions, we assumed that the prediction of card colour on a given
 860 trial k is a function of arbitration and of the predictions afforded by each source (see
 861 Eq. 21). The response model predicts two components of the behavioural response: (i)
 862 the participant's decision to accept or reject the advice and (ii) the number of points
 863 wagered on every trial. Responses were coded as $y = 1$ when participants took the
 864 advice and chose the card colour indicated by the advisor, and $y = 0$ when
 865 participants decided against following the advice and chose the opposite card colour.
 866 The expected outcome probability is thus a precision-weighted sum of the two
 867 information sources, the estimates of advice accuracy and rewarding colour
 868 probability.

$$\mu_{1,b}^{(k)} = \xi_{i,a}^{(k)} \cdot \hat{\mu}_{1,a}^{(k)} + \xi_{1,c}^{(k)} \cdot \hat{\mu}_{1,c}^{(k)} \quad (21)$$

869 where $\xi_{i,a}^{(k)}$ and $\xi_{1,c}^{(k)}$ are the arbitration for each information source; $\hat{\mu}_{1,a}^{(k)}$ is the
 870 expected advice accuracy (Eq. 4) and $\hat{\mu}_{1,c}^{(k)}$ is the transformed expected card colour
 871 probability from the perspective of the advice (i.e., the estimated card colour
 872 probability indicated by the advisor).

873 It follows from Eqn 21, that social weighting is represented by the first term of this
 874 integrated sum – i.e, $\xi_{i,a}^{(k)} \cdot \hat{\mu}_{1,a}^{(k)}$ whereas card colour weighting is represented by the
 875 second term or $\xi_{1,c}^{(k)} \cdot \hat{\mu}_{1,c}^{(k)}$.

876 The probability that participants chose a particular card colour according to their
 877 expectations about the outcome (Equation 21) was modelled by a softmax function:

$$p(y_{choice} = 1 | \hat{\mu}_{1,b}^{(k)}) = \frac{\hat{\mu}_{1,b}^{(k)\beta_{choice}}}{\hat{\mu}_{1,b}^{(k)\beta_{choice}} + (1 - \hat{\mu}_{1,b}^{(k)})^{\beta_{choice}}} \quad (22)$$

878 where $\beta_{choice} > 0$ is the participant-specific inverse decision temperature parameter. A
 879 low decision temperature (high β_{choice}) means always choosing the highest probability
 880 colour, whereas a high decision temperature (low β_{choice}) means sampling randomly
 881 from a uniform distribution.

882 The number of points wagered provided us with a behavioural readout of decision
 883 confidence. We aimed to formally explain trial-wise wager responses as a linear
 884 function of various sources of uncertainty and precision associated with the lottery
 885 outcome prediction: (i) irreducible decision uncertainty or $\hat{\sigma}_b^{(k)}$ about the outcome,
 886 (ii) arbitration, (iii) informational uncertainty about the card colour or the advice, and
 887 (iv) environmental uncertainty/volatility about the card colour or the advice. We
 888 transformed these computational quantities down to the first level in the hierarchy
 889 using the sigmoid transformation and used them to predict the trial-by-trial wager
 890 (Figure 5 for the group average of each of these quantities):

$$\log(y_{wager}) = \beta_0 + \beta_1 \hat{\sigma}_b^{(k)} + \beta_2 \xi_1^{(k)} + \beta_3 I_{2,a}^{(k)} + \beta_4 I_{2,c}^{(k)} + \beta_5 V_{3,a}^{(k)} + \beta_6 V_{3,c}^{(k)} \quad (23)$$

891

892

893 with

$$\hat{\sigma}_b^{(k)} = \hat{\mu}_{1,b}^{(k)} (1 - \hat{\mu}_{1,b}^{(k)}). \quad (24)$$

894 Parameter ζ captures the social bias in arbitration (equation 19) and $I_{2,a}^{(k)}$ is the
 895 informational uncertainty about the advisor fidelity

$$I_{2,a}^{(k)} = \hat{\mu}_{1,a}^{(k)} (1 - \hat{\mu}_{1,a}^{(k)}) \hat{\sigma}_{2,a}^{(k)} \quad (25)$$

896 where $\hat{\sigma}_{2,a}^{(k)}$ is the inverse of $\hat{\pi}_{2,a}^{(k)}$ and represents the informational uncertainty of the
 897 prediction about the advisor's fidelity (Eq. 6).

898 The environmental volatility is defined as:

$$V_{3,a}^{(k)} = \hat{\mu}_{1,a}^{(k)} (1 - \hat{\mu}_{1,a}^{(k)}) \exp(\mu_{3,a}^{(k-1)}). \quad (26)$$

899

900 Equivalent equations can be derived for the individual information source.

901

902 The trial-wise wager amount predicted by the model is then defined as:

$$\hat{y}_{wager} \stackrel{\text{def}}{=} \log(y_{wager}) + \sqrt{\beta_{wager}} \quad (28)$$

903

904 where β_{wager} is a stochasticity parameter associated with the wager amount. For the
 905 priors of all β parameters estimated here, please refer to Table 1.

906 Competing Models

907 To contrast competing mechanisms underlying learning and arbitration, our model
 908 space consisted of a total of 9 models (*Figure 3*). On the one hand, we included non-
 909 normative perceptual models varying in the degree of volatility processing (3-level full
 910 HGF vs. 2-level no-volatility HGF) and normative perceptual models assuming optimal
 911 Bayesian inference (normative HGF). On the other hand, we included response
 912 models varying in the level of arbitration (arbitration; no arbitration: advice only; no
 913 arbitration: card information only).

914 We considered three families of perceptual models. The first family included the full,
 915 three-level version of the HGF (as described above). By contrast, the second family
 916 lacked the third level, and assumed that agents do not estimate the volatility of the
 917 card probabilities or the advice. Thus, comparing families with and without volatility
 918 tested whether volatility mattered for arbitrated behaviour. Finally, the third family
 919 assumed a Bayes-optimal, normative process of learning from the advice and card
 920 outcomes.

921 In terms of response models, we also considered three families, capturing different
 922 ways in which participants may arbitrate between social and individual sources of

923 information to make decisions. These included: (i) an “Arbitrated” model, which
924 assumed that participants combine and arbitrate between the two information
925 sources, possibly unequally, (ii) an “Advice only” model, assuming arbitration-free
926 reliance on social information only, and (iv) a “Card only” model, representing
927 arbitration-free reliance on the inferred card colour probabilities only (*Figure 3a*).

928 All models were compared formally using Bayesian model selection (BMS; (Stephan et
929 al., 2009). Random effects BMS results in a posterior probability for each model given
930 the participants’ data. The relative goodness of models is denoted by the “protected
931 exceedance probability” reflecting how likely it is that a given model has a higher
932 posterior probability than any other model in the set of models considered (Stephan et
933 al., 2009; Rigoux et al., 2014).

934 We adopted a similar set of priors over the perceptual model parameters as in our
935 previous studies (Diaconescu et al., 2014) (see Table 1). Maximum-a-posteriori (MAP)
936 estimates of model parameters were obtained using the HGF toolbox version 3.0,
937 freely available as part of the open source software package TAPAS at
938 <http://www.translationalneuromodeling.org/tapas>.

939 **FMRI Data Analysis**

940 ***Single-subject Level***

941 Our fMRI data analysis focused on the neural mechanisms of arbitration. Specifically,
942 we conducted two types of analyses on the pre-processed fMRI data:

943 First, we performed a model-based fMRI analysis, in which we constructed a general
944 linear model (GLM), which sought to explain the high-pass filtered voxel time-series
945 with several parametric modulators. The parametric modulators are listed below and
946 were derived from the winning model (i.e., arbitrated three-level version of the HGF,
947 which had the highest posterior probability at the group level). The GLMs were
948 individualized, as the regressors were obtained from fitting the model to the
949 behavioural data of each of the 38 participants. We individualized GLMs because
950 participants differed in how much they relied on each information source and in the
951 extent to which volatility influenced their trial-by-trial wagers (*Figures 4-5*). To
952 investigate the unique contribution of each parametric modulator, we did not
953 orthogonalise them (see *Figure 1 - figure supplement 2* for correlations between them).
954 Moreover, we also included movement and the physiological noise regressors obtained
955 from the PhysIO toolbox (Kasper et al., 2017) based on ECG and respiration recordings
956 as regressors of no interest.

957 In addition to arbitration at the time of advice presentation, we modelled the wager
958 and the outcome phases to examine the effects of hierarchical precision-weighted PEs,
959 and thus test the validity of the computational model and the reproducibility of

960 previous findings, see *Figure 6 – figure supplement 1-2* (Iglesias et al., 2013; Diaconescu
961 et al., 2017). Specifically, the following regressors were included in the GLM:

- 962 1. **Social information** – time when the advice was presented (regressor
963 duration two seconds);
- 964 2. **Arbitration** – parametric modulator of (1), using the trial-specific
965 arbitration quantity (Eq. 19-20);
- 966 3. **Social Weighting** – parametric modulator of (1), using the precision-
967 weighted prediction of the advisor fidelity (first term of Eq. 21);
- 968 4. **Non-social Weighting** – parametric modulator of (1), using the
969 precision-weighted prediction of the individual card weighting (second
970 term of Eq. 21);
- 971 5. **Wager presentation** – time when the option to wager was presented
972 (regressor duration zero seconds);
- 973 6. **Wager** - parametric modulator of (3), using the trial-specific amount of
974 points wagered;
- 975 7. **Outcome** – time when the winning card colour was presented (regressor
976 duration zero seconds);
- 977 8. **Advice Precision-weighted PE** – parametric modulator of (5), using the
978 trial-specific precision-weighted PE of advice validity (Eq. 8);
- 979 9. **Outcome Precision-weighted PE** – parametric modulator of (5), using
980 the trial-specific precision-weighted PE arising from comparing actual
981 and predicted card colour (Eq. 14).
- 982 10. **Volatility Advisor Precision-weighted PE** – parametric modulator of
983 (5), using the trial-specific precision-weighted PE of advice volatility (Eq.
984 13);
- 985 11. **Volatility Card Precision-weighted PE** – parametric modulator of (5),
986 using the trial-specific precision-weighted PE of card colour volatility (see
987 Eq. 15).

988 We observed no significant correlations between response times (RTs) and any of the
989 parametric modulators ($|r| < 0.3$, $p > 0.05$) and therefore did not model RT explicitly.
990 The lack of effects on RTs may be due to the temporal structure of our task (*Figure 1*).
991 Specifically, participants responded long after having received individual information
992 (card outcome in previous trial) and social information had fixed duration (video).
993 Therefore, they are likely to have simply conveyed the decision in the response phase
994 but made it at some time during the video or even before.

995 Second, we predicted that arbitration should be sensitive to volatility, and favour one
996 or the other source of information as a function of perceived relative reliability. Based
997 on this hypothesis, we also performed a non-model based, factorial analysis by
998 dividing the 160 trials into four conditions corresponding to those factors (*Figure 10a*).
999 This GLM included for each of the four conditions the time when the advice was

1000 presented (the social information phase) and the trial-wise wager amount as a
1001 parametric modulator. We assumed that the difference between the four conditions
1002 will be expressed in the advice phase, before participants make their predictions.

1003 *Group Level*

1004 Contrast images from the 38 participants entered a random effects group analysis
1005 (Penny and Holmes, 2007). We used F-tests to identify undirected arbitration signals.
1006 Moreover, one-sample t-tests to investigate directed social or individual arbitration
1007 signals and positive or negative BOLD responses for each of the computational
1008 trajectories of interest described above.

1009 Participant gender and age were included as covariates of no interest at the group
1010 level (the findings remained the same without these covariates). To investigate
1011 individual variability in the representation of social arbitration as a function of
1012 reliance on advice, we used parameter ζ to perform a median split of the group of
1013 participants.

1014 For all analyses, we report results that survived whole-brain family-wise error (FWE)
1015 correction at the cluster level at $p < 0.05$, under a cluster-defining threshold of $p <$
1016 0.001 at the voxel level using Gaussian random field theory (Worsley et al., 1996).
1017 Given recent debate regarding the vulnerabilities of cluster-level FWE procedures
1018 (Eklund et al., 2016), it is worth emphasising that this cluster-defining threshold
1019 ensures adequate control of cluster-level FWE rates in SPM (Flandin and Friston,
1020 2016). The coordinates of all brain regions were expressed in Montreal Neurological
1021 Institute (MNI) space.

1022 Based on recent results that precisions at different levels of a computational hierarchy
1023 may be encoded by distinct neuromodulatory systems (Payzan-LeNestour et al., 2013;
1024 Schwartenbeck et al., 2015), we also performed ROI analyses based on anatomical
1025 masks. We included (i) the dopaminergic midbrain nuclei substantia nigra (SN) and
1026 ventral tegmental area (VTA) using an anatomical atlas based on magnetization
1027 transfer weighted structural MR images (Bunzeck and Düz el, 2006), (ii) the
1028 cholinergic nuclei in the basal forebrain and the tegmentum of the brainstem using
1029 the anatomical toolbox in SPM12 with anatomical landmarks from the literature
1030 (Naidich and Duvernoy, 2009) and (iii) the noradrenergic locus coeruleus based on a
1031 probabilistic map (Keren et al., 2009) (see *Figure S8* for this neuromodulatory ROI).

1032

1033

1034

1035 **References:**

- 1036 Andreou C, Kelm L, Bierbrodt J, Braun V, Lipp M, Yassari AH, Moritz S (2015) Factors
1037 contributing to social cognition impairment in borderline personality disorder
1038 and schizophrenia. *Psychiatry Res* 229:872–879.
- 1039 Apps MAJ, Lockwood PL, Balsters JH (2013) The role of the midcingulate cortex in
1040 monitoring others' decisions. *Front Neurosci* 7 Available at:
1041 <http://journal.frontiersin.org/article/10.3389/fnins.2013.00251/abstract>
1042 [Accessed October 30, 2015].
- 1043 Apps MAJ, Rushworth MFS, Chang SWC (2016) The Anterior Cingulate Gyrus and
1044 Social Cognition: Tracking the Motivation of Others. *Neuron* 90:692–707.
- 1045 Ashburner J, Friston KJ (2005) Unified segmentation. *NeuroImage* 26:839–851.
- 1046 Baker CL, Saxe RR, Tenenbaum JB (2011) Bayesian theory of mind: Modeling joint
1047 belief-desire attribution. In: *Proceedings of the thirty-second annual*
1048 *conference of the cognitive science society*, pp 2469–2474.
- 1049 Behrens TEJ, Hunt LT, Woolrich MW, Rushworth MFS (2008) Associative learning of
1050 social value. *Nature* 456:245–U45.
- 1051 Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS (2007) Learning the value of
1052 information in an uncertain world. *Nat Neurosci* 10:1214–1221.
- 1053 Biele G, Rieskamp J, Gonzalez R (2009) Computational models for the combination of
1054 advice and individual learning. *Cogn Sci* 33:206–242.
- 1055 Braams BR, Güroğlu B, de Water E, Meuwese R, Koolschijn PC, Peper JS, Crone EA
1056 (2014) Reward-related neural responses are dependent on the beneficiary. *Soc*
1057 *Cogn Affect Neurosci* 9:1030–1037.
- 1058 Bunzeck N, Düz el E (2006) Absolute Coding of Stimulus Novelty in the Human
1059 Substantia Nigra/VTA. *Neuron* 51:369–379.
- 1060 Burke CJ, Soutschek A, Weber S, Raja Beharelle A, Fehr E, Haker H, Tobler PN (2018)
1061 Dopamine Receptor-Specific Contributions to the Computation of Value.
1062 *Neuropsychopharmacology* 43:1415–1424.
- 1063 Campbell-Meiklejohn D, Bach D, Roepstorff A, Dolan R, Frith C (2010) How the
1064 opinion of others affects our valuation of objects. *Curr Biol* 20:1165–1170.
- 1065 Carrington SJ, Bailey AJ (2009) Are there theory of mind regions in the brain? A review
1066 of the neuroimaging literature. *Hum Brain Mapp* 30:2313–2335.
- 1067 Charness G, Karni E, Levin D (2013) Ambiguity attitudes and social interactions: An
1068 experimental investigation. *J Risk Uncertain* 46:1–25.

Arbitrating between social and individual learning systems

- 1069 Collins AGE, Frank MJ (2016) Neural signature of hierarchically structured
1070 expectations predicts clustering and transfer of rule sets in reinforcement
1071 learning. *Cognition* 152:160–169.
- 1072 de Berker AO, Rutledge RB, Mathys C, Marshall L, Cross GF, Dolan RJ, Bestmann S
1073 (2016) Computations of uncertainty mediate acute stress responses in humans.
1074 *Nat Commun* 7:10996.
- 1075 Delgado MR, Frank RH, Phelps EA (2005a) Perceptions of moral character modulate
1076 the neural systems of reward during the trust game. *Nat Neurosci* 8:1611–1618.
- 1077 Delgado MR, Frank RH, Phelps EA (2005b) Perceptions of moral character modulate
1078 the neural systems of reward during the trust game. *Nat Neurosci* 8:1611–1618.
- 1079 Deuse L, Rademacher LM, Winkler L, Schultz R, Gründer G, Lammertz SE (2016)
1080 Neural correlates of naturalistic social cognition: Brain-behavior relationships
1081 in healthy adults. *Soc Cogn Affect Neurosci*:nsw094.
- 1082 Devaine M, Hollard G, Daunizeau J (2014a) The Social Bayesian Brain: Does
1083 Mentalizing Make a Difference When We Learn? *PLoS Comput Biol*
1084 10:e1003992.
- 1085 Devaine M, Hollard G, Daunizeau J (2014b) Theory of Mind: Did Evolution Fool Us?
1086 Zalla T, ed. *PLoS ONE* 9:e87619.
- 1087 Diaconescu AO, Mathys C, Weber LAE, Daunizeau J, Kasper L, Lomakina EI, Fehr E,
1088 Stephan KE (2014) Inferring on the Intentions of Others by Hierarchical
1089 Bayesian Learning. *PLoS Comput Biol* 10:e1003810.
- 1090 Diaconescu AO, Mathys C, Weber LAE, Kasper L, Mauer J, Stephan KE (2017)
1091 Hierarchical prediction errors in midbrain and septum during social learning.
1092 *Soc Cogn Affect Neurosci* 12:618–634.
- 1093 Dreher J-C, Kohn P, Berman KF (2006) Neural Coding of Distinct Statistical Properties
1094 of Reward Information in Humans. *Cereb Cortex* 16:561–573.
- 1095 Dunbar RIM (2009) The social brain hypothesis and its implications for social
1096 evolution. *Ann Hum Biol* 36:562–572.
- 1097 Eklund A, Nichols T, Knutsson H (2016) Can parametric statistical methods be trusted
1098 for fMRI based group studies? *Proc Natl Acad Sci* 113:7900–7905.
- 1099 Fareri DS, Chang LJ, Delgado MR (2015) Computational Substrates of Social Value in
1100 Interpersonal Collaboration. *J Neurosci* 35:8170–8180.
- 1101 Flandin G, Friston KJ (2016) Analysis of family-wise error rates in statistical parametric
1102 mapping using random field theory. *ArXiv160608199 Stat* Available at:
1103 <http://arxiv.org/abs/1606.08199> [Accessed September 20, 2016].

Arbitrating between social and individual learning systems

- 1104 Freeman D, Garety P (2014) Advances in understanding and treating persecutory
1105 delusions: a review. *Soc Psychiatry Psychiatr Epidemiol* 49:1179–1189.
- 1106 Friston K, Schwartenbeck P, FitzGerald T, Moutoussis M, Behrens T, Dolan RJ (2014)
1107 The anatomy of choice: dopamine and decision-making. *Philos Trans R Soc*
1108 *Lond B Biol Sci* 369:20130481.
- 1109 Frith C, Frith U (2005) Theory of mind. *Curr Biol* 15:R644–R645.
- 1110 Frith CD (2012) The role of metacognition in human social interactions. *Philos Trans R*
1111 *Soc B Biol Sci* 367:2213–2223.
- 1112 Frith U, Frith C (2010) The social brain: allowing humans to boldly go where no other
1113 species has been. *Philos Trans R Soc B Biol Sci* 365:165–176.
- 1114 Glover GH, Li T-Q, Ress D (2000) Image-based method for retrospective correction of
1115 physiological motion effects in fMRI: RETROICOR. *Magn Reson Med* 44:162–
1116 167.
- 1117 Güroğlu B, Haselager GJT, van Lieshout CFM, Takashima A, Rijpkema M, Fernández G
1118 (2008) Why are friends special? Implementing a social interaction simulation
1119 task to probe the neural correlates of friendship. *NeuroImage* 39:903–910.
- 1120 Hauser TU, Iannaccone R, Ball J, Mathys CD, Brandeis D, Walitza S, Brem S (2014)
1121 Role of the medial prefrontal cortex in impaired decision making in juvenile
1122 attention-deficit/hyperactivity disorder. *JAMA Psychiatry* 71:1165–1173.
- 1123 Henco L, Brandi M-L, Lahnakoski JM, Diaconescu AO, Mathys C, Schilbach L (2020)
1124 Bayesian modelling captures inter-individual differences in social belief
1125 computations in the putamen and insula. *Cortex* Available at:
1126 <http://www.sciencedirect.com/science/article/pii/S0010945220301453> [Accessed
1127 April 27, 2020].
- 1128 Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HEM, Stephan
1129 KE (2013) Hierarchical Prediction Errors in Midbrain and Basal Forebrain
1130 during Sensory Learning. *Neuron* 80:519–530.
- 1131 Kasper L, Bollmann S, Diaconescu AO, Hutton C, Heinzle J, Iglesias S, Hauser TU,
1132 Sebold M, Manjaly Z-M, Pruessmann KP, Stephan KE (2017) The PhysIO
1133 Toolbox for Modeling Physiological Noise in fMRI Data. *J Neurosci Methods*
1134 276:56–72.
- 1135 Keren NI, Lozar CT, Harris KC, Morgan PS, Eckert MA (2009) In vivo mapping of the
1136 human locus coeruleus. *NeuroImage* 47:1261–1267.
- 1137 King-Casas B, Sharp C, Lomax-Bream L, Lohrenz T, Fonagy P, Montague PR (2008)
1138 The rupture and repair of cooperation in borderline personality disorder.
1139 *Science* 321:806–810.

Arbitrating between social and individual learning systems

- 1140 Klucharev V, Hytönen K, Rijpkema M, Smidts A, Fernández G (2009) Reinforcement
1141 learning signal predicts social conformity. *Neuron* 61:140–151.
- 1142 Lebreton M, Abitbol R, Daunizeau J, Pessiglione M (2015) Automatic integration of
1143 confidence in the brain valuation signal. *Nat Neurosci* 18:1159–1167.
- 1144 Lee SW, Shimojo S, O’Doherty JP (2014) Neural Computations Underlying Arbitration
1145 between Model-Based and Model-free Learning. *Neuron* 81:687–699.
- 1146 Lockwood PL (2016) The anatomy of empathy: Vicarious experience and disorders of
1147 social cognition. *Behav Brain Res* 311:255–266.
- 1148 Marshall L, Mathys C, Ruge D, de Berker AO, Dayan P, Stephan KE, Bestmann S (2016)
1149 Pharmacological Fingerprints of Contextual Uncertainty Rushworth MFS, ed.
1150 *PLOS Biol* 14:e1002575.
- 1151 Mathys C, Daunizeau J, Friston KJ, Stephan KE (2011) A Bayesian foundation for
1152 individual learning under uncertainty. *Front Hum Neurosci* 5 Available at:
1153 [://000290218000001](http://dx.doi.org/10.3389/fnhum.2011.00025).
- 1154 Mathys CD, Lomakina EI, Daunizeau J, Iglesias S, Brodersen KH, Friston KJ, Stephan
1155 KE (2014) Uncertainty in perception and the Hierarchical Gaussian Filter. *Front*
1156 *Hum Neurosci* 8 Available at:
1157 [http://www.frontiersin.org/Human_Neuroscience/10.3389/fnhum.2014.00825/a](http://www.frontiersin.org/Human_Neuroscience/10.3389/fnhum.2014.00825/abstract)
1158 [bstract](http://www.frontiersin.org/Human_Neuroscience/10.3389/fnhum.2014.00825/abstract) [Accessed November 22, 2014].
- 1159 Montag C, Dziobek I, Richter IS, Neuhaus K, Lehmann A, Sylla R, Heekeren HR, Heinz
1160 A, Gallinat J (2011) Different aspects of theory of mind in paranoid
1161 schizophrenia: Evidence from a video-based assessment. *Psychiatry Res*
1162 186:203–209.
- 1163 Naidich TP, Duvernoy HM (2009) Duvernoy’s atlas of the human brain stem and
1164 cerebellum high-field MRI: surface anatomy, internal structure, vascularization
1165 and 3D sectional anatomy. Wien; New York: Springer.
- 1166 Nassar MR, Wilson RC, Heasley B, Gold JJ (2010) An Approximately Bayesian Delta-
1167 Rule Model Explains the Dynamics of Belief Updating in a Changing
1168 Environment. *J Neurosci* 30:12366–12378.
- 1169 Ojala KE, Janssen LK, Hashemi MM, Timmer MHM, Geurts DEM, ter Huurne NP,
1170 Cools R, Sescousse G (2018) Dopaminergic Drug Effects on Probability
1171 Weighting during Risky Decision Making. *eNeuro* 5 Available at:
1172 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5889481/> [Accessed August 8,
1173 2019].
- 1174 Payzan-LeNestour E, Dunne S, Bossaerts P, O’Doherty JP (2013) The Neural
1175 Representation of Unexpected Uncertainty during Value-Based Decision
1176 Making. *Neuron* 79:191–201.

Arbitrating between social and individual learning systems

- 1177 Penny WD, Holmes AJ (2007) Chapter 12 - Random Effects Analysis. In: Statistical
1178 Parametric Mapping (Karl Friston, John Ashburner, Stefan Kiebel, Thomas
1179 Nichols and William PennyA2 - Karl Friston JA, William Penny, eds), pp 156–
1180 165. London: Academic Press. Available at:
1181 <http://www.sciencedirect.com/science/article/pii/B9780123725608500127>
1182 [Accessed October 8, 2012].
- 1183 Powell J, Lewis PA, Roberts N, García-Fiñana M, Dunbar RIM (2012) Orbital prefrontal
1184 cortex volume predicts social network size: an imaging study of individual
1185 differences in humans. *Proc R Soc Lond B Biol Sci*:rspb20112574.
- 1186 Powell JL, Lewis PA, Dunbar RIM, García-Fiñana M, Roberts N (2010) Orbital
1187 prefrontal cortex volume correlates with social cognitive competence.
1188 *Neuropsychologia* 48:3554–3562.
- 1189 Preuschoff K, Quartz SR, Bossaerts P (2008) Human Insula Activation Reflects Risk
1190 Prediction Errors As Well As Risk. *J Neurosci* 28:2745–2752.
- 1191 Pulcu E, Browning M (2017) Affective bias as a rational response to the statistics of
1192 rewards and punishments Frank MJ, ed. *eLife* 6:e27879.
- 1193 Rigoux L, Stephan KE, Friston KJ, Daunizeau J (2014) Bayesian model selection for
1194 group studies — Revisited. *NeuroImage* 84:971–985.
- 1195 Rushworth MFS, Behrens TEJ (2008) Choice, uncertainty and value in prefrontal and
1196 cingulate cortex. *Nat Neurosci* 11:389–397.
- 1197 Rushworth MFS, Behrens TEJ, Rudebeck PH, Walton ME (2007) Contrasting roles for
1198 cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends*
1199 *Cogn Sci* 11:168–176.
- 1200 Schaafsma SM, Pfaff DW, Spunt RP, Adolphs R (2015) Deconstructing and
1201 reconstructing theory of mind. *Trends Cogn Sci* 19:65–72.
- 1202 Scheuerecker J, Meisenzahl EM, Koutsouleris N, Roesner M, Schöpf V, Linn J,
1203 Wiesmann M, Brückmann H, Möller H-J, Frodl T (2010) Orbitofrontal volume
1204 reductions during emotion recognition in patients with major depression. *J*
1205 *Psychiatry Neurosci JPN* 35:311–320.
- 1206 Schmitgen MM, Walter H, Drost S, Rückl S, Schnell K (2016) Stimulus-dependent
1207 amygdala involvement in affective theory of mind generation. *NeuroImage*
1208 129:450–459.
- 1209 Schurz M, Radua J, Aichhorn M, Richlan F, Perner J (2014) Fractionating theory of
1210 mind: A meta-analysis of functional brain imaging studies. *Neurosci Biobehav*
1211 *Rev* 42:9–34.

Arbitrating between social and individual learning systems

- 1212 Schwartenbeck P, FitzGerald THB, Mathys C, Dolan R, Friston K (2015) The
1213 Dopaminergic Midbrain Encodes the Expected Certainty about Desired
1214 Outcomes. *Cereb Cortex*:bhu159.
- 1215 Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model
1216 selection for group studies. *NeuroImage* 46:1004–1017.
- 1217 Tobler PN, Christopoulos GI, O’Doherty JP, Dolan RJ, Schultz W (2009) Risk-
1218 dependent reward value signal in human prefrontal cortex. *Proc Natl Acad Sci*
1219 106:7185–7190.
- 1220 Vossel S, Mathys C, Daunizeau J, Bauer M, Driver J, Friston KJ, Stephan KE (2013)
1221 Spatial Attention, Precision, and Bayesian Inference: A Study of Saccadic
1222 Response Speed. *Cereb Cortex* Available at:
1223 <http://cercor.oxfordjournals.org/content/early/2013/01/14/cercor.bhs418>
1224 [Accessed March 30, 2013].
- 1225 Wang JX, Kurth-Nelson Z, Kumaran D, Tirumala D, Soyer H, Leibo JZ, Hassabis D,
1226 Botvinick M (2018) Prefrontal cortex as a meta-reinforcement learning system.
1227 *Nat Neurosci* 21:860.
- 1228 Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC, others (1996) A
1229 unified statistical approach for determining significant signals in images of
1230 cerebral activation. *Hum Brain Mapp* 4:58–73.
- 1231 Zerubavel N, Bearman PS, Weber J, Ochsner KN (2015) Neural mechanisms tracking
1232 popularity in real-world social networks. *Proc Natl Acad Sci* 112:15072–15077.
- 1233 Zink CF, Tong Y, Chen Q, Bassett DS, Stein JL, Meyer-Lindenberg A (2008) Know your
1234 place: neural processing of social hierarchy in humans. *Neuron* 58:273–283.
- 1235
- 1236

1237

Tables

1238

Table 1. Prior mean and variance of the perceptual and response model parameters

Model		Prior mean	Prior variance
Perceptual Models:			
3-level HGF	κ_a, κ_c	0.5	1
	ϑ_a, ϑ_c	0.62	1
Normative HGF	κ_a, κ_c	0.5	0
	ϑ_a, ϑ_c	0.62	0
2-level HGF	ϑ_a, ϑ_c	0.00062	0
Response Models:			
	β_{1-6}	0	4
	β_{ch}	48	1
	β_0	6.21	4
	β_{wager}	1.50	100
1. Arbitrated	ζ	1	25
2. Advice Only	ζ	Inf	0
3. Card Only	ζ	0	0

1239

Note: The prior variances are given in the numeric space in which parameters are estimated. κ , ϑ , and $\mu_3^{(k=0)}$ are estimated in logit-space, while the other parameters are estimated in log-space. Whereas the prior variances for all parameters are set to be rather broad, we selected a shrinkage prior mean and variance for the decision noise parameter β_{ch} such that behaviour is explained more by variations in the remaining parameters rather than decision noise.

1240

1241

1242

1243

1244

1245

1246

1247 **Table 2a.** Results of Bayesian model selection: Model probability ($p(m|y)$) and
 1248 protected exceedance probabilities (ϕ_p). Please refer to the participants' LME and
 1249 BMS results in Table 2-source data files 1 and 2, respectively.
 1250

<i>Perceptual Models:</i>			
Response Models:	Arbitrated	Advice Only	Card Only
<i>3-level HGF</i>			
$p(m y)$	0.63	0.04	0.02
ϕ_p	0.99	4.7e-12	4.7e-12
<i>Normative HGF</i>			
$p(m y)$	0.03	0.03	0.02
ϕ_p	4.7e-12	4.7e-12	4.7e-12
<i>2-level HGF</i>			
$p(m y)$	0.15	0.06	0.02
ϕ_p	6.2e-05	4.7e-12	4.7e-12

1251
 1252
 1253 **Table 2b.** Average maximum a-posteriori estimates of the learning and arbitration
 1254 parameters of the winning model (Arbitrated 3-level HGF). Please refer to participants'
 1255 individual posterior parameter estimates for perceptual and response model
 1256 parameters in Table 2-source data files 3 and 4, respectively.

1257
 1258
 1259

Perceptual Model Parameters	Mean	SD	Response Model Parameters	Mean	SD
κ_c	0.58	0.17	ζ	1.03	1.24
ϑ_c	0.59	0.07	$\beta \xi$	1.42	1.69
κ_a	0.56	0.27	$\beta \hat{\sigma}_1$	-1.59	0.94
ϑ_a	0.62	0.09	$\beta \hat{\sigma}_{2,a}$	0.23	1.37
			$\beta \hat{\sigma}_{2,r}$	0.63	1.24
			$\beta \hat{\mu}_{3,a}$	-2.97	2.47
			$\beta \hat{\mu}_{3,r}$	-0.51	1.83
			β_{ch}	2.25	0.92

1260
 1261

1262
1263
1264

	Hemisphere	x	y	z	# Voxels	F-statistic
$\xi^{(k)}$						
midbrain	L	-6	-18	-12	20	23.49
thalamus	L	-12	-18	8	490	59.87
anterior insula	L	-44	2	0	1744	52.97
anterior insula	R	48	6	-2	813	31.56
fusiform gyrus	R	28	-78	-10	1327	75.32
fusiform gyrus	L	-28	-76	-10	227	39.55
inferior occipital gyrus	R	48	-68	-10	810	52.70
inferior occipital gyrus	L	-42	-68	-4	1519	67.56
calcarine sulcus	R	12	-86	6	22285	199.99
superior temporal gyrus	L	-60	-30	-2	79	24.02
superior temporal sulcus	R	52	-18	-8	104	30.35
amygdala	R	18	-10	-16	76	27.01
precuneus	R	4	-52	30	238	38.50
dorsal medial PFC	L	-10	44	52	108	23.14
superior medial PFC	R	4	56	28	493	39.83
ventrolateral PFC	R	50	36	0	202	24.28
frontopolar cortex	R	4	54	30	138	24.28
orbitofrontal cortex	R	26	34	-10	80	30.47
ventromedial PFC	L	-2	46	-10	393	37.43
supramarginal gyrus	R	54	-30	50	46.46	952
cerebellum	R	18	-48	-18	1919	166.69

1265 **Table 3.** MNI coordinates and F-statistic of maxima of activations induced by either form of
1266 arbitration (Eqs. 19-20; $p < 0.05$, cluster-level whole-brain FWE corrected). Related to *Figure 7*.
1267
1268
1269

	Hemisphere	x	y	z	# Voxels	t-statistic
$\xi_c^{(k)}$: positive correlations						
midbrain	L	-6	-18	-10	95	4.94
thalamus	L	-16	-18	8	232	5.10
	R	22	-30	4	206	5.10
anterior insula	L	-44	2	0	2232	7.28
	R	36	16	8	943	6.23
supplementary motor area/ anterior cingulate sulcus	L	-2	-8	52	1688	6.29
dorsolateral PFC	R	36	46	30	136	5.93
middle occipital gyrus	R	12	-86	6	237	11.70
	L	-32	-82	16	136	8.26
superior occipital gyrus	R	28	-78	30	343	11.00
	L	-26	-82	32	143	8.73
cerebellum	R	18	-48	-18	21557	12.91

1270 **Table 4.** MNI coordinates and t-statistic of maxima of activations induced by arbitration for
1271 the individually estimated card reward probability (Eq. 20; $p < 0.05$, cluster-level whole-brain
1272 corrected). Related to *Figure 8a*.
1273
1274

1275
1276
1277
1278

	Hemisphere	x	y	z	# Voxels	t -statistic
$\xi_a^{(k)}$: positive correlations						
precuneus	R	6	-51	32	284	6.25
amygdala	R	18	-10	-16	107	5.20
anterior cingulate cortex	L	-2	44	-10	136	4.82
ventromedial PFC	R	8	52	14	231	5.72
ventrolateral PFC	R	50	36	0	305	4.93
frontopolar cortex	R	4	62	22	153	4.59
orbitofrontal cortex	R	28	26	-16	126	5.11
middle frontal gyrus	R	38	14	28	305	5.36
superior temporal gyrus	L	-60	-30	-2	107	4.90
superior temporal sulcus	R	52	-18	-8	152	5.51
anterior temporoparietal junction	R	56	-52	24	173	4.18
cerebellum	L	-24	-84	-34	121	4.11

1279 **Table 5.** MNI coordinates and t-statistic of maxima of activations induced by arbitration for
1280 the social advice (Eq. 19; $p < 0.05$, cluster-level whole-brain FWE corrected). Related to *Figure*
1281 *8b*.

1282
1283

	Hemisphere	x	y	z	# Voxels	F-statistic
Stability > Volatility						
supramarginal gyrus	R	46	-28	42	1199	38.16
inferior occipital gyrus	R	46	-66	0	580	33.99
	L	-46	-70	4	256	20.82
anterior insula	R	34	20	2	98	29.30
postcentral gyrus	L	-52	2	34	107	28.97
	R	54	-22	34	129	5.59
precentral gyrus	L	-60	-20	32	512	40.21
	R	50	4	32	129	20.58
middle frontal gyrus	L	-26	0	58	117	20.18

1284 **Table 6.** MNI coordinates and F-statistic for main effects of stability ($p < 0.05$, FWE whole-
1285 brain corrected). Related to *Figure 11* (activations in cyan).

1286
1287
1288

	Hemisphere	x	y	z	# Voxels	F-statistic
Information Source × Task Phase						
midbrain	L	-4	-22	-8	154	48.03
thalamus	L	-12	-24	0	189	116.73
	R	16	-30	2	154	104.27
middle cingulate gyrus	L	-10	16	32	94	37.10
anterior insula	L	-34	-2	10	88	26.71
supplementary motor area/ anterior cingulate sulcus	L	-6	-2	56	736	104.45
dorsolateral PFC	L	-38	52	8	133	22.96
	R	34	34	34	94	21.02

Arbitrating between social and individual learning systems

inferior occipital gyrus	R	44	-66	6	3600	190.83
	L	-40	-76	-12	3300	162.67
superior occipital gyrus	R	28	-78	30	80	23.54
	L	-26	-82	32	81	28.64
orbitofrontal cortex	L	0	48	-22	189	100.84
	R	2	40	-24	180	34.66
ventrolateral prefrontal cortex	L	-46	48	-12	81	37.69
	R	50	44	-8	80	23.53
cerebellum	R	30	-86	-42	95	25.15

1289 **Table 7.** MNI coordinates and F-statistic for interactions between task phases and stimulus
 1290 type ($p < 0.05$, FWE whole-brain corrected). Related to *Figure 11* (activations in magenta).
 1291
 1292

1293 **Supplementary Table Captions**

1294 **Supplementary file 1A.** MNI coordinates and F-statistic of activations induced by precision-
 1295 weighted prediction error about individually estimated card colour probability (equation 14).
 1296 Related to Figure 6 – figure supplement 1a.

1297
 1298 **Supplementary File 1B.** MNI coordinates and F-statistic of activations induced by precision-
 1299 weighted prediction error about advice validity (equation 8). Related to Figure 6 – figure
 1300 supplement 1b.

1301

1302 **Supplementary Figure Captions**

1303 **Figure 1 – figure supplement 1 | Behaviour influenced by volatility.** Average lottery
 1304 prediction accuracy (a), decisions to take the advice (b), and amount of points wagered per
 1305 trial (c) were reduced during volatile phases of the paradigm, particularly with regard to social
 1306 information. The average values across all trials were $68.2\% \pm 6.2\%$ (mean accuracy \pm standard
 1307 deviation) lottery prediction accuracy, $62.1\% \pm 6.9\%$ advice-taking, and 5.6 ± 1.5 points
 1308 wagered (participants on average accumulated 378.6 ± 173.2 points). Jittered raw data (i.e.,
 1309 means over all trials of each behavioural measure per subject) are plotted for each behavioural
 1310 measure. Red lines indicate the mean, grey areas reflect 1 SD of the mean, and coloured areas
 1311 the 95% confidence intervals of the mean. $**p < 0.001$ is indicated to emphasize the phase \times cue
 1312 interactions.

1313 **Figure 2 – figure supplement 1 | Parameter recovery when using empirical parameter**
 1314 **values (Binary HGF).** Parameter recovery for perceptual (a) and response model parameters
 1315 (b). The correlation coefficients (with corresponding p-values) and Cohen's f values are
 1316 included to quantify and compare the parameter recovery results. We saved the seed of the
 1317 random number generator to ensure reproducibility of the results.

1318 **Figure 5 - figure supplement 1 | Model validity with regard to wager amount:** The z -
 1319 transformed wager amount predicted by the model strongly correlated with the z -transformed
 1320 number of points participants actually wagered across all four conditions of the task ((i) $r_1 =$
 1321 0.62 , $p_1 = 3e-05$; (ii) $r_2 = 0.63$, $p_2 = 2e-05$; (iii) $r_3 = 0.81$, $p_4 = 9e-10$; (iv) $r_4 = 0.80$, $p_4 = 1e-09$). The
 1322 regression line is plotted to illustrate the relationship between the actual and predicted
 1323 wagers.

1324 **Figure 1 – figure supplement 2 | Average pairwise correlations between regressors.**
 1325 Using the Fisher-transformation, we computed averages of the pairwise correlations between
 1326 regressors. Overall, the correlations between time periods and between parametric
 1327 modulators were small to moderate, with the exception of the correlation between second-
 1328 and third-level precision-weighted prediction errors about the card colour outcome
 1329 (Epsi2Card with Epsi3Card).

1330 **Figure 7 – figure supplement 1 | Social versus non-social weighting** (equation 21). Whole-
 1331 brain activations by non-social weighting (one's individual predictions about the card colour
 1332 outcome) compared to social weighting were detected in bilateral cerebellum, occipital

Arbitrating between social and individual learning systems

1333 cortices (lingual gyrus, superior occipital cortex), left anterior cingulate sulcus, right
1334 supramarginal gyrus, and left postcentral gyrus (blue). Conversely, activation by social
1335 weighting was significantly larger in the subgenual ACC (green) (whole-brain FWE cluster-
1336 level corrected, $p < 0.05$).

1337 **Figure 6 – figure supplement 1 | Main effects of precision-weighted PEs about card and**
1338 **advice outcomes** (equations 8 and 14). (a) Whole-brain activation by ε_2 : Activations by
1339 unsigned precision-weighted PE about the card probabilities (blue) were detected in the
1340 bilateral inferior/middle occipital gyri, anterior insula, bilateral inferior, medial and middle
1341 frontal gyri, and the bilateral intraparietal sulcus (whole-brain FWE peak- and cluster-level
1342 corrected, $p < 0.05$). Activations by signed precision-weighted PE about the adviser fidelity
1343 (green) were observed in the bilateral fusiform gyrus, lingual gyrus, anterior insula, bilateral
1344 supplementary motor area, left middle temporal cortex, right posterior superior temporal
1345 sulcus, temporal-parietal junction, bilateral dorsolateral and left dorsomedial prefrontal cortex
1346 (whole-brain FWE peak- and cluster-level corrected, $p < 0.05$). (b) Activation of the right VTA
1347 was associated with the unsigned precision-weighted PE about the card probabilities (blue)
1348 and activation of bilateral VTA/SN associated with the signed precision-weighted prediction
1349 error about the adviser fidelity (green). This activation is shown at $p < 0.05$ FWE corrected for
1350 the volume of our anatomical mask comprising dopaminergic nuclei (yellow).

1351 **Figure 6 – figure supplement 2 | Main effects of precision-weighted PEs about card and**
1352 **advice volatility**. (a) Whole-brain activation by ε_3 : Whole-brain activations by signed
1353 precision-weighted volatility PEs about the card probabilities (blue) were detected in the right
1354 superior temporal gyrus, supramarginal gyrus, and posterior insula. Whole-brain activations
1355 by signed precision-weighted volatility PEs about the adviser fidelity (green) were detected in
1356 the right anterior SMA and anterior insula (whole-brain FWE cluster-level corrected, $p < 0.05$).
1357 (b) Whole-brain activation by ε_3 in the PPT/LDT nuclei: Activation of the right cholinergic
1358 PPT/LDT associated with the signed precision-weighted volatility prediction error about the
1359 adviser fidelity is shown at $p < 0.05$ FWE corrected for the volume of our anatomical mask
1360 comprising cholinergic nuclei (yellow).

1361 **Figure 8 – figure supplement 1 | Neuromodulatory nuclei anatomical mask**. The mask
1362 for ROI analyses included (i) the dopaminergic midbrain (substantia nigra, SN, and ventral
1363 tegmental area, VTA), (ii) the cholinergic basal forebrain, (iii) cholinergic nuclei in the
1364 tegmentum of the brainstem, i.e., the pedunculo-pontine tegmental (PPT) and laterodorsal
1365 tegmental (LDT) nuclei, and (iv) the noradrenergic locus coeruleus (LC).

1366

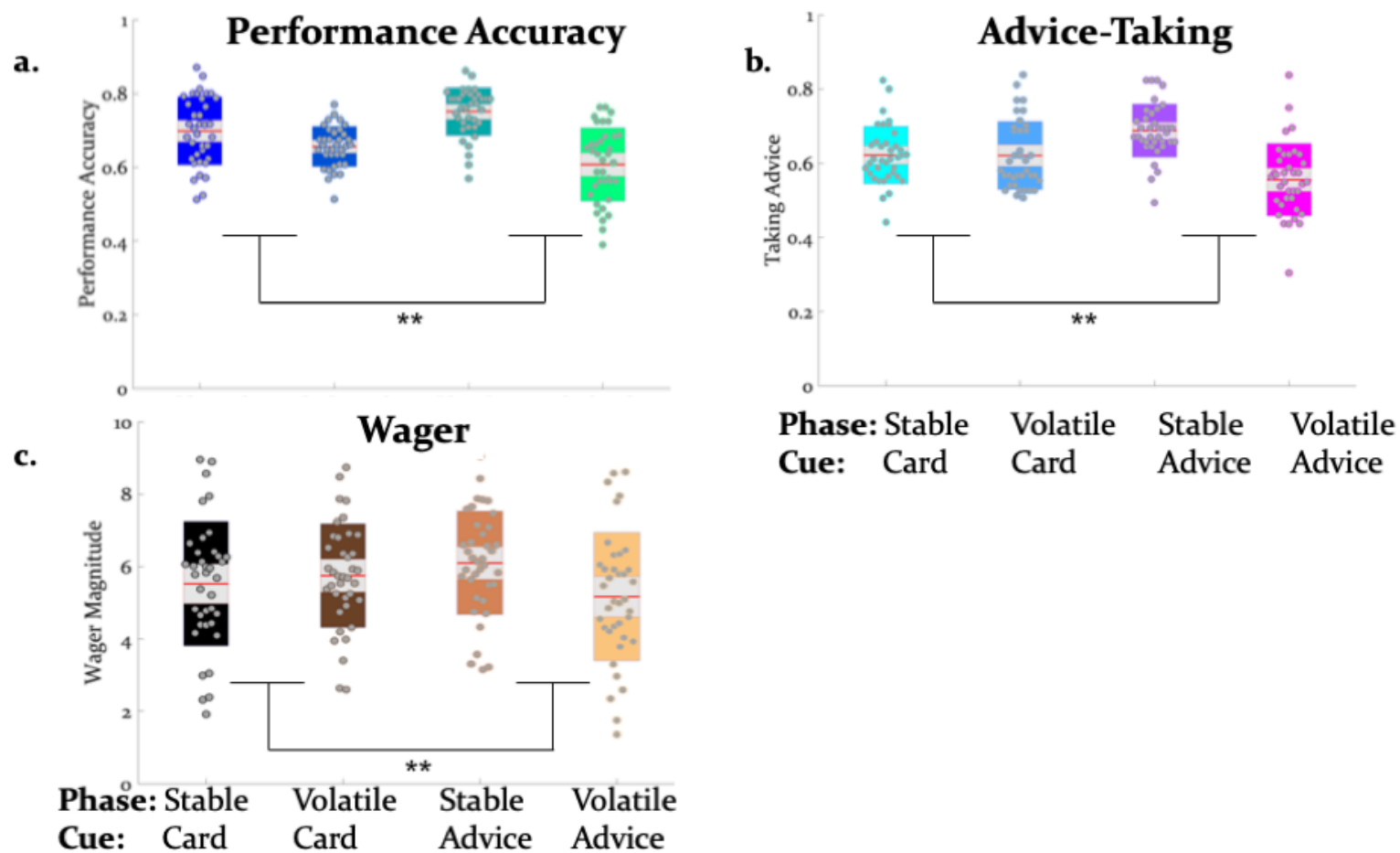


Figure 1 – figure supplement 1 | Behaviour influenced by volatility. Average lottery prediction accuracy (a), decisions to take the advice (b), and amount of points wagered per trial (c) were reduced during volatile phases of the paradigm, particularly with regard to social information. The average values across all trials were $68.2\% \pm 6.2\%$ (mean accuracy \pm standard deviation) lottery prediction accuracy, $62.1\% \pm 6.9\%$ advice-taking, and 5.6 ± 1.5 points wagered (participants on average accumulated 378.6 ± 173.2 points). Jittered raw data (i.e., means over all trials of each behavioural measure per subject) are plotted for each behavioural measure. Red lines indicate the mean, grey areas reflect 1 SD of the mean, and coloured areas the 95% confidence intervals of the mean. ** $p < 0.001$ is indicated to emphasize the phase \times cue interactions.

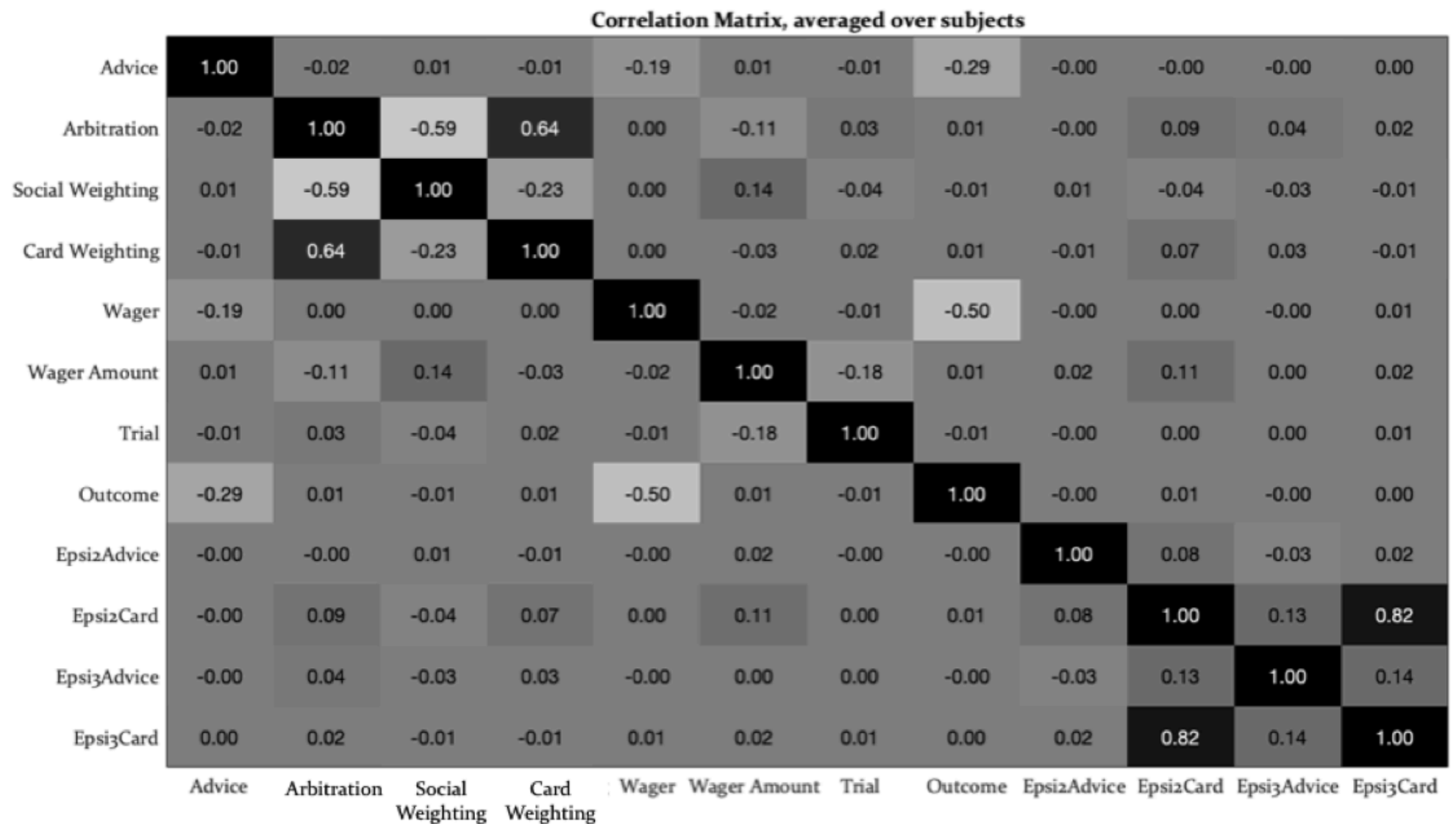


Figure 1 – figure supplement 2 | Average pairwise correlations between regressors. Using the Fisher-transformation, we computed averages of the pairwise correlations between regressors. Overall, the correlations between time periods and between parametric modulators were small to moderate, with the exception of the correlation between second- and third-level precision-weighted prediction errors about the card colour outcome (Epsi2Card with Epsi3Card).

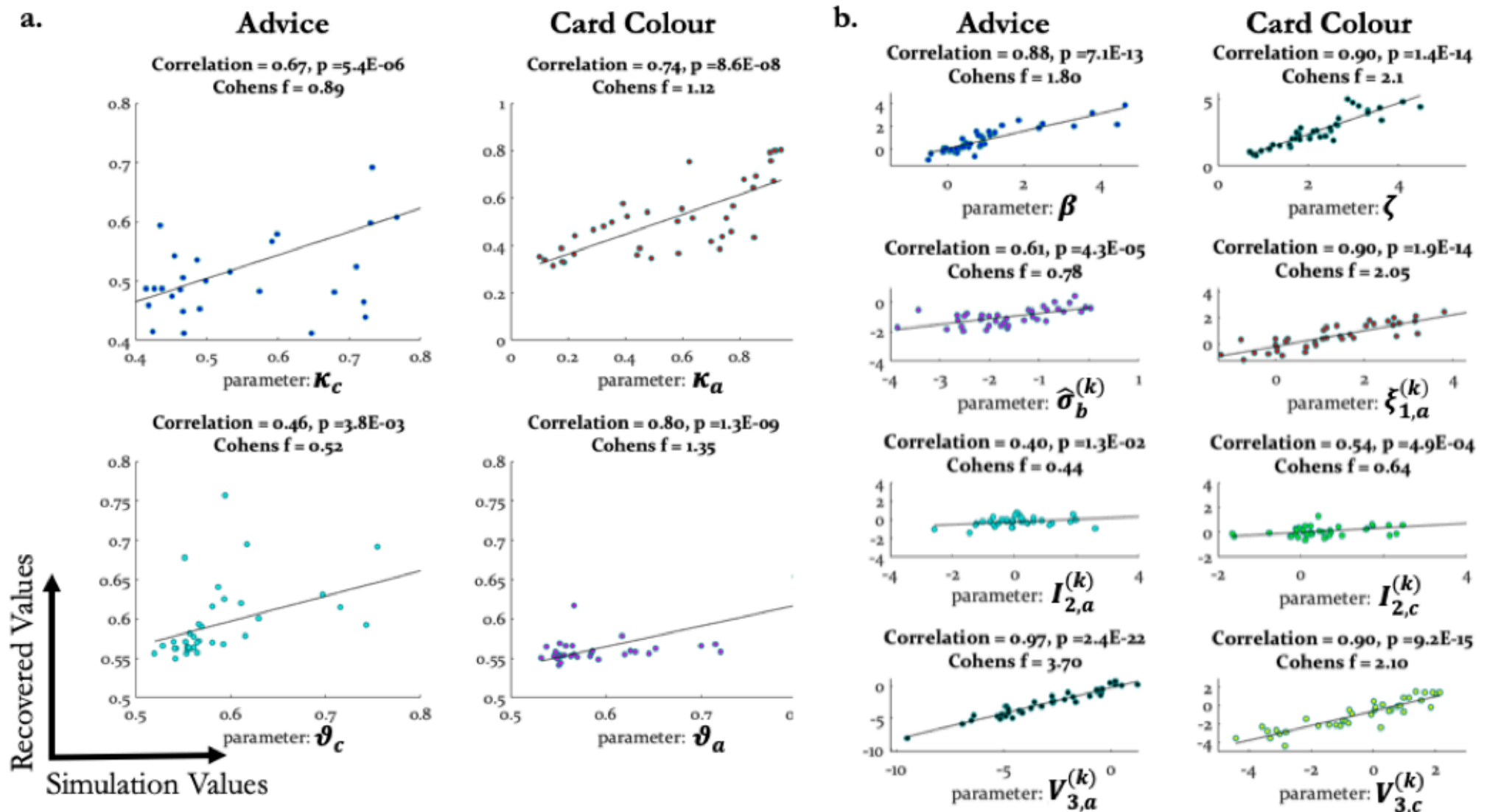


Figure 2 – figure supplement 1 | Parameter recovery when using empirical parameter values (Binary HGF). Parameter recovery for perceptual (a) and response model parameters (b). The correlation coefficients (with corresponding p-values) and Cohen's f values are included to quantify and compare the parameter recovery results. We saved the seed of the random number generator to ensure reproducibility of the results.

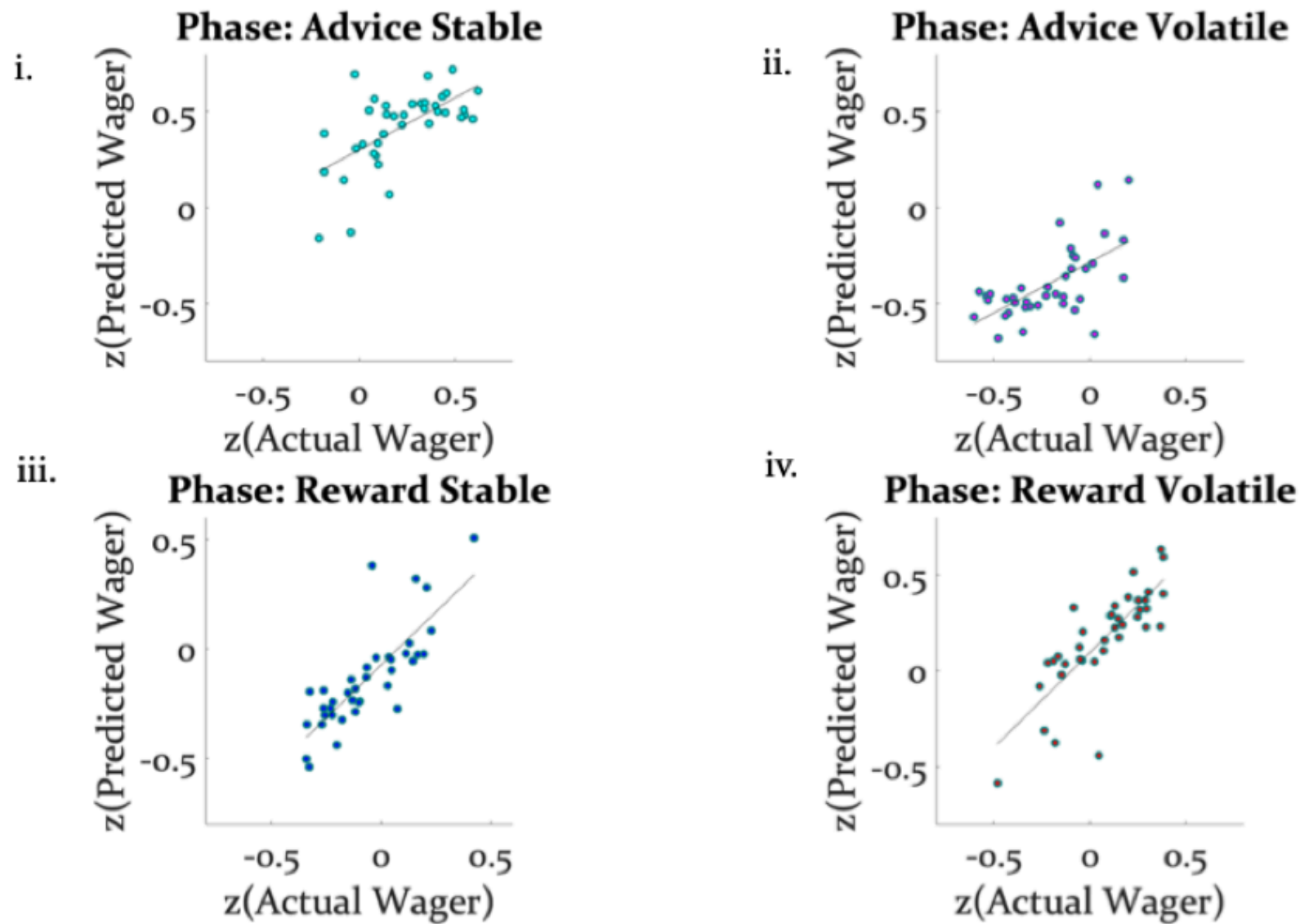


Figure 5 – figure supplement 1 | Model validity with regard to wager amount: The z-transformed wager amount predicted by the model strongly correlated with the z-transformed number of points participants actually wagered across all four conditions of the task ((i) $r_1 = 0.62$, $p_1 = 3e-05$; (ii) $r_2 = 0.63$, $p_2 = 2e-05$; (iii) $r_3 = 0.81$, $p_4 = 9e-10$; (iv) $r_4 = 0.80$, $p_4 = 1e-09$). The regression line is plotted to illustrate the relationship between the actual and predicted wagers.

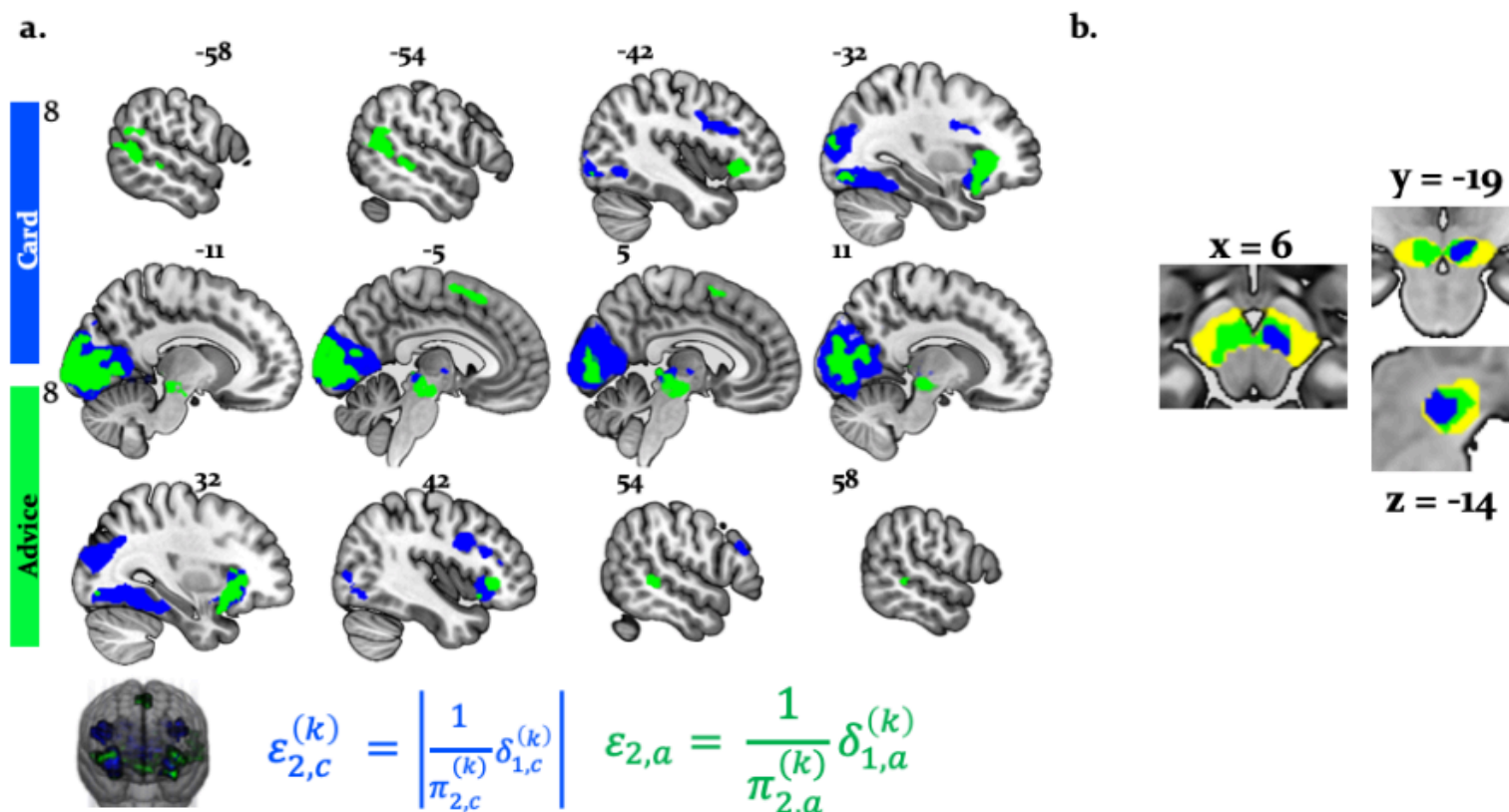


Figure 6 – figure supplement 1 | Main effects of precision-weighted PEs about card and advice outcomes (equations 8 and 14). (a) Whole-brain activation by ε_2 : Activations by unsigned precision-weighted PE about the card probabilities (**blue**) were detected in the bilateral inferior/middle occipital gyri, anterior insula, bilateral inferior, medial and middle frontal gyri, and the bilateral intraparietal sulcus (whole-brain FWE peak- and cluster-level corrected, $p < 0.05$). Activations by signed precision-weighted PE about the adviser fidelity (**green**) were observed in the bilateral fusiform gyrus, lingual gyrus, anterior insula, bilateral supplementary motor area, left middle temporal cortex, right posterior superior temporal sulcus, temporal-parietal junction, bilateral dorsolateral and left dorsomedial prefrontal cortex (whole-brain FWE peak- and cluster-level corrected, $p < 0.05$). (b) Activation of the right VTA was associated with the unsigned precision-weighted PE about the card probabilities (**blue**) and activation of bilateral VTA/SN associated with the signed precision-weighted prediction error about the adviser fidelity (**green**). This activation is shown at $p < 0.05$ FWE corrected for the volume of our anatomical mask comprising dopaminergic nuclei (**yellow**).

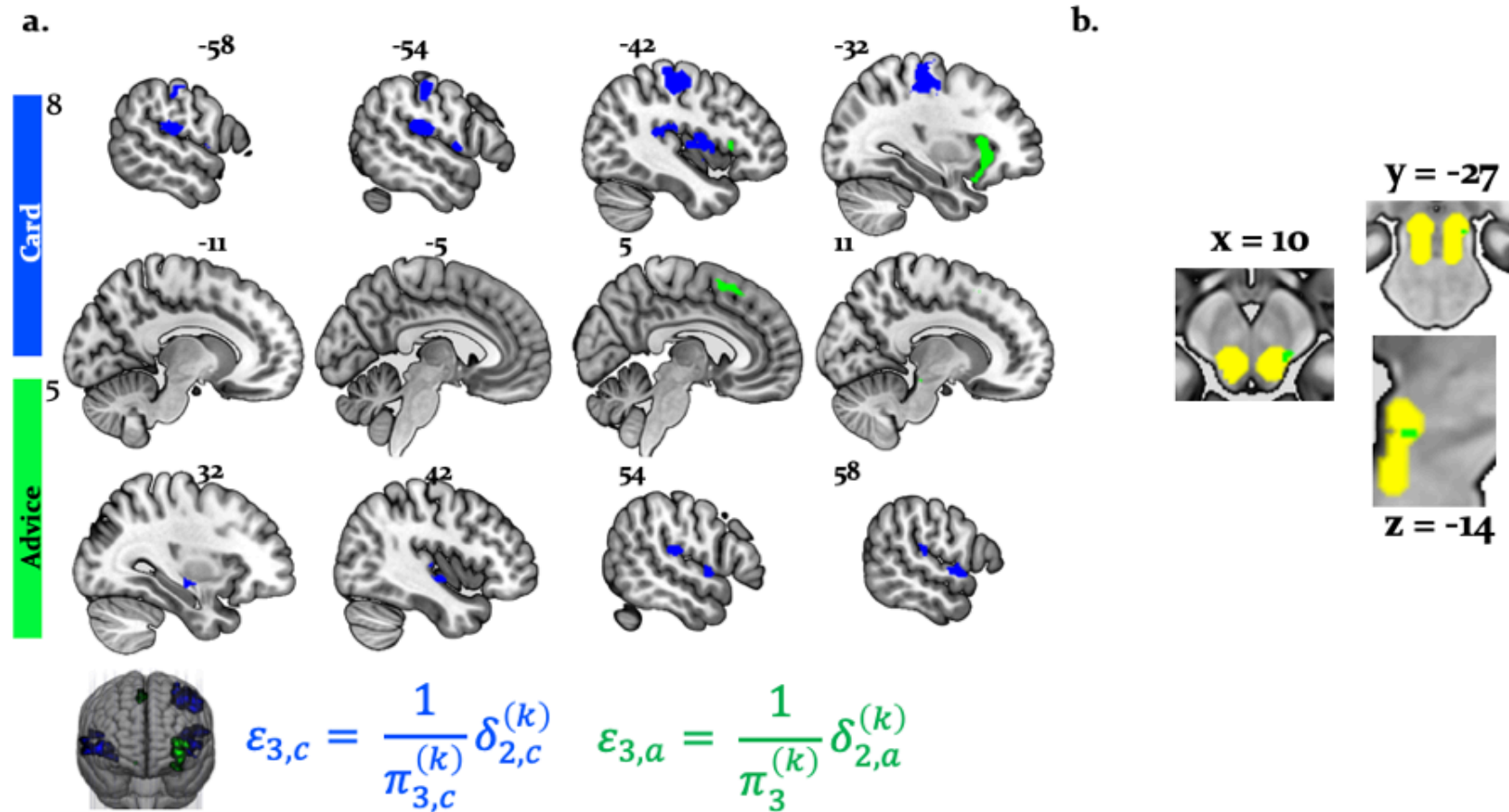


Figure 6 – figure supplement 2 | Main effects of precision-weighted PEs about card and advice volatility. (a) Whole-brain activation by ε_3 : Whole-brain activations by signed precision-weighted volatility PEs about the card probabilities (**blue**) were detected in the right superior temporal gyrus, supramarginal gyrus, and posterior insula. Whole-brain activations by signed precision-weighted volatility PEs about the adviser fidelity (**green**) were detected in the right anterior SMA and anterior insula (whole-brain FWE cluster-level corrected, $p < 0.05$). (b) Whole-brain activation by ε_3 in the PPT/LDT nuclei: Activation of the right cholinergic PPT/LDT associated with the signed precision-weighted volatility prediction error about the adviser fidelity is shown at $p < 0.05$ FWE corrected for the volume of our anatomical mask comprising cholinergic nuclei (**yellow**).

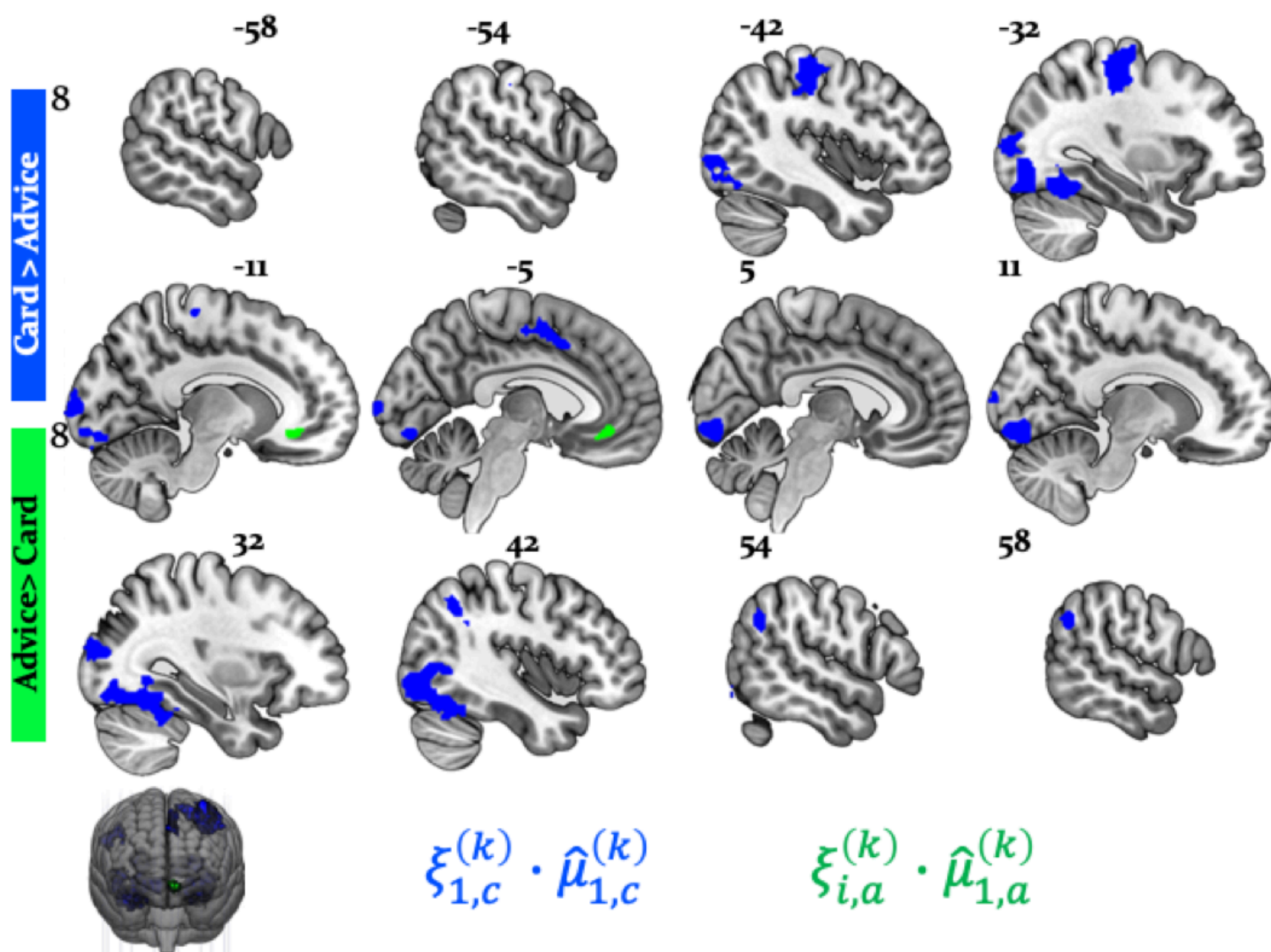
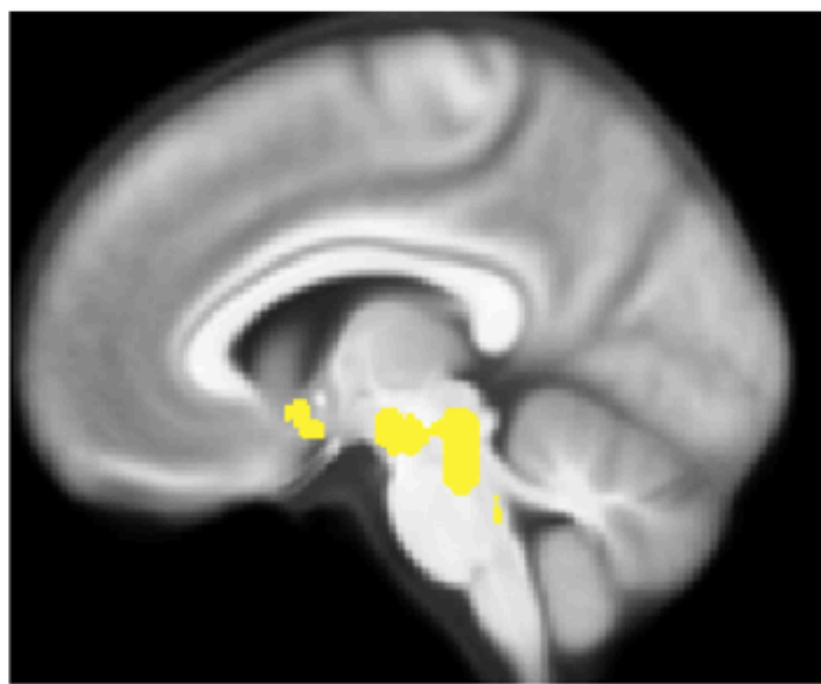
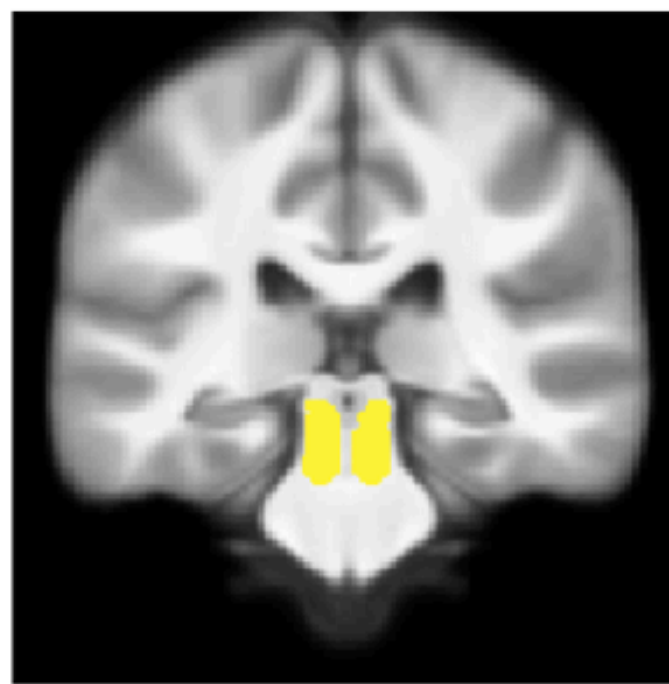


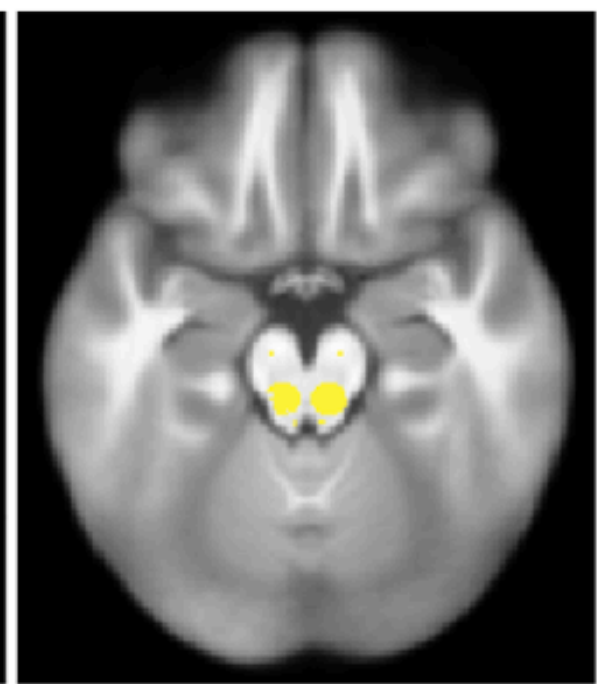
Figure 7 – figure supplement 1 | Social versus non-social weighting (equation 21). Whole-brain activations by non-social weighting (one’s individual predictions about the card colour outcome) compared to social weighting were detected in bilateral cerebellum, occipital cortices (lingual gyrus, superior occipital cortex), left anterior cingulate sulcus, right supramarginal gyrus, and left postcentral gyrus (**blue**). Conversely, activation by social weighting was significantly larger in the subgenual ACC (**green**) (whole-brain FWE cluster-level corrected, $p < 0.05$).



$x = 5$



$y = -16$



$z = -15$


 neuromodulatory system mask

Figure 8 - figure supplement 1 | Neuromodulatory nuclei anatomical mask. The mask for ROI analyses included (i) the dopaminergic midbrain (substantia nigra, SN, and ventral tegmental area, VTA), (ii) the cholinergic basal forebrain, (iii) cholinergic nuclei in the tegmentum of the brainstem, i.e., the pedunculopontine tegmental (PPT) and laterodorsal tegmental (LDT) nuclei, and (iv) the noradrenergic locus coeruleus (LC).