

1 **Metacognition and the effect of incentive motivation in two compulsive**
2 **disorders: gambling disorder and obsessive-compulsive disorder**

3

4 **Authors:**

5 Monja Hoven^{1†}, Nina S. de Boer², Anna E. Goudriaan^{1,3}, Damiaan Denys¹, Mael
6 Lebreton^{4,5,6}, Ruth J. van Holst^{1*}, Judy Luigjes^{1*}

7 †: Corresponding author: m.hoven@amsterdamumc.nl

8 *: Shared last authors

9

10 **Affiliations:**

11 ¹Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam, The
12 Netherlands

13 ²Department of Philosophy, Radboud University, Nijmegen, The Netherlands

14 ³Arkin and Jellinek, Mental Health Care, and Amsterdam Institute for Addiction Research,
15 Amsterdam, The Netherlands

16 ⁴Paris School of Economics, Paris, France

17 ⁵Swiss Center for Affective Science, University of Geneva, Switzerland

18 ⁶Laboratory for Behavioral Neurology and Imaging of Cognition, Department of Fundamental
19 Neurosciences, University of Geneva, Switzerland

20

21 **Nr. of words:** 4852

22 **Keywords:** confidence, metacognition, VMPFC, obsessive-compulsive disorder, gambling
23 disorder, compulsivity

24

25 **Abstract:**

26 Compulsivity is a common phenotype amongst various psychiatric disorders, such as
27 obsessive-compulsive disorder (OCD) and gambling disorder (GD). Deficiencies in
28 metacognition, such as the inability to properly estimate ones' own performance via well-
29 calibrated confidence judgments could contribute to pathological decision-making in these
30 psychiatric disorders. Earlier research has indeed suggested that OCD and GD patients reside
31 at opposite ends of the confidence spectrum, with OCD patients exhibiting underconfidence,
32 and GD patients exhibiting overconfidence. Recently, several studies established that
33 motivational states (e.g. monetary incentives) influence metacognition, with gain (respectively
34 loss) prospects increasing (respectively decreasing) confidence judgments. Here, we
35 reasoned that the OCD and GD symptomatology might correspond to an exacerbation of this
36 interaction between metacognition and motivational states. We hypothesized GD's
37 overconfidence to be exaggerated during gain prospects, while OCD's underconfidence to be
38 worsened in loss context, which we expected to see represented in ventromedial prefrontal
39 cortex (VMPFC) blood-oxygen-level-dependent (BOLD) activity. We tested those hypotheses
40 in a task-based functional magnetic resonance imaging (fMRI) design. Our initial analyses
41 showed increased confidence levels for GD versus OCD patients, that could partly be
42 explained by sex and IQ. Although our primary analyses did not support the hypothesized
43 interaction between incentives and groups, exploratory analyses did show increased
44 confidence in GD patients specifically in gain context. fMRI analyses confirmed a central role
45 for VMPFC in the processing of confidence and incentives, but with no differences between
46 the clinical samples. The trial is registered in the Dutch Trial Register (Trial NL6171,
47 registration number: NTR6318) (<https://www.trialregister.nl/trial/6171>).

48

49

50 **Introduction**

51 Compulsive behaviors are defined as “repetitive acts that are characterized by the feeling
52 that one ‘has to’ perform them while being aware that these acts are not in line with one’s
53 overall goal”¹. Various psychiatric disorders are associated with compulsion, of which
54 obsessive-compulsive disorder (OCD) is the most typical², but it’s also seen in addictive
55 disorders such as gambling disorder (GD)³. Both disorders are characterized by a loss of
56 control over their compulsive behaviors, albeit originating from distinct motivations, serving
57 different purposes and relating to distinct symptomatology^{4,5}. Hence, compulsion seems to
58 be a common phenotype in otherwise symptomatically different disorders.

59 Dysfunctions in metacognition could explain distinct features of compulsive behaviors.
60 Metacognition is the ability to monitor, reflect upon and think about our own behavior⁶. One
61 metacognitive computation is the judgment of confidence, defined as the subjective estimate
62 of the probability of being correct about a choice⁷. Confidence plays a key role in decision-
63 making and learning⁶⁻⁸, and therefore in steering our future behavior^{9,10}. It is crucial for
64 behavioral control that one’s confidence is in line with reality. Nonetheless, discrepancies
65 between actual behavior (e.g. choice accuracy) and confidence in that behavior (subjective
66 estimate of accuracy) have been consistently described, which could contribute to
67 pathological (compulsive) decision-making as seen in various psychiatric disorders¹¹. Clinical
68 presentations of OCD and GD indeed suggest confidence abnormalities in opposite direction,
69 under- and overconfidence, respectively, which could both promote detrimental decision-
70 making, such as checking behavior and compulsive gambling¹²⁻¹⁵. In a recent review we
71 showed that both people with subclinical and clinical OCD consistently showed a decrease in
72 confidence level, which was especially profound in OCD-symptom contexts¹¹. Oppositely, in
73 pathological gamblers, there was evidence for overconfidence in rewarding gambling
74 contexts, which was also related to symptom severity^{16,17}. In sum, GD and OCD patients
75 seem to function at opposite sides of the confidence continuum, respectively over- and

76 under-estimating their performance, which could explain how opposite traits may underlie
77 similar pathological behavior (i.e. compulsive behavior).

78 Reward processes are important for learning and decision-making and interact with
79 cognition¹⁸. Many studies have implicated subcortical regions such as the ventral striatum
80 (VS) and cortical regions such as the ventromedial prefrontal cortex (VMPFC) in reward
81 processing, forming a 'brain valuation system'^{19–21} whose activity relates to value-based
82 decision-making²² and motivates behavior²³. Both OCD and GD patients show deficits in
83 reward processes and accompanying dysregulated neural circuitries. A recent review on
84 neuroimaging of reward mechanisms by Clark et al. (2019) clearly indicated dysregulated
85 reward circuitries, especially focused on the VMPFC and VS in GD, with mixed evidence
86 regarding the direction of these effects²⁴. In OCD, a recent review showed that the ventral
87 affective circuit, consisting of medial frontal cortex and VS was consistently shown to be
88 dysregulated, showing decreased activity in response to rewards, which was increased in
89 response to losses²⁵. This is particularly relevant to the question of how confidence might
90 contribute to those pathologies' symptoms, as an increasing number of studies show that
91 affective and motivational states can influence confidence^{26–28}. Recently, we demonstrated
92 that monetary incentives bias confidence judgments in healthy individuals, where prospects
93 of gain (respectively loss) increase (respectively decrease) confidence, whilst performance
94 levels remained unaffected in both perceptual and reinforcement-learning contexts^{29–32}.

95 We therefore reasoned that an interaction between incentive and confidence processing
96 could cause or fuel the compulsive behaviors in GD and OCD. On the one hand, prospects
97 of high monetary incentives could exaggerate overconfidence in GD patients, leading to
98 continuation of compulsive gambling. On the other hand, in OCD this could lead to
99 exaggerated decreased confidence in negative value context as harm avoidance is
100 considered one of the core motivations of compulsive behavior in OCD^{33–35}.

101 On the neurobiological side, a growing number of functional magnetic resonance imaging
102 (fMRI) studies have associated metacognitive processes with activity in the frontal-parietal
103 network^{36–40}, and activity in the dorsomedial prefrontal cortex (dmPFC), insula and dorsal
104 anterior cingulate cortex (dACC) has been negatively associated to confidence, suggesting a
105 role for these areas in representing uncertainty-related variables^{41–45}. Interestingly, recent
106 studies have also found activity in the VS, the VMPFC and perigenual anterior cingulate
107 cortex (pgACC) - to be positively associated with confidence^{41,46–51}. Importantly, this latter
108 network has been previously positively associated with value-based processes^{20,21,52,53}.
109 Actually, both confidence judgments and value information seem to be automatically
110 integrated into VMPFC's activity^{20,22,47,54,55}. Yet, little is known about if and how the behavioral
111 interaction observed between incentives and confidence can be explained by their shared
112 association with the VMPFC. In an attempt to answer this question, we recently reported an
113 important interaction between incentive and metacognitive signals in the VMPFC in healthy
114 subjects: confidence signals in the VMPFC were observed in trials with gain prospects, but
115 disrupted in trials with no – or negative (loss) monetary prospects³⁰. This suggest that the
116 VMPFC has a key role in mediating the relation between incentives and metacognition.
117 Given the crucial roles of the VMPFC and VS in reward processes and metacognition, which
118 were found to be dysregulated in GD and OCD, we hypothesized that both regions would
119 show disrupted activation patterns related to incentive processing and metacognition and
120 their interaction in patients compared to healthy controls.

121 Overall, in the present study we investigate metacognitive ability and its interaction with
122 incentive motivation in OCD and GD, behaviorally and neurobiologically.

123

124

125 **Methods**

126 *Ethics*

127 Experimental procedures were approved by the Medical Ethics Committee of the Academic

128 Medical Center, University of Amsterdam. All subjects provided written informed consent.

129

130 *Participants*

131 We recruited a total of 31 GD patients, 29 OCD patients and 55 HCs between 18 and 65

132 years old. Of our HC sample of 55 subjects, 25 subjects were included in our earlier work³⁰.

133 HCs were recruited through online advertisements and from our participant database. GD

134 patients were recruited from a local treatment center (Jellinek Addiction Treatment Center

135 Amsterdam) and were recently diagnosed with GD. OCD patients were recruited through the

136 department of psychiatry at the Academic Medical Center in Amsterdam and were diagnosed

137 with OCD.

138

139 *Exclusion criteria*

140 After applying all exclusion criteria (see Supplementary Materials), we included 27 GD

141 patients, 28 OCD patients and 55 HCs for the behavioral analyses, of which four, two and

142 two participants contributed only one of two task sessions, respectively. For the fMRI

143 analyses we included 24 GD patients, 27 OCD patients and 53 HCs, of which seven, three

144 and two participants contributed only one of two task sessions, respectively.

145

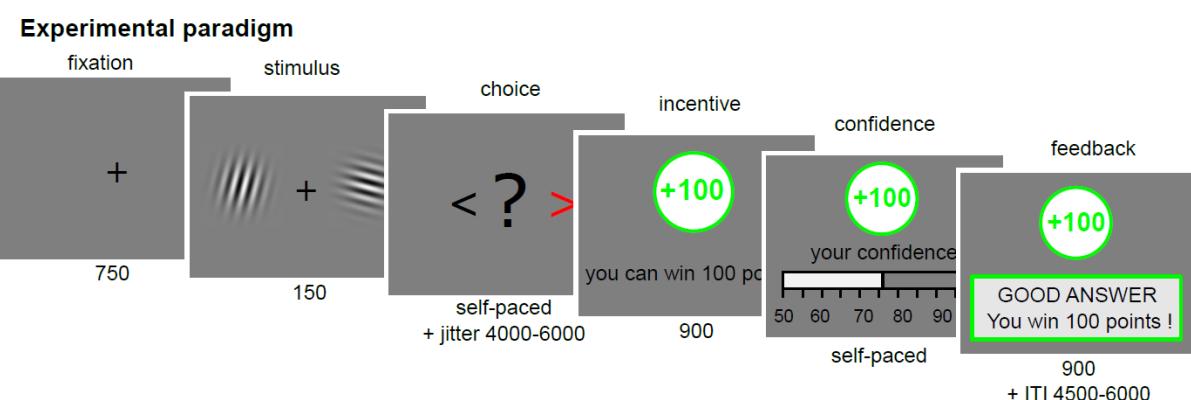
146 *Experimental Design and Study Procedure*

147 We used a similar experimental design and study procedure as previously described³⁰. For

148 details on the experimental design and study procedure, see Hoven et al. (2020) and **Figure**

149 1. In sum, subjects performed a simple perceptual decision-making task, with a 2-alternative
150 forced choice of contrast discrimination followed by a confidence judgment. In each trial,
151 participants could either win (gain context) or lose (loss context) points, or not (neutral
152 context), conditional on the accuracy of the choice in that trial. Importantly, this
153 incentivization was administered after the choice moment, but before the confidence rating.
154 The task was implemented using MATLAB® (MathWorks Inc., Sherborn, MA, USA) and the
155 COGENT toolbox.

156



157

158 **Figure 1 | Experimental paradigm.** Participants viewed two Gabor patches on both sides of the screen
159 (150 ms) and then chose which had the highest contrast (left/right, self-paced) (for more information,
160 see Hoven et al., 2020). After a jitter of a random interval between 4500 to 6000 ms, the incentive was
161 shown (900 ms; green frame for win trials, grey frame for neutral trials, red frame for loss trials).
162 Afterwards, participants were asked to report their confidence in their choice on a rating scale ranging
163 from 50% to 100% with steps of 5%. The initial position of the cursor was randomized between 65%
164 and 85%. Finally, subjects received feedback. The inter trial interval (ITI) had a random duration
165 between 4500 and 6000 ms. The calibration session only consisted of Gabor discrimination, without
166 confidence rating, incentives or feedback and was used to adjust difficulty so that every individual
167 reached a performance of 70%.

168

169 *Behavioral Measures*

170 We extracted trial-by-trial experimental factors: incentive condition, evidence and behavioral
171 measures: accuracy, confidence ratings, reaction times. Evidence was calculated by
172 normalizing the unsigned difference of the two Gabor patches' contrast intensities by their
173 sum to adjust for saturation effects (for more details see³¹). In addition, we computed an
174 extra *latent* variable: early certainty.

175 The early certainty variable was computed in order to analyze BOLD activity at choice
176 moment, when the brain encodes a confidence signal that is not yet biased by incentives.
177 This was done by making a trial-by-trial prediction of early certainty based on stimulus
178 features (reaction times, evidence and accuracy) at choice moment. This resulted in an early
179 certainty signal that was highly correlated with confidence, but showed no statistical
180 relationship with incentives (see Supplementary Materials). For more details, see³⁰.
181 Next to confidence ratings we also assessed additional metacognitive metrics:(1) Confidence
182 calibration, the difference between average confidence and average performance as an
183 indicator of over- or underconfidence, (2) Metacognitive sensitivity, the ability to discriminate
184 between correct answers and errors using confidence judgments (see Supplementary
185 Materials).

186

187 *Behavioral Analyses*

188 All analyses were performed in the R environment (RStudio Team (2015). RStudio:
189 Integrated Development for R. RStudio, Inc., Boston, MA). We used linear mixed effects
190 models (LMMs) as implemented in the lmer function from the lme4 and afex packages^{56,57}.
191 To determine p-values for the fixed effects, we performed Type 3 F tests with Satterthwaite
192 approximation for degrees of freedom as implemented in the afex package. When relevant,
193 we used the ‘emmeans’ package to perform post-hoc tests that were corrected for multiple
194 comparisons using Tukey’s method⁵⁸.

195 To answer our main research questions, we built several LMMs and performed a model
196 selection procedure (see Supplementary Materials). The final model (Model 1) included fixed
197 effects of incentive, group, accuracy and evidence (z-scored) and interactions between
198 incentive and group, as well as two-way and three-way interactions between evidence,
199 accuracy and group. Moreover, a random subject intercept and a random slope of incentives
200 per subject were included in the final model as well. To confirm that the incentive condition or

201 group did not influence accuracy or reaction time, we modelled additional LMEMs with
202 performance and reaction time as dependent variables (Model 2, Model 3).

203 Lastly, we added IQ (z-scored) and sex as fixed effects to our original Model 1 (Model 4) to
204 control for differences in the distribution of these demographic variables. Model fit was
205 assessed and compared using Chi-square tests on log-likelihood values. Additional control
206 analyses on the properties of confidence, early certainty, confidence calibration and
207 metacognitive sensitivity are reported in the Supplementary Materials.

208 Due to a technical bug, our design was not fully balanced as the level of perceptual evidence
209 was not equal across the incentive conditions. ANOVA and post-hoc testing indeed showed
210 that evidence was highest in neutral condition, followed by gain and loss. There were no
211 group differences, nor an interaction between group and incentive. These effects cannot
212 account for any group differences we find in our data, since evidence did not differ between
213 groups. Importantly, the evidence differences did not affect performance, since performance
214 is equal across conditions. See Supplementary Materials for more details.

215

216 *fMRI analyses*

217 For details on fMRI acquisition and preprocessing see Supplementary Materials and Hoven
218 et al (2020)³⁰.

219 All fMRI analyses were conducted using SPM12. Critically, our design allowed us to
220 distinguish between our two timepoints of interest: 1) the moment of stimulus presentation
221 and choice in which implicit (un)certainty about the choice is formed, and 2) the moment of
222 incentive presentation and confidence rating, in which the value of incentives and the
223 confidence rating are encoded. We built a general linear model (GLM 1) estimated on
224 subject-level with these two moments of interest: the moment of choice (i.e. stimulus
225 presentation) and the moment of incentive presentation/confidence rating. We chose to

226 analyze the incentive presentation and confidence rating as a single timepoint since the
227 rating moment followed the presentation of the incentive after 900 ms, with regressors time-
228 locked to the onset of incentive presentation. We also included a regressor for the moment of
229 feedback to explain variance in neural responses related to feedback on accuracy and value
230 that was not related to the decision-making process, but this regressor was not of interest for
231 the current analyses. All whole-brain activation maps were thresholded using family-wise
232 error correction (FWE) at cluster level ($PFWE_{clu} < 0.05$), with a voxel cluster-defining
233 threshold of $p < .001$ uncorrected.

234 Using GLM 1, with regressors for choice modulated by early certainty, for incentive/rating
235 modulated by incentive and confidence, and for feedback modulated by accuracy we were
236 able to investigate our contrasts of interest: (1) choice moment modulated by early certainty,
237 (2) incentive/rating moment modulated by incentive value and (3) incentive/rating moment
238 modulated by confidence rating. For details see Supplementary Materials.

239 In order to study the interaction between incentive motivation and metacognitive ability on the
240 neurobiological level we leveraged the factorial design of our task to build GLM 2. We used
241 GLM 2 to explicate the effect of incentive motivation on both the integration of evidence at
242 choice moment, as well as on confidence formation, and compare those between groups.
243 GLM 2 consisted of regressors for each time point (choice and incentive/rating moments)
244 and for each incentive condition, as well as a single regressor at feedback moment, resulting
245 in seven regressors. For all these events we examined both baseline activity and regression
246 slopes relating to their pmod of interest: signed evidence for choice and confidence for
247 incentive/rating. See Supplementary Materials for more details.

248

249 Since the results by Hoven et al., 2020 suggested that the VMPFC plays an important role in
250 the interaction between incentive motivation and metacognition, we created a functional
251 region of interest (ROI) that represented the confidence-related activity in the VMPFC cluster

252 from our GLM 1 across groups results (see **Figure 4C, Table 5**). We then extracted
253 individual *t*-statistics within this ROI (i.e. normalized beta estimates⁵⁹) from our contrasts of
254 interest and performed one-sample *t*-tests against 0 to check for positive or negative
255 activation patterns. Then, we compared them between incentive conditions, groups, and
256 studied their interactions using mixed ANOVAs implemented in the afex package. When
257 appropriate, we performed post-hoc testing using the emmeans package, correcting for
258 multiple comparisons using Tukey's method. Since we also hypothesized that the VS would
259 play a role in the interaction between incentives and metacognition, we performed the same
260 ROI analysis in the VS with a functional ROI that represented the incentive-related activity in
261 the VS cluster from our GLM 1 across group results (see **Table 5**).

262

Model	Model notation	AIC	BIC	Comparison	χ^2	P-value	Winning model
A	Confidence ~ Incentive * Group + (Incentive Subject)	122919	123041				
B	Confidence ~ Incentive * Group + Accuracy + (Incentive Subject)	122273	122402	A vs. B	648.59	< 2.2e-16	B
C	Confidence ~ Incentive * Group + Accuracy + Evidence + (Incentive Subject)	122004	122141	B vs. C	271.00	< 2.2e-16	C
D	Confidence ~ Incentive * Group + Accuracy*Evidence + (Incentive Subject)	121791	121936	C vs. D	214.53	< 2.2e-16	D
E	Confidence ~ Incentive * Group + Accuracy*Evidence*Group + (Incentive Subject)	121751	121942	D vs. E	52.141	1.747e-09	E
F	Confidence ~ Incentive * Group + Accuracy*Evidence*Group + Sex + IQ + (Incentive Subject)	121752	121958	E vs. F	2.7018	0.259	E

263 **Table 1 | Model descriptions and comparison.** Shown here are the model notations of all models
264 with their respective Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)
265 values, as well as model comparison outcomes with corresponding χ^2 and P-values, resulting in the
266 winning model 'E', which is referred to as Model 1 in the manuscript.

267

268

269 **Results**

270 *Demographics*

271 IQ and sex distributions differed between groups (IQ: $F_{2,107} = 3.222$, $p=0.0438$; sex: $X =$
272 14.483 , $df = 2$, $p<.001$), with higher IQ scores for HC subjects compared with GD patients (t
273 $= 2.53$, $p=0.014$) and with mostly men in the GD group, and relatively more women in the
274 OCD group (Table 2). This corresponds to the natural distribution observed in
275 epidemiological studies for OCD and GD, showing higher prevalence of GD amongst men,
276 and a slightly higher prevalence of OCD in women^{60–63}. Age did not differ between groups.
277 For post-hoc group differences on questionnaire scores, see Supplementary Materials.

278

279

	HC	GD	OCD	Statistics
Age	33.51 +- 12.32	33.22 +- 10.40	31.93 +- 8.21	$F_{2,107} = 0.25, p = 0.777$
IQ*	91.18 +- 10.96	85.22 +- 9.53	89.54 +- 8.32	$F_{2,107} = 3.22, p = 0.0438$
Y-BOCS***	0.25 +- 1.76	1.19 +- 2.60	20.36 +- 6.15	$F_{2,107} = 322.2, p < .001$
PGSI***	0.05 +- 0.40	14.85 +- 4.80	0.64 +- 1.91	$F_{2,107} = 380.5, p < .001$
HAMA***	1.09 +- 1.97	3.93 +- 5.88	11.43 +- 6.28	$F_{2,107} = 48.02, p < .001$
HDRS***	1.31 +- 2.31	5.07 +- 6.24	7.71 +- 4.04	$F_{2,107} = 24.97, p < .001$
Sex (m/f)***	33 / 22	24 / 3	11 / 17	$\chi^2(2) = 14.483, p < .001$

280 **Table 2 | Demographics:** Means +- standard deviations of various demographic variables are shown
281 per group, for sex counts are displayed. Statistics for group comparisons are shown, including F and
282 χ^2 statistics, degrees of freedom and p-values. IQ= estimated Intelligence Quotient, GD = gambling
283 disorder, HAMA = Hamilton Anxiety Rating Scale, HC = healthy control, HDRS = Hamilton Depression
284 Rating Scale, OCD = obsessive-compulsive disorder PGSI = Problem Gamblers Severity Index, Y-
285 BOCS = Yale-Brown Obsessive Compulsive Scale. *p<.05, ***p<.001

286

287

288 *Behavioral Results*

289 To start, we answered our main questions: (1) are there group differences in confidence, and
290 (2) what is the influence of incentive motivation on confidence. Model 1 showed a main effect
291 of group ($F_{2,112} = 4.7910$, $p=.01$) and incentive ($F_{2,112} = 20.9371$, $p<.001$) on confidence
292 (**Figure 2, Supplementary Table 3**). We also found a main effect of accuracy ($F_{1,15107} =$
293 608.8906 , $p<0.001$), with subjects showing higher confidence for correct answers. Moreover,
294 there was a significant two-way interaction of group and evidence ($F_{2,15099} = 3.5094$,
295 $p=0.02994$). As expected, we also found a significant interaction between accuracy and
296 evidence, replicating the 'X-pattern' signature of evidence integration where confidence
297 increases with increasing evidence when correct, and vice versa ($F_{1,15097}=185.3245$,
298 $p<0.001$)⁶⁴. Interestingly, the evidence integration effect differed per group, as signaled by a
299 significant three-way interaction between accuracy, evidence and group ($F_{2,15094} = 3.0533$,
300 $p=0.04723$) (**Supplementary Figure 3, Supplementary Table 3**, for post-hoc tests see
301 Supplementary Materials). Lastly, the interaction between incentive and group revealed a
302 trend towards an effect ($F_{4,112}= 2.2821$, $p=0.06487$).

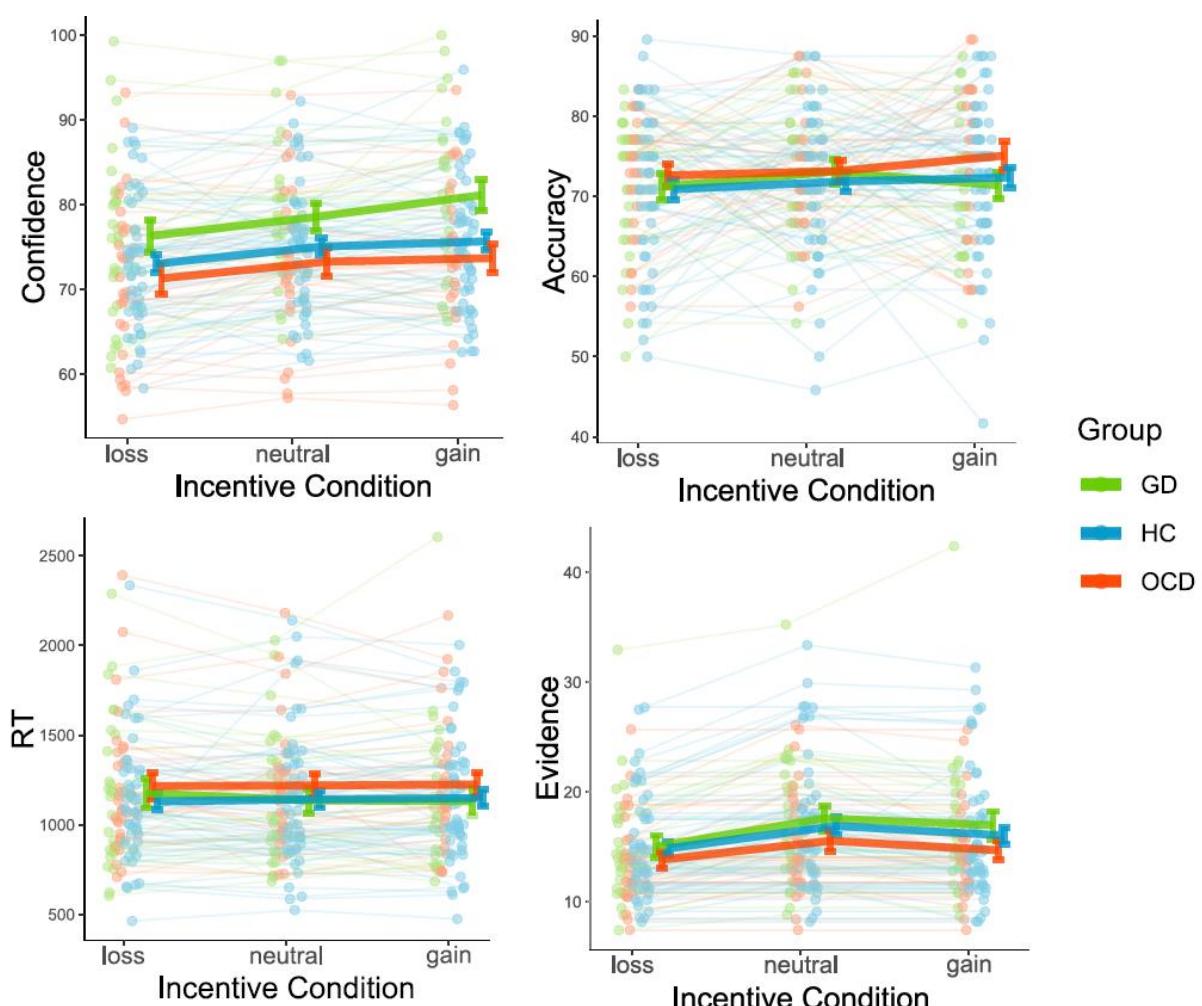
303 Post-hoc tests indicated a significantly higher confidence in GD patients versus OCD patients
304 (GD-OCD = 6.38 ± 2.12 , Z-ratio = 3.014 , $p=0.0073$), and a trend towards higher confidence
305 in GD compared to HC subjects (GD-HC = 4.30 ± 1.84 , Z-ratio = 2.333 , $p=0.0513$), whereas
306 OCD patients did not differ from HC subjects. Moreover, we replicated the parametric effect
307 of incentive value on confidence (loss-neutral = -1.80 ± 0.429 , Z-ratio = -4.192 , $p<0.001$;
308 loss-gain = -3.14 ± 0.486 , Z-ratio = -6.460 , $p<0.001$; neutral-gain = -1.34 ± 0.363 , Z-ratio = $-$
309 3.683 , $p<0.001$). With regards to the three way interaction we found that GD patients'
310 confidence was less influenced by evidence for correct answers compared to both HCs and
311 OCD patients (see Supplementary Materials, **Supplementary Figure 3**). Exploratory post-
312 hoc analyses on the group*incentive interaction effect showed that, especially in context of
313 possible gains, GD patients were more confident than OCD patients (GD - OCD = $8.12 \pm$
314 2.24 , Z-ratio = 3.621 , $p<0.001$) and HC subjects (GD - HC = 5.83 ± 1.95 , Z-ratio = 2.989 ,

315 p=0.0079), with no differences between HC and OCD patients in any incentive condition

316 (**Table 3**).

317 As control analyses we estimated Model 2 and 3 with accuracy and reaction time as
318 dependent variables. (**Table 4**). No effect of group, incentive or an interaction effect on
319 accuracy or reaction time were found, as expected from our design (where incentives follow
320 choices), confirming that accuracy and response times cannot confound any effect of
321 incentives that we found on confidence.

322



323

324 **Figure 2 | Behavioral results.** Individual-averaged confidence, accuracy, reaction times and evidence
325 as a function of incentive condition (loss, neutral and gain) per group. Green dots and lines represent
326 gambling disorder patients, blue dots and lines represent healthy controls and red dots and lines
327 represent obsessive-compulsive disorder patients. Dots represent individuals, and lines highlight
328 within subject variation across conditions. Error bars represent sample mean \pm SEM per group. GD =
329 gambling disorder, HC = healthy control, OCD = obsessive-compulsive disorder

330

331

Model 1	Confidence
Incentive	$F(2.00, 112.34) = 20.94, p < .001$
Group	$F(2.00, 112.51) = 4.79, p = .010$
Accuracy	$F(1.00, 15107.05) = 608.89, p < .001$
Evidence	$F(1.00, 15104.05) = 0.04, p = .848$
Incentive:Group	$F(4.00, 112.10) = 2.28, p = .065$
Accuracy:Evidence	$F(1.00, 15097.33) = 185.32, p < .001$
Group:Accuracy	$F(2.00, 15106.28) = 2.27, p = .103$
Group:Evidence	$F(2.00, 15099.41) = 3.51, p = .030$
Group:Accuracy:Evidence.	$F(2.00, 15094.35) = 3.05, p = .047$
Model 4	Confidence
Incentive	$F(2.00, 112.34) = 20.93, p < .001$
Group	$F(2.00, 112.50) = 2.75, p = .068$
Sex	$F(1.00, 110.26) = 2.88, p = .093$
IQ	$F(1.00, 109.80) = 0.03, p = .865$
Accuracy	$F(1.00, 15107.01) = 609.14, p < .001$
Evidence	$F(1.00, 15104.51) = 0.04, p = .845$
Incentive:Group	$F(4.00, 112.11) = 2.29, p = .064$
Accuracy:Evidence	$F(1.00, 15097.16) = 185.42, p < .001$
Group:Accuracy	$F(2.00, 15106.06) = 2.30, p = .100$
Group:Evidence	$F(2.00, 15098.91) = 3.45, p = .032$
Group:Accuracy:Evidence	$F(2.00, 15094.15) = 3.09, p = .046$

332 **Table 3 | Results of linear mixed-effects models.** Shown here are the results of Model 1 (without
333 demographics) and Model 4 (with demographics) acquired using Type 3 F tests with Satterthwaite
334 approximation for degrees of freedom using the afex package. Shown are F values, with
335 corresponding degrees of freedom and P-values.

336

337

Model 2: Accuracy ~ Incentive*Group + (1+Incentive|Subject)

Group	$F_{2,109} = 0.5827, P = 0.5601$
Incentive	$F_{2,1591} = 1.0319, P = 0.3566$
Group*Incentive	$F_{4,1586} = 0.8671, P = 0.4830$

Model 3: RT ~ Incentive*Group + (1+Incentive|Subject)

Group	$F_{2,110} = 0.5207, P = 0.5956$
Incentive	$F_{2,220} = 0.0994, P = 0.9054$
Group*Incentive	$F_{4,219} = 0.4269, P = 0.7891$

338 **Table 4 | Results of control models.** Shown here are the results of Model 2 and Model 3 linear
339 mixed-effects models, acquired using Type 3 F tests with Satterthwaite approximation for degrees of
340 freedom using the afex package. Shown are F values, with corresponding degrees of freedom and P-
341 values

342

343

344 Since sex and IQ were significantly different between the groups, we aimed to control for
345 these variables by adding them as fixed effects, resulting in Model 4. The main effect of
346 group did not remain significant, but showed a trend towards an effect ($F_{2,112} = 2.7465$,
347 $p=0.06846$), while the main effect of incentive did remain significant ($F_{2,112} = 20.9326$, $p<$
348 0.001). We found no evidence for a significant effect of sex ($F_{1,110} = 2.8776$, $p=0.09264$), or
349 IQ ($F_{1,109} = 0.0291$, $p=0.86489$). The interaction effect between group and incentive
350 remained non-significant at trend-level ($F_{4,112} = 2.2898$, $p=0.06412$). The significant three-
351 way interaction between accuracy, evidence and group persisted ($F_{2,15094} = 3.0871$,
352 $p=0.04566$). Importantly, when performing a Chi-square test on the log-likelihood values of
353 the models excluding and including the demographic variables to compare model fit, the
354 model without demographics showed a better model fit ($\chi^2 = 2.7018$, $df=2$, $p=0.259$), thereby
355 favoring this simpler model. Additionally, to investigate how confidence was differently
356 affected by sex in our healthy controls, we performed a two-sample t-test which showed that
357 males were generally more confident than females (males: 76.51 ± 1.04 ; females: $71.70 \pm$
358 0.77) ($t_{52} = 2.6518$, $p\text{-value}=0.01057$). However, both sex and IQ did not show a significant
359 influence on confidence level in Model 4.

360 Next to confidence, we also examined calibration and metacognitive sensitivity (see
361 **Supplementary Materials**). In short, we showed that GD patients were more overconfident
362 than OCD patients, without an effect of incentive condition. No differences in metacognitive
363 sensitivity were found between groups or incentive conditions.

364

365 **fMRI results GLM 1**

366 We analyzed functional neuroimaging data to test for differences in brain activity between
367 groups for our contrasts of interest: (1) choice moment modulated by early certainty, (2)
368 rating/incentive moment modulated by incentive value, and (3) rating/incentive moment

369 modulated by confidence. The results from the fMRI group analysis revealed no significant
370 differences between the groups for any of our contrasts.

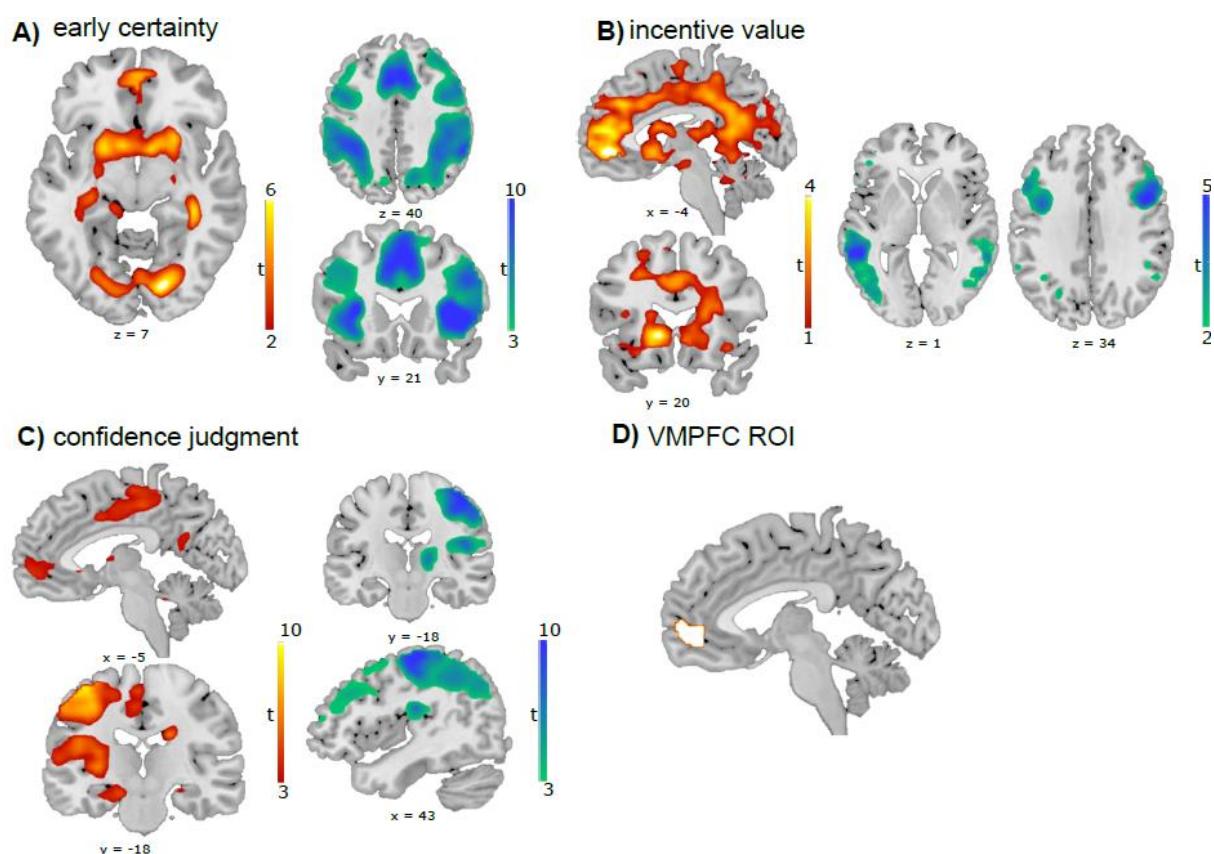
371

372 Next, we grouped all subjects together and performed one-sample t-tests on our contrasts of
373 interest to examine the results across groups (cluster-generating voxel threshold $p < .001$
374 uncorr.; clusterwise correction for multiple comparisons $p_{FWE} < 0.05$). During choice, early
375 certainty positively correlated with activation in the precuneus, VMPFC, bilateral VS and
376 putamen, and bilateral visual areas (**Figure 3A**). The dorsal anterior cingulate cortex,
377 bilateral dorsomedial- and dorsolateral prefrontal cortex, bilateral insula, thalamus, middle
378 frontal gyrus, bilateral sensorimotor cortex, superior and inferior parietal lobe related
379 negatively to early certainty (**Figure 3A**).

380 At the moment of incentive presentation, the incentive value correlated positively with
381 activation in the VS and VMPFC stretching into more dorsal areas, as well as the superior
382 temporal gyrus (**Figure 3B**). Incentive value was negatively related to activity in the right
383 (pre)motor cortex and dorsolateral PFC, as well as the left middle and superior temporal
384 gyrus, left occipitotemporal gyrus, and left middle and inferior frontal gyrus. Moreover, activity
385 in right lateral occipitotemporal gyrus and middle temporal gyrus were negatively related to
386 incentive value (**Figure 3B**).

387 During rating moment, confidence was positively related to activity in the VMPFC, left motor
388 cortex and putamen and bilateral visual areas (**Figure 3C**). The following areas were
389 negatively related to confidence: the left superior and inferior parietal lobes, right dorsolateral
390 PFC, right supramarginal gyrus and thalamus, right motor cortex stretching into the
391 dorsolateral PFC, left visual cortex and cerebellum (**Figure 3C**). See **Table 5** for details of
392 across group fMRI results.

393



394

Figure 3 | Whole brain statistical bold-oxygen-level-dependent (BOLD) activity across groups. Red/yellow areas represent areas with a positive relationship, while green/blue areas represent areas that have a negative relationship. (A) Areas correlating significantly with early certainty at choice moment. Shown are positive activations in ventromedial prefrontal cortex, ventral striatum and visual cortices. Negative activations in dorsal anterior cingulate cortex, dorsolateral prefrontal cortices, insula, parietal cortices. (B) Areas correlating significantly with incentive value at incentive/rating moment. Shown are positive activations in ventromedial prefrontal cortex, anterior cingulate cortex, ventral striatum. Negative activations in dorsolateral prefrontal cortices and temporal gyri (C) Areas correlating significantly with confidence judgments at incentive/rating moment. Positive actions are shown in ventromedial prefrontal cortex, motor cortex and putamen. Negative clusters in motor cortex and dorsolateral prefrontal cortex. All clusters survived $P < 0.05$ FWE cluster correction. Voxel-wise cluster-defining threshold was set at $P < .001$, uncorrected. For whole brain activation table see table 5. (D) Region of interest (ROI) of the VMPFC used for GLM2 analyses.

408

Effect	Brain Region	k	Peak z- score	P (FWE cluster corrected)	Peak MNI x	y	z	Hemi- sphere
Early Certainty +	Precuneus	2180	6.66	<.001	-6	-34	11	LR
	Ventromedial PFC							
	Ventral Striatum							
	Putamen							
	Lingual gyrus (visual cortex)	154	6.39	<.001	18	-81	-4	R
Early Certainty -	Lingual gyrus (visual cortex)	54	4.49	0.045	-21	-79	-4	L
	Dorsal Anterior Cingulate	13299	Inf (>8)	<.001	45	14	2	LR
	Dorsomedial PFC							
	Dorsolateral PFC							
	Insula							
	Thalamus							
	Middle Frontal Gyrus							
	Precentral Gyrus							
	Postcentral Gyrus							
	Supramarginal Gyrus							
	Superior Parietal Lobe							
	Inferior Parietal Lobe							
	Calcarine gyrus (visual cortex)							
	Middle Occipital Lobe	451	7.06	<.001	-30	-91	-4	L
	Middle Temporal Gyrus				-48	-67	-1	
	Lateral Occipito-temporal Gyrus				-45	-61	-10	
	Right Cerebellum	144	6.64	<.001	33	-55	-31	R
	Ventral Striatum	74	4.75	.004	-12	11	-4	L

Incentive Value +	Ventromedial PFC	212	4.53	<.001	-3 -9 0	44 50 35	-4 -4 14	LR
	Dorsomedial PFC							
	Superior Temporal Gyrus	48	4.25	.026	-45 -39	-16 -22	-1 5	L
Incentive Value -	Precentral gyrus stretching into premotor cortex and dorsolateral PFC	283	5.81	<.001	39 45 48	11 5 14	26 32 29	R
	Middle temporal gyrus	277	5.26	<.001	-54 -51 -48	-43 -52 -25	2 11 -7	L
	Superior temporal gyrus							
	Lateral occipitotemporal gyrus	183	5.06	<.001	-45 -24 -24	-61 -73 -82	-13 -7 -10	L
	Medial occipitotemporal gyrus							
	Middle frontal gyrus	299	4.93	<.001	-45 -39 -54	2 17 17	53 23 14	L
	Inferior frontal gyrus							
	Lateral occipitotemporal gyrus	116	4.90	<.001	42 45	-58 -49	-13 -13	R
	Middle temporal gyrus	47	3.74	.029	57 60 57	-46 -46 -61	11 2 2	R
Confidence +	Middle occipitotemporal gyrus	1947	Inf (>8)	<.001	12 21	-73 -70	-10 -7	R
	Lateral occipitotemporal gyrus				15	-52	-16	

	Cerebellum							
	Motor cortex (precentral gyrus)	993	Inf (>8)	<.001	-36 -36 -54	-25 -19 -16	65 47 47	L
	Putamen	968	5.91	<.001	-30 -45 -30	-19 -16 -22	2 20 14	L
	Rolandic operculum							
	Occipital lobe	65	4.58	.011	42	-67	5	R
	Ventromedial PFC	92	4.39	.002	-3 -12 -19	56 47 41	-4 8 -1	LR
Confidence -	Lingual gyrus (visual cortex)	1144	Inf (>8)	<.001	-9	-79	-7	L
	Cerebellum				-15 -24	-52 -67	-22 -28	
	Motor cortex (precentral gyrus)	2421	Inf (>8)	<.001	45 42 39	-16 -37 -52	59 62 41	R
	Stretching into dorsolateral PFC							
	Supramarginal gyrus	262	6.92	<.001	45 15	-19 -22	20 2	R
	Thalamus							
	Superior parietal lobe	168	5.09	<.001	-33 -39 -39	-58 -52 -43	41 47 41	L
	Inferior parietal lobe							
	Middle frontal gyrus (Dorsolateral PFC)	71	4.49	.007	-45 -45	32 23	32 35	R

410 **Table 5 | Whole brain activation tables.** Brain activations (whole brain analyses) showing activity
 411 related to early certainty at choice moment, as well as activity related to incentive and confidence at
 412 incentive/rating moment. All whole-brain activation maps were thresholded using family-wise error
 413 correction for multiple correction (FWE) at cluster level ($P_{\text{FWE_clu}} < 0.05$), with a voxel cluster-
 414 defining threshold of $P < 0.001$ uncorrected. Activity that positively correlates to given variable is
 415 denoted by '+', whereas negative correlations are denoted by '-'. PFC = prefrontal cortex.

416

417 *Interaction between metacognition and incentives in VMPFC (GLM 2)*

418 Our recent study suggested an important role of the VMPFC in the interaction between

419 incentive-processing and metacognitive signals³⁰. To investigate how this interaction takes

420 effect in and differs between our clinical groups, we performed an ROI analysis by leveraging

421 our factorial design. We extracted VMPFC activations for both time points (choice and

422 rating), all incentives (loss, neutral and gain), and all groups (HC, OCD and GD), for both

423 baseline activity and a regression slope with (1) signed evidence and (2) confidence

424 judgments (see **Figure 3D** for the ROI).

425 First, one-sample t-tests showed that, overall, VMPFC baseline activations were negative at

426 choice and rating moment (choice: $t_{100} = -3.611$, $p < 0.001$; baseline: $t_{100} = -4.9287$, $p < 0.001$).

427 The correlations between VMPFC activity and both signed evidence at choice moment and

428 confidence at rating moment, however, were significantly positive (choice: $t_{100} = 3.057$,

429 $p = 0.003$; baseline: $t_{100} = 3.7399$, $p < 0.001$) (**Figure 4**). This implies that the VMPFC

430 represents both confidence judgments and signed evidence (i.e. interaction between

431 accuracy and evidence: increased VMPFC activity with increased evidence when correct and

432 vice versa).

433 Then, we investigated whether there were effects of incentive condition and group around

434 this general signal. As expected, at choice moment there were no effects of incentive

435 condition on VMPFC baseline activity, nor on its correlation with the signed evidence signal

436 (i.e. slope) (**Figure 4, Table 6**). Despite the behavioral group effect on evidence integration,

437 we did not find a group nor interaction effect on both baseline VMPFC activity and the

438 correlation with signed evidence. At rating moment, however, incentive condition had a

439 significant effect on both the baseline VMPFC activity, as well as its correlation with

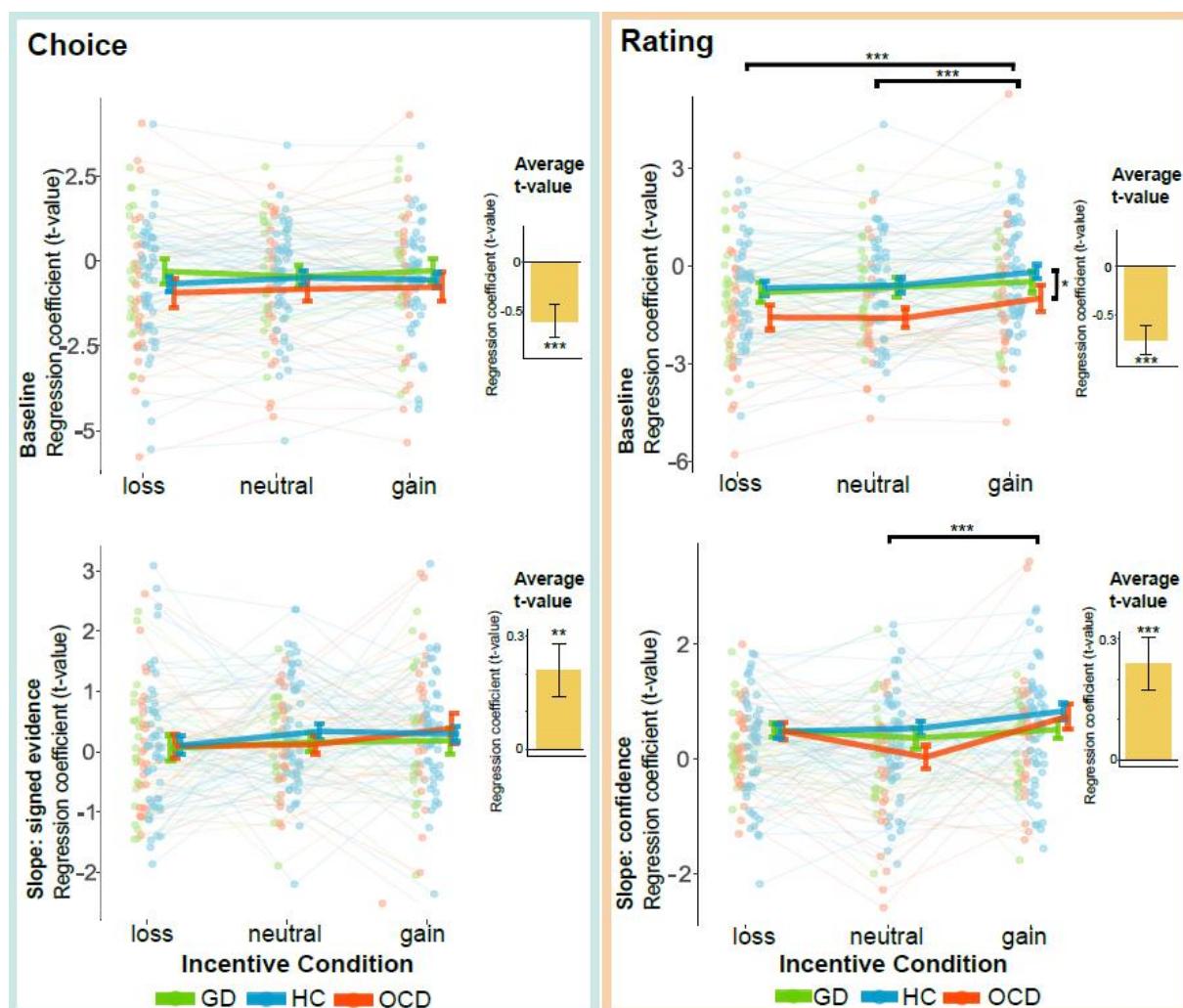
440 confidence. Post-hoc testing showed that the baseline VMPFC activity was higher during

441 gain versus loss ($t_{196} = -3.874$, $p < 0.001$), and during gain versus neutral ($t_{196} = -3.228$,

442 p<0.001), but no differences between neutral and loss conditions were found ($t_{196} = -0.646$,
443 p=0.7948). The correlation of VMPFC activity with confidence was significantly higher (i.e.
444 increased slope) in gain versus neutral ($t_{196} = -3.053$, p=0.0072), while no differences
445 between gain and loss, or between neutral and loss were found. Moreover, there was a
446 significant group effect on VMPFC baseline activity during rating moment. The post-hoc tests
447 revealed that OCD subjects had significantly decreased activity compared with HCs,
448 averaged over incentive conditions ($t_{98} = -2.515$, p=0.0358). No interaction effects between
449 group and incentive were found on baseline activity or its correlation with confidence at rating
450 moment.

451 Similar analyses using a ROI of the VS were performed (see Supplementary Materials), with
452 similar results: VS activity correlated with signed evidence, but no incentive, group or
453 interaction effects were found at choice moment. Similarly, the correlation of VS activity with
454 confidence was significantly higher in gain versus neutral, with no group difference at rating
455 moment.

456



457

Figure 4 | Ventromedial prefrontal cortex region of interest (ROI) analysis. T-values corresponding to baseline and regression slopes were extracted for all three groups and three incentive conditions, at two time points of interest: choice and incentive/rating moment. Green dots and lines represent gambling disorder patients, blue dots and lines represent healthy controls and red dots and lines represent obsessive-compulsive disorder patients. Dots represent individual t-statistics, and error bars represent sample mean \pm SEM per group. Black bars represent significant post-hoc tests. Yellow bars represent average t-values, with corresponding significance level of one-sample t-tests against 0. (* $p<0.05$, ** $p<0.01$, *** $p<0.001$). GD = gambling disorder, HC = healthy control, OCD = obsessive-compulsive disorder.

467

468

	Incentive	Group	Incentive:Group
Choice Baseline	$F(1.99, 195.28) = 0.37, p = 0.687$	$F(2, 98) = 0.54, p = 0.582$	$F(3.99, 195.28) = 0.41, p = 0.803$
Choice Slope	$F(1.99, 195) = 1.15, p = 0.320$	$F(2, 98) = 0.20, p=0.819$	$F(3.98, 195) = 0.31, p = 0.869$
Rating Baseline	$F(1.91, 186.81) = 8.61, p < 0.001$	$F(2, 98) = 3.24, p = 0.044$	$F(3.81, 186.81) = 0.44, p = 0.771$
Rating Slope	$F(1.92, 187.68) = 4.67, p = 0.012$	$F(2, 98) = 0.99, p = 0.375$	$F(3.83, 187.68) = 1.29, p = 0.277$

469 **Table 6: Results of VMPFC ROI analysis.** Shown here are the results of the mixed ANOVAs of t-
470 statistics in the ventromedial prefrontal cortex (VMPFC) region of interest (ROI) using the afex
471 package. Shown are the main effects of incentive condition, group and their interaction effect on the
472 choice and rating time points, focusing on both the baseline activity as well as the slope of signed
473 evidence and confidence judgments, respectively. F-values, with corresponding degrees of freedom
474 and p-values are reported.

475

476 **Discussion**

477 In this study we investigated the (neural signatures of) metacognitive ability and its
478 interaction with incentive motivation in two compulsive disorders: OCD and GD. First, we
479 replicated the biasing effect of incentives on confidence estimation in all groups, showing that
480 confidence was higher in the gain context and lower in the loss context. This is a robust
481 effect, that has now been independently replicated multiple times^{29–32}. We initially found
482 evidence for a significantly higher confidence in GD patients versus OCD patients, although
483 this effect diminished after controlling for sex and IQ differences between groups. Hence, we
484 only found moderate evidence for our hypothesis of group differences in confidence, as well
485 as for our hypothesis that incentive motivation would affect confidence judgments differently
486 in the groups. Future research should address the role of the demographic confounding
487 factors more specifically.

488 When looking into the computational signatures of confidence formation in more detail, GD
489 patients interestingly showed less integration of evidence into their confidence judgments for
490 correct choices compared to both HCs and OCD patients. This suggests that GD patients
491 were less able to use evidence they received to form confidence judgments. This decreased
492 sensitivity to objective evidence could fit GD's symptomatology of cognitive inflexibility^{3,65},
493 and cognitive distortions^{66,67}. Illusion of control leads pathological gamblers to believe they
494 can predict outcomes, rendering them less influenced by objective evidence, which may
495 promote continuation of (overconfident) gambling behavior^{13,68}.

496 Notably, our patient groups seemed to be situated on opposite sides of the confidence
497 spectrum, with GD patients being more confident than OCD patients. However, this effect
498 was partly driven by sex and IQ differences between groups. The GD group consisted mostly
499 of males, whereas the OCD group had a more mixed composition. mirroring the prevalence
500 distribution of these disorders^{69–72}. Consistent with our findings of increased confidence in HC
501 male subjects, recent studies have shown that males are more confident than females,

502 despite equal performance^{74,75}. Therefore, the effect of sex might have explained some
503 variance in our data, but does not fully explain the group differences, since we do find a trend
504 toward a group effect. The importance of taking into account sex and gender as factors in
505 both neuroscience and psychiatry research is increasingly recognized and acted upon⁷⁶,
506 since sex differences play a role in the incidence, treatment and manifestation of
507 psychopathology^{77,78}. The precise role of sex and gender in metacognition deserves more
508 attention and should be characterized further in future research.

509 Our data shows no convincing evidence for an exaggerated decrease/increase in confidence
510 during loss/gain anticipation in OCD/GD, respectively. However, the group*incentive
511 interaction approached significance, with increased confidence in GD patients compared to
512 both OCD patients and HCs, specifically in the gain condition. This finding agrees with
513 literature demonstrating increased reward sensitivity in GD^{79,80}. Confidence in OCD patients
514 has been mostly studied using metamemory paradigms, and abnormalities were most
515 profound in OCD-relevant contexts⁸¹⁻⁸⁶. Earlier studies probing confidence in GD are sparse,
516 and whilst they all did show an effect of overconfidence in (sub)clinical problem gamblers,
517 none of the studies actively controlled for performance differences, making it difficult to draw
518 strong conclusions about confidence biases^{16,17,87}.

519 Since confidence in GD and OCD did not differ from the healthy population we cannot
520 technically speak of confidence 'abnormalities' in GD and OCD. Future work is necessary to
521 study the link between compulsivity and confidence more directly. One interesting method is
522 transdiagnostic research to study metacognition in psychiatry. Transdiagnostic research
523 methods are useful, since (meta)cognition might relate more closely to symptoms than
524 diagnoses, due to high levels of comorbidity and heterogeneity of symptoms within disorders.
525 Indeed, a transdiagnostic factor of 'anxious-depression' was negatively related to confidence,
526 whereas 'compulsive behavior and intrusive thoughts' were positively related to confidence
527 and showed decoupling of confidence and behavior by diminished utilizing of perceptual

528 evidence for confidence judgments⁸⁸. This latter result is in line with our findings of
529 diminished evidence integration into confidence judgments in GD patients.

530
531 The brain areas we found to be related to confidence and incentive processing converge with
532 earlier work. Confidence was found to be positively related to the VMPFC via automatic
533 processing at the choice moment^{20,46,47,55}. Early certainty processing was also positively
534 related to activity in the VS and precuneus^{39,49,51}. We also observed a wide-spread network of
535 areas negatively related to early certainty, containing the dACC, dorsolateral PFC, insula,
536 inferior parietal lobe and midfrontal gyrus, a network repeatedly associated with uncertainty
537 and metacognitive processes^{39,44,45,51}. Also, well-known relationships between reward
538 processing and activity in both VS and VMPFC^{21,22} were replicated. Moreover, we found
539 negative relationships between incentive value and BOLD activity in the central executive
540 network (i.e. lateral PFC and middle frontal gyrus), as well as superior temporal gyrus^{89,90}.
541 Confidence was found to be related to VMPFC activity, not only at choice moment, but also
542 during rating^{20,46,47}. Overall, our fMRI findings closely resemble activation patterns previously
543 shown in healthy populations.

544 We also replicated the effect of incentive condition on VMPFC baseline activity and on the
545 correlation of VMPFC activity with confidence, which was highest in gain conditions, which
546 we also found in the VS³⁰. While we found aberrant evidence integration in GD patients on a
547 behavioral level, we did not find any group differences in evidence processing on
548 neurobiological level. Interestingly, OCD patients showed a decreased baseline VMPFC
549 activity during incentive/rating moment, which fits with earlier work showing neurobiological
550 deficits in a ‘ventral motivational circuit’ including the VMPFC^{91,92}. However, we did not find
551 any interactions with incentive condition in the VMPFC activity related to either signed
552 evidence or confidence.

553 In sum, contrary to our hypotheses, we did not find neurobiological deficits directly related to
554 confidence or to the effects of incentive on confidence in our clinical samples. This might not
555 be surprising, given that the behavioral group effects were small (and disappeared when
556 controlling for demographics), which limited our ability a priori to find impairments in neural
557 circuits mediating confidence processes. Because, to our knowledge, the present study
558 represents the first attempt in investigating the joint neural basis of metacognitive and reward
559 processes in both GD and OCD, further study - e.g. looking into transdiagnostic variations of
560 symptoms - might be more powerful in detecting clinically useful neurocognitive signatures of
561 those processes than the present clinical case-control comparisons⁹³.

562

563 **Acknowledgements:** Data collection for this work was funded by two independent personal
564 Amsterdam Brain and Cognition (ABC) Talent grants to JL and RvH, and a NWO Veni
565 Fellowship (grant 451-15-015) granted to ML. ML is supported by a Swiss National Fund
566 Ambizione Grant (PZ00P3_174127) and an ERC Starting Grant (ERC-StG-948671), JL is
567 supported by a NWO VENI Fellowship grant (916-18-119).

568 **Disclosures:** None of the authors have any conflicts of interest to declare.

569

570 **Literature**

571 1 Luigjes J, Lorenzetti V, de Haan S, Youssef GJ, Murawski C, Sjoerds Z *et al.* Defining Compulsive
572 Behavior. *Neuropsychol Rev* 2019; **29**: 4–13.

573 2 Stein DJ. Obsessive-compulsive disorder. *Lancet* 2002; **360**: 397–405.

574 3 van Timmeren T, Daams JG, van Holst RJ, Goudriaan AE. Compulsivity-related neurocognitive
575 performance deficits in gambling disorder: A systematic review and meta-analysis. *Neurosci
576 Biobehav Rev* 2018; **84**: 204–217.

577 4 Chamberlain SR, Blackwell AD, Fineberg NA, Robbins TW, Sahakian BJ. The neuropsychology
578 of obsessive compulsive disorder: The importance of failures in cognitive and behavioural
579 inhibition as candidate endophenotypic markers. *Neurosci Biobehav Rev* 2005; **29**: 399–419.

580 5 Figue M, Patti T, Willuhn I, Luigjes J, van den Brink W, Goudriaan A *et al.* Compulsivity in
581 obsessive-compulsive disorder and addictions. *Eur Neuropsychopharmacol* 2016; **26**: 856–
582 868.

583 6 Fleming SM, Dolan RJ, Frith CD. Metacognition: Computation, biology and function. *Philos
584 Trans R Soc B Biol Sci* 2012; **367**: 1280–1286.

585 7 Pouget A, Drugowitsch J, Kepcs A. Confidence and certainty: distinct probabilistic quantities
586 for different goals. *Nat Neurosci* 2016; **19**: 366–374.

587 8 Meyniel F, Sigman M, Mainen ZF. Perspective Confidence as Bayesian Probability: From
588 Neural Origins to Behavior. *Neuron* 2015; **88**: 78–92.

589 9 Folke T, Jacobsen C, Fleming SM, De Martino B. Explicit representation of confidence informs
590 future value-based decisions. *Nat Hum Behav* 2017; **1**. doi:10.1038/s41562-016-0002.

591 10 Samaha J, Switzky M, Postle BR. Confidence boosts serial dependence in orientation
592 estimation. *J Vis* 2019; **19**: 25–25.

593 11 Hoven M, Lebreton M, Engelmann JB, Denys D, Luigjes J, van Holst RJ. Abnormalities of
594 confidence in psychiatry: an overview and future perspectives. *Transl Psychiatry* 2019; **9**: 1–
595 18.

596 12 Fortune EE, Goodie AS. Cognitive distortions as a component and treatment focus of
597 pathological gambling: A review. *Psychol Addict Behav* 2012; **26**: 298–310.

598 13 Goodie AS, Fortune EE. Measuring cognitive distortions in pathological gambling: Review and
599 meta-analyses. *Psychol Addict Behav* 2013; **27**: 730–743.

600 14 Nestadt G, Kamath V, Maher BS, Krasnow J, Nestadt P, Wang Y *et al.* Doubt and the decision-
601 making process in obsessive-compulsive disorder. *Med Hypotheses* 2016; **96**: 1–4.

602 15 Samuels J, Bienvenu OJ, Krasnow J, Wang Y, Grados MA, Cullen B *et al.* An investigation of
603 doubt in obsessive-compulsive disorder. *Compr Psychiatry* 2017; **75**: 117–124.

604 16 Goodie AS. The role of perceived control and overconfidence in pathological gambling. *J
605 Gambl Stud* 2005; **21**: 481–502.

606 17 Lakey CE, Goodie AS, Campbell WK. Frequent card playing and pathological gambling: The
607 utility of the Georgia Gambling Task and Iowa Gambling Task for predicting pathology. *J
608 Gambl Stud* 2007; **23**: 285–297.

609 18 Pessoa L, Engelmann JB. Embedding reward signals into perception and cognition. *Front. Neurosci.* 2010. doi:10.3389/fnins.2010.00017.

611 19 Rangel A, Camerer C, Montague PR. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 2008; **9**: 545–556.

613 20 Lopez-Persem A, Bastin J, Petton M, Abitbol R, Lehongre K, Adam C *et al.* Four core properties of the human brain valuation system demonstrated in intracranial signals. *Nat Neurosci* 2020; **23**: 664–675.

616 21 Bartra O, McGuire JT, Kable JW. The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* 2013; **76**: 412–427.

619 22 Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M. An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging. *Neuron* 2009; **64**: 431–439.

621 23 Pessiglione M, Lebreton M. From the Reward Circuit to the Valuation System: How the Brain Motivates Behavior. 2015; : 1–421.

623 24 Clark L, Boileau I, Zack M. Neuroimaging of reward mechanisms in Gambling disorder: an integrative review. *Mol Psychiatry* 2019; **24**: 674–693.

625 25 Shephard E, Stern ER, van den Heuvel OA, Costa DLC, Batistuzzo MC, Godoy PBG *et al.* Toward a neurocircuit-based taxonomy to guide treatment of obsessive-compulsive disorder. *Mol Psychiatry* 2021 269 2021; **26**: 4583–4604.

628 26 Allen M, Frank D, Samuel Schwarzkopf D, Fardo F, Winston JS, Hauser TU *et al.* Unexpected arousal modulates the influence of sensory noise on confidence. *Elife* 2016; **5**: 1–17.

630 27 Koellinger P, Treffers T. Joy leads to overconfidence, and a simple countermeasure. *PLoS One* 2015; **10**: 1–22.

632 28 Massoni S. Emotion as a boost to metacognition: How worry enhances the quality of confidence. *Conscious Cogn* 2014; **29**: 189–198.

634 29 Ting CC, Palminteri S, Engelmann JB, Lebreton M. Robust valence-induced biases on motor response and confidence in human reinforcement learning. *Cogn Affect Behav Neurosci* 2020; **20**: 1184–1199.

637 30 Hoven M, Brunner G, De Boer NS, Goudriaan A, Denys D, Van Holst RJ *et al.* How motivational signals disrupt metacognitive signals in the human VMPFC. *bioRxiv* 2020; : 2020.10.02.323550.

639 31 Lebreton M, Langdon S, Slieker MJ, Nooitgedacht JS, Goudriaan AE, Denys D *et al.* Two sides of the same coin: Monetary incentives concurrently improve and bias confidence judgments. *Sci Adv* 2018; **4**: eaqq0668.

642 32 Lebreton M, Bacilly K, Palminteri S, Engelmann JB. Contextual influence on confidence judgments in human reinforcement learning. *PLoS Comput Biol* 2019; **15**: e1006973.

644 33 Bey K, Lennertz L, Riesel A, Klawohn J, Kaufmann C, Heinzel S *et al.* Harm avoidance and childhood adversities in patients with obsessive-compulsive disorder and their unaffected first-degree relatives. *Acta Psychiatr Scand* 2017; **135**: 328–338.

647 34 Bey K, Weinhold L, Grützmann R, Heinzel S, Kaufmann C, Klawohn J *et al.* The polygenic risk for obsessive-compulsive disorder is associated with the personality trait harm avoidance. *Acta Psychiatr Scand* 2020; **142**: 326–336.

650 35 Summerfeldt LJ, Kloosterman PH, Antony MM, Swinson RP. Examining an obsessive-
651 compulsive core dimensions model: Structural validity of harm avoidance and
652 incompleteness. *J Obsessive Compuls Relat Disord* 2014; **3**: 83–94.

653 36 Baird B, Smallwood J, Gorgolewski KJ, Margulies DS. Medial and lateral networks in anterior
654 prefrontal cortex support metacognitive ability for memory and perception. *J Neurosci* 2013;
655 **33**: 16657–16665.

656 37 Allen M, Glen JC, Müllensiefen D, Schwarzkopf DS, Fardo F, Frank D *et al.* Metacognitive ability
657 correlates with hippocampal and prefrontal microstructure. *Neuroimage* 2017; **149**: 415–423.

658 38 Hilgenstock R, Weiss T, Witte OW. You'd Better Think Twice: Post-Decision Perceptual
659 Confidence. *Neuroimage* 2014; **99**: 323–331.

660 39 Vaccaro AG, Fleming SM. Thinking about thinking: A coordinate-based meta-analysis of
661 neuroimaging studies of metacognitive judgements. *Brain Neurosci Adv* 2018; **2**:
662 239821281881059.

663 40 Fleming SM, Weil RS, Nagy Z, Dolan RJ, Rees G. Relating introspective accuracy to individual
664 differences in brain structure. *Science (80-)* 2010; **329**: 1541–1543.

665 41 Rouault M, Fleming SM. Formation of global self-beliefs in the human brain. *Proc Natl Acad Sci
666 U S A* 2020; **117**: 27268–27276.

667 42 Shenhav A, Cohen JD, Botvinick MM. Dorsal anterior cingulate cortex and the value of control.
668 *Nat Neurosci* 2016; **19**: 1286–1291.

669 43 Fleming SM, Van Der Putten EJ, Daw ND. Neural mediators of changes of mind about
670 perceptual decisions. *Nat Neurosci* 2018; **21**: 617–624.

671 44 Molenberghs P, Trautwein F-M, Böckler A, Singer T, Kanske P. Neural correlates of
672 metacognitive ability and of feeling confident: a large-scale fMRI study. *Soc Cogn Affect
673 Neurosci* 2016; **11**: 1942–1951.

674 45 Morales J, Lau H, Fleming SM. Domain-General and Domain-Specific Patterns of Activity
675 Supporting Metacognition in Human Prefrontal Cortex. *J Neurosci* 2018; **38**: 2360–17.

676 46 De Martino B, Fleming SM, Garrett N, Dolan RJ. Confidence in value-based choice. *Nat
677 Neurosci* 2013; **16**: 105–110.

678 47 Lebreton M, Abitbol R, Daunizeau J, Pessiglione M. Automatic integration of confidence in the
679 brain valuation signal. *Nat Neurosci* 2015; **18**: 1159–67.

680 48 Gherman S, Philiastides MG. Human VMPFC encodes early signatures of confidence in
681 perceptual decisions. *Elife* 2018; **7**. doi:10.7554/eLife.38293.

682 49 Rouault M, Mcwilliams A, Allen MG, Fleming SM. Human Metacognition Across Domains:
683 Insights from Individual Differences and Neuroimaging. *Personal Neurosci* 2018; **1**: 17.

684 50 Bang D, Fleming SM. Distinct encoding of decision confidence in human medial prefrontal
685 cortex. *Proc Natl Acad Sci U S A* 2018; **115**: 6082–6087.

686 51 Hebart MN, Schriever Y, Donner TH, Haynes J-D. The Relationship between Perceptual
687 Decision Variables and Confidence in the Human Brain. *Cereb Cortex* 2016; **26**: 118–130.

688 52 Haber SN, Knutson B. The Reward Circuit: Linking Primate Anatomy and Human Imaging.
689 *Neuropsychopharmacol* 2010 351 2009; **35**: 4–26.

690 53 Haber SN, Behrens TEJ. The Neural Network Underlying Incentive-Based Learning:
691 Implications for Interpreting Circuit Disruptions in Psychiatric Disorders. *Neuron* 2014; **83**:
692 1019–1039.

693 54 Gherman S, Philiastides MG. Neural representations of confidence emerge from the process
694 of decision formation during perceptual choices. *Neuroimage* 2015; **106**: 134–143.

695 55 Shapiro AD, Grafton ST. Subjective value then confidence in human ventromedial prefrontal
696 cortex. *PLoS One* 2020; **15**: e0225617.

697 56 Bates D, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using lme4. *J
698 Stat Softw* 2015; **67**: 1–48.

699 57 Singmann H, Bolker B, Westfall J. Analysis of Factorial Experiments, package ‘afex’. 2015; : 1–
700 44.

701 58 Lenth R, Singmann H, Love J, Buerkner P, Herve M. Emmeans: Estimated marginal means, aka
702 least-squares means. *R Packag* 2018; **3**. doi:10.1080/00031305.1980.10483031.

703 59 Lebreton M, Bavard S, Daunizeau J, Palminteri S. Assessing inter-individual differences with
704 task-related functional neuroimaging. *Nat Hum Behav* 2019; **3**: 897–905.

705 60 Black DW, Shaw M. The Epidemiology of Gambling Disorder. *Gambl Disord* 2019; : 29–48.

706 61 Calado F, Griffiths MD. Problem gambling worldwide: An update and systematic review of
707 empirical research (2000-2015). *J Behav Addict* 2016; **5**: 592–613.

708 62 Ruscio AM, Stein DJ, Chiu WT, Kessler RC. The epidemiology of obsessive-compulsive disorder
709 in the National Comorbidity Survey Replication. *Mol Psychiatry* 2010 151 2008; **15**: 53–63.

710 63 Mathes BM, Morabito DM, Schmidt NB. Epidemiological and Clinical Gender Differences in
711 OCD. *Curr Psychiatry Reports* 2019 215 2019; **21**: 1–7.

712 64 Sanders JI, Hangya BB, Kepecs A. Signatures of a Statistical Computation in the Human Sense
713 of Confidence. *Neuron* 2016; **90**: 499–506.

714 65 Perandrés-Gómez A, Navas JF, van Timmeren T, Perales JC. Decision-making (in)flexibility in
715 gambling disorder. *Addict Behav* 2021; **112**: 106534.

716 66 Ledgerwood DM, Dyshniku F, McCarthy JE, Ostojic-Aitkens D, Forfitt J, Rumble SC. Gambling-
717 Related Cognitive Distortions in Residential Treatment for Gambling Disorder. *J Gambl Stud*
718 2020; **36**: 669–683.

719 67 Mallorquí-Bagué N, Vintró-Alcaraz C, Verdejo-García A, Granero R, Fernández-Aranda F,
720 Magaña P *et al.* Impulsivity and cognitive distortions in different clinical phenotypes of
721 gambling disorder: Profiles and longitudinal prediction of treatment outcomes. *Eur Psychiatry*
722 2019; **61**: 9–16.

723 68 Cowley E, Briley DA, Farrell C. How do gamblers maintain an illusion of control? *J Bus Res*
724 2015; **68**: 2181–2188.

725 69 Howe PDL, Vargas-Sáenz A, Hulbert CA, Boldero JM. Predictors of gambling and problem
726 gambling in Victoria, Australia. *PLoS One* 2019; **14**: e0209277.

727 70 Subramaniam M, Wang P, Soh P, Vaingankar JA, Chong SA, Browning CJ *et al.* Prevalence and
728 determinants of gambling disorder among older adults: A systematic review. *Addict Behav*
729 2015; **41**: 199–209.

730 71 Welte JW, Barnes GM, Tidwell MCO, Wieczorek WF. Predictors of Problem Gambling in the
731 U.S. *J Gambl Stud* 2017; **33**: 327–342.

732 72 Swedo SE, Rapoport JL, Leonard H, Lenane M, Cheslow D. Obsessive-Compulsive Disorder in
733 Children and Adolescents: Clinical Phenomenology of 70 Consecutive Cases. *Arch Gen
734 Psychiatry* 1989; **46**: 335–341.

735 73 Dowling NA, Merkouris SS, Greenwood CJ, Oldenhof E, Toumbourou JW, Youssef GJ. Early risk
736 and protective factors for problem gambling: A systematic review and meta-analysis of
737 longitudinal studies. *Clin Psychol Rev* 2017; **51**: 109–124.

738 74 Ariel R, Lembeck NA, Moffat S, Hertzog C. Are there sex differences in confidence and
739 metacognitive monitoring accuracy for everyday, academic, and psychometrically measured
740 spatial ability? *Intelligence* 2018; **70**: 42–51.

741 75 Rivers ML, Fitzsimmons CJ, Fisk SR, Dunlosky J, Thompson CA. Gender differences in
742 confidence during number-line estimation. *Metacognition Learn* 2021; **16**: 157–178.

743 76 Cahill L. Why sex matters for neuroscience. *Nat Rev Neurosci* 2006; **7**: 477–484.

744 77 Cosgrove KP, Mazure CM, Staley JK. Evolving Knowledge of Sex Differences in Brain Structure,
745 Function, and Chemistry. *Biol Psychiatry* 2007; **62**: 847–855.

746 78 Gobinath AR, Choleris E, Galea LAM. Sex, hormones, and genotype interact to influence
747 psychiatric disease, treatment, and behavioral research. *J Neurosci Res* 2017; **95**: 50–64.

748 79 Navas JF, Billieux J, Perandrés-Gómez A, López-Torrecillas F, Cándido A, Perales JC. Impulsivity
749 traits and gambling cognitions associated with gambling preferences and clinical status. *Int
750 Gambl Stud* 2017; **17**: 102–124.

751 80 Van Holst RJ, Veltman DJ, Bichel C, Van Den Brink W, Goudriaan AE. Distorted expectancy
752 coding in problem gambling: Is the addictive in the anticipation? *Biol Psychiatry* 2012; **71**:
753 741–748.

754 81 Boschen MJ, Vuksanovic D. Deteriorating memory confidence, responsibility perceptions and
755 repeated checking: Comparisons in OCD and control samples. *Behav Res Ther* 2007; **45**: 2098–
756 2109.

757 82 Bucarelli B, Purdon C. Stove checking behaviour in people with OCD vs. anxious controls. *J
758 Behav Ther Exp Psychiatry* 2016; **53**: 17–24.

759 83 Hermans D, Engelen U, Grouwels L, Joos E, Lemmens J, Pieters G. Cognitive confidence in
760 obsessive-compulsive disorder: Distrusting perception, attention and memory. *Behav Res Ther*
761 2008; **46**: 98–113.

762 84 Moritz S, Wahl K, Zurowski B, Jelinek L, Hand I, Fricke S. Enhanced perceived responsibility
763 decreases metamemory but not memory accuracy in obsessive-compulsive disorder (OCD).
764 *Behav Res Ther* 2007; **45**: 2044–2052.

765 85 Radomsky AS, Rachman S, Hammond D. Memory bias, confidence and responsibility in
766 compulsive checking. *Behav Res Ther* 2001; **39**: 813–822.

767 86 Tolin DF, Abramowitz JS, Brigidi BD, Amir N, Street GP, Foa EB. Memory and memory
768 confidence in obsessive – compulsive disorder. *Behav Res Ther* 2001; **39**: 913–927.

769 87 Brevers D, Cleeremans A, Bechara A, Greisen M, Kornreich C, Verbanck P *et al.* Impaired
770 Metacognitive Capacities in Individuals with Problem Gambling. *J Gambl Stud* 2014; **30**: 141–

771 152.

772 88 Seow TXF, Gillan CM. Transdiagnostic Phenotyping Reveals a Host of Metacognitive Deficits
773 Implicated in Compulsivity. *Sci Rep* 2020; **10**: 1–11.

774 89 Liu X, Hairston J, Schrier M, Fan J. Common and distinct networks underlying reward valence
775 and processing stages: A meta-analysis of functional neuroimaging studies. *Neurosci Biobehav
776 Rev* 2011; **35**: 1219–1236.

777 90 Wilson RP, Colizzi M, Bossong MG, Allen P, Kempton M, Abe N *et al.* The Neural Substrate of
778 Reward Anticipation in Health: A Meta-Analysis of fMRI Findings in the Monetary Incentive
779 Delay Task. *Neuropsychol Rev* 2018; **28**: 496–506.

780 91 Thorsen AL, Hagland P, Radua J, Mataix-Cols D, Kvale G, Hansen B *et al.* Emotional Processing
781 in Obsessive-Compulsive Disorder: A Systematic Review and Meta-analysis of 25 Functional
782 Neuroimaging Studies. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2018; **3**: 563–571.

783 92 Stein DJ, Costa DLC, Lochner C, Miguel EC, Reddy YCJ, Shavitt RG *et al.* Obsessive-compulsive
784 disorder. *Nat Rev Dis Prim* 2019 51 2019; **5**: 1–21.

785 93 Parkes L, Tiego J, Aquino K, Braganza L, Chamberlain SR, Fontenelle LF *et al.* Transdiagnostic
786 variations in impulsivity and compulsivity in obsessive-compulsive disorder and gambling
787 disorder correlate with effective connectivity in cortical-striatal-thalamic-cortical circuits.
788 *Neuroimage* 2019; **202**. doi:10.1016/j.neuroimage.2019.116070.

789

790

791

792