Metacognition and the effect of incentive motivation in two compulsive disorders: Gambling disorder and obsessive–compulsive disorder

Monja Hoven, MSc [1]* Nina S. de Boer, MSc, MA,[2] Anna E. Goudriaan, PhD,[1,3] Damiaan Denys, PhD,[1] Maël Lebreton, PhD,[4,5,6] Ruth J. van Holst, PhD[1†] and Judy Luigjes, PhD[1†]

Aims: Compulsivity is a common phenotype among psychiatric disorders, such as obsessive–compulsive disorder (OCD) and gambling disorder (GD). Deficiencies in metacognition, such as the inability to estimate one's performance via confidence judgments could contribute to pathological decision-making. Earlier research has shown that patients with OCD exhibit underconfidence, while patients with GD exhibit overconfidence. Moreover, it is known that motivational states (e.g. monetary incentives) influence metacognition, with gain (respectively loss) prospects increasing (respectively decreasing) confidence. Here, we reasoned that OCD and GD symptoms might correspond to an exacerbation of this interaction between metacognition and motivation.

Methods: We hypothesized GD's overconfidence to be exaggerated during gain prospects, while OCD's underconfidence to be worsened in loss context, which we expected to see represented in ventromedial prefrontal cortex (VMPFC) blood-oxygen-level-dependent activity. We tested those hypotheses in a task-based functional magnetic resonance imaging (fMRI) design (27 patients with GD,

28 patients with OCD, 55 controls). The trial is registered in the Dutch Trial Register (NL6171).

Results: We showed increased confidence for patients with GD versus patients with OCD, which could partly be explained by sex and IQ. Although our primary analyses did not support the hypothesized interaction between incentives and groups, exploratory analyses did show increased confidence in patients with GD specifically in gain context. fMRI analyses confirmed a central role for VMPFC in the processing of confidence and incentives, but no differences between the groups.

Conclusion: Patients with OCD and those with GD reside at opposite ends of the confidence spectrum, while no interaction with incentives was found, nor group differences in neuronal processing of confidence.

Keywords: confidence, gambling disorder, metacognition, obsessive–compulsive disorder, VMPFC.

http://onlinelibrary.wiley.com/doi/10.1111/pcn.13434/full

Compulsive behaviors are defined as "repetitive acts that are characterized by the feeling that one 'has to' perform them while being aware that these acts are not in line with one's overall goal."[1] Various psychiatric disorders are associated with compulsivity, of which obsessive–compulsive disorder (OCD) is the most typical,[2] but it is also seen in addictive disorders such as gambling disorder (GD).[3] Both disorders are characterized by a loss of control over their compulsive behaviors, albeit originating from distinct motivations, serving different purposes and relating to distinct symptoms.[4,5] Hence, compulsivity seems to be a common phenotype in otherwise symptomatically different disorders.

Dysfunctions in metacognition could explain distinct features of compulsive behaviors. Metacognition is the ability to monitor, reflect on, and think about our own behavior.[6] One metacognitive computation

is the judgment of confidence, defined as the subjective estimate of the probability of being correct about a choice.[7] Confidence plays a key role in decision-making and learning,[6–8] and therefore in steering our future behavior.[9,10] It is crucial for behavioral control that one's confidence is in line with reality. Nonetheless, discrepancies between actual behavior (e.g. choice accuracy) and confidence in that behavior (subjective estimate of accuracy) have been consistently described, which could contribute to pathological (compulsive) decision-making as seen in various psychiatric disorders.[11] Clinical presentations of OCD and GD indeed suggest confidence abnormalities in the opposite direction, underconfidence and overconfidence, respectively, which could both promote detrimental decision-making, such as checking behavior and compulsive gambling.[12–15] In a recent review we showed that both patients with subclinical and clinical OCD consistently showed a

decrease in confidence level, which was especially profound in OCD symptom contexts.[11] Oppositely, in pathological gamblers, there was evidence for overconfidence in rewarding gambling contexts, which was also related to symptom severity.[16,17] In sum, patients with GD and those with OCD seem to function at opposite sides of the confidence continuum, respectively overestimating and underestimating their performance, which could explain how opposite traits may underlie similar pathological behavior (i.e. compulsive behavior).

Reward processes are important for learning and decision-making and interact with cognition.[18] Many studies have implicated subcortical regions such as the ventral striatum (VS) and cortical regions such as the ventromedial prefrontal cortex (VMPFC) in reward processing, forming a "brain valuation system"[19–21] whose activity relates to value-based decision-making[22] and motivates behavior.[23] Both patients with OCD and those with GD show deficits in reward processes and accompanying dysregulated neural circuitries. A recent review on neuroimaging of reward mechanisms by Clark et al.[24] clearly indicated dysregulated reward circuitries, especially focused on the VMPFC and VS in patients with GD, with mixed evidence regarding the direction of these effects. In patients with OCD, a recent review showed that the ventral affective circuit, consisting of medial frontal cortex and VS was consistently shown to be dysregulated, showing decreased activity in response to rewards, which was increased in response to losses.[25] This is particularly relevant to the question of how confidence might contribute to those pathologies' symptoms, as an increasing number of studies show that affective and motivational states can influence confidence.[26–28] Recently, we demonstrated that monetary incentives bias confidence judgments in healthy individuals, where prospects of gain (respectively loss) increase (respectively decrease) confidence, while performance levels remained unaffected in both perceptual and reinforcement-learning contexts.[29–32]

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

On the neurobiological side, a growing number of functional magnetic resonance imaging (fMRI) studies have associated metacognitive processes with activity in the frontal–parietal network,[36–40]

and activity in the dorsomedial prefrontal cortex (PFC), insula, and dorsal anterior cingulate cortex (dACC) has been negatively associated with confidence, suggesting a role for these areas in representing uncertainty-related variables.[41–45] Interestingly, recent studies have also found activity in the VS, the VMPFC, and perigenual anterior cingulate cortex to be positively associated with confidence.[41,46–51] Importantly, this latter network has been previously positively associated with value-based processes.[20,21,52,53] Actually, both confidence judgments and value information seem to be automatically integrated into VMPFC's activity.[20,22,47,54,55] Yet, little is known about whether and how the behavioral interaction observed between incentives and confidence can be explained by their shared association with the VMPFC. In an attempt to answer this question, we recently reported an important interaction between incentive and metacognitive signals in the VMPFC in healthy individuals: confidence signals in the VMPFC were observed in trials with gain prospects, but disrupted in trials with no or negative (loss) monetary prospects.[30] This suggests that the VMPFC has a key role in mediating the relationship between incentives and metacognition. Given the crucial roles of the VMPFC and VS in reward processes and metacognition, which were found to be dysregulated in GD and OCD, we hypothesized that both regions would show disrupted activation patterns related to incentive processing and metacognition and their interaction in patients compared with healthy controls (HCs).

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

Methods

Ethics

Experimental procedures were approved by the medical ethics committee of the Academic Medical Center, University of Amsterdam. All patients provided written informed consent.

Participants

We recruited a total of 31 patients with GD, 29 patients with OCD, and 55 HCs between the ages of 18 and 65 years. Of our HC sample of 55 participants, 25 were included in our earlier work.[30] HCs were recruited through online advertisements and from our participant database. Patients with GD were recruited from a local treatment center (Jellinek Addiction Treatment Center Amsterdam) and were recently diagnosed with GD. Patients with OCD were recruited
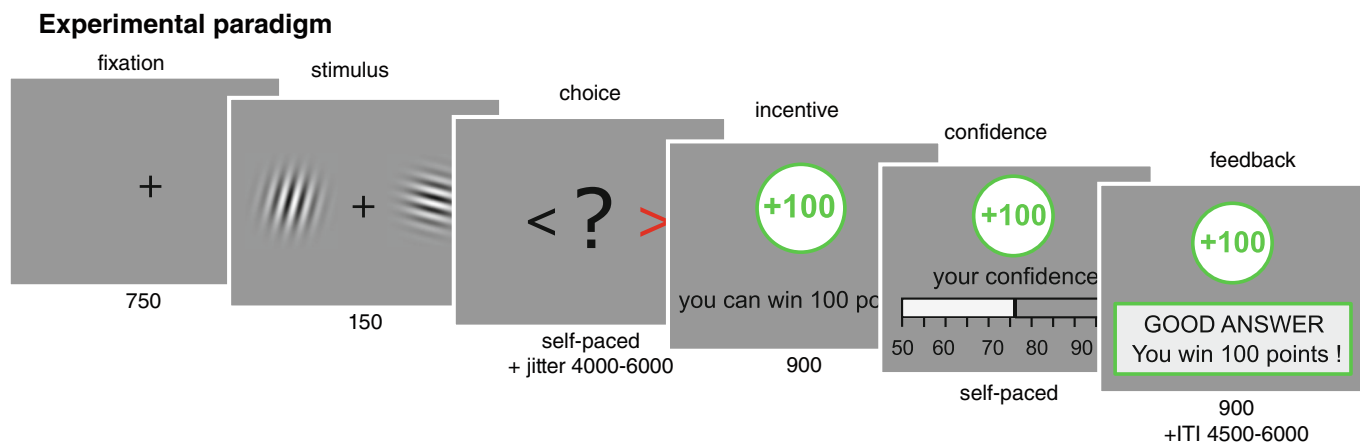
Experimental paradigm

Fig. 1 Experimental paradigm. Participants viewed two Gabor patches on both sides of the screen (150 ms) and then chose which had the highest contrast (left/right, self-paced) (for more information, see Hoven et al.[30]) After a jitter of a random interval between 4000 ms and 6000 ms, the incentive was shown (900 ms; green frame for win trials, gray frame for neutral trials, red frame for loss trials). Afterward, participants were asked to report their confidence in their choice on a rating scale ranging from 50% to 100% with steps of 5%. The initial position of the cursor was randomized between 65% and 85%. Finally, participants received feedback. The intertrial interval (ITI) had a random duration between 4500 ms and 6000 ms. The calibration session only consisted of Gabor discrimination, without confidence rating, incentives, or feedback and was used to adjust difficulty so that every individual reached a performance of 70%.

through the Department of Psychiatry at the Academic Medical Center in Amsterdam and were diagnosed with OCD.

Exclusion criteria

After applying all exclusion criteria (see Appendix S1), we included 27 patients with GD, 28 patients with OCD, and 55 HCs for behavioral analyses, of which four, two, and two participants contributed only one of two task sessions, respectively. For fMRI analyses we included 24 patients with GD, 27 patients with OCD, and 53 HCs, of which seven, three, and two participants contributed only one of two task sessions, respectively.

Experimental design and study procedure

We used a similar experimental design and study procedure as previously described.[30] For details on the experimental design and study procedure, see Hoven et al.[30] and Figure 1. In sum, participants performed a simple perceptual decision-making task, with a two-alternative forced-choice of contrast discrimination followed by a confidence judgment. In each trial, participants could either win (gain context) or lose (loss context) points, or not (neutral context), conditional on the accuracy of the choice in that trial. Importantly, this incentivization was administered after the choice moment but before the confidence rating. The task was implemented using MATLAB (The MathWorks, Inc.) and the COGENT toolbox.

Behavioral measures

We extracted trial-by-trial experimental factors including incentive condition, evidence, and behavioral measures (accuracy, confidence ratings, reaction times). Evidence was calculated by normalizing the unsigned difference of the two Gabor patches' contrast intensities by their sum to adjust for saturation effects (for more details see Lebreton et al[31]). In addition, we computed an extra *latent* variable: early certainty.

The early certainty variable was computed to analyze blood-oxygen-level-dependent (BOLD) activity at choice moment, when the brain encodes a confidence signal that is not yet biased by incentives. This was done by making a trial-by-trial prediction of early certainty based on stimulus features (reaction times, evidence, and accuracy) at choice moment. This resulted in an early certainty signal that was highly correlated with confidence, but showed no statistical

relationship with incentives (see Appendix S1). For more details, see Hoven et al.[30]

Next to confidence ratings we also assessed additional metacognitive metrics: (i) confidence calibration—the difference between average confidence and average performance as an indicator of overconfidence or underconfidence; and (ii) metacognitive sensitivity—the ability to discriminate between correct answers and errors using confidence judgments (see Appendix S1).

Behavioral analyses

All analyses were performed in the R environment (RStudio Team [2015]). We used linear mixed-effects models (LMEMs) as implemented in the lmer function from the lme4 and afex packages.[56,57] To determine P-values for the fixed effects, we performed type 3 F tests with Satterthwaite approximation for degrees of freedom, as implemented in the afex package. When relevant, we used the emmeans package to perform post hoc tests that were corrected for multiple comparisons using the Tukey method.[58]

To answer our main research questions, we built several LMEMs and performed a model selection procedure (see Table 1). The final model (model 1) included fixed effects of incentive, group, accuracy, and evidence (z scored) and interactions between incentive and group, as well as two-way and three-way interactions between evidence, accuracy, and group. Moreover, a random-subject intercept and a random slope of incentives per subject were included in the final model as well. To confirm that the incentive condition or group did not influence accuracy or reaction time, we modeled additional LMEMs with performance and reaction time as dependent variables (model 2, model 3).

Last, we added IQ (z scored) and sex as fixed effects to our original model 1 (model 4) to control for differences in the distribution of these demographic variables. Model fit was assessed and compared using chi-square tests on log-likelihood values. Additional control analyses on the properties of confidence, early certainty, confidence calibration, and metacognitive sensitivity are reported in the Appendix S1.

Because of a technical bug, our design was not fully balanced as the level of perceptual evidence was not equal across the incentive conditions. ANOVA and post hoc testing indeed showed that evidence was highest in neutral condition, followed by gain and loss. There

Table 1. Model descriptions and comparison

Model	Model notation	AIC	BIC	Comparison	Chi-square	P-value	Winning model
A	Confidence ~ incentive * group + (incentive|subject)	1 22 919	1 23 041				
B	Confidence ~ incentive * group + accuracy + (incentive|subject)	1 22 273	1 22 402	A vs B	648.59	<2.2e-16	B
C	Confidence ~ incentive * group + accuracy + evidence + (incentive|subject)	1 22 004	1 22 141	B vs C	271.00	<2.2e-16	C
D	Confidence ~ incentive * group + accuracy * evidence + (incentive|subject)	1 21 791	1 21 936	C vs D	214.53	<2.2e-16	D
E	Confidence ~ incentive * group + accuracy * evidence * group + (incentive|subject)	1 21 751	1 21 942	D vs E	52.141	1.747e-09	E
F	Confidence ~ incentive * group + accuracy * evidence * group + sex + IQ + (incentive|subject)	1 21 752	1 21 958	E vs F	2.7018	0.259	E

Shown here are the model notations of all models with their respective Akaike information criterion (AIC) and Bayesian information criterion (BIC) values, as well as model comparison outcomes with corresponding chi-square and P-values, resulting in the winning model "E," which is referred to as model 1 in the article.

were no group differences, nor an interaction between group and incentive. These effects cannot account for any group differences we find in our data, since evidence did not differ between groups. Importantly, the evidence differences did not affect performance, since performance was equal across conditions. See Appendix S1 for more details.

fMRI Analyses

For details on fMRI acquisition and preprocessing see Appendix S1 and Hoven et al.[30]

All fMRI analyses were conducted using SPM12. Critically, our design allowed us to distinguish between our two time points of interest: (i) the moment of stimulus presentation and choice, in which implicit (un)certainty about the choice is formed; and (ii) the moment of incentive presentation and confidence rating, in which the value of incentives and the confidence rating are encoded. We built a general linear model (GLM 1) estimated on subject-level with these two moments of interest: the moment of choice (i.e. stimulus presentation) and the moment of incentive presentation/confidence rating. We chose to analyze the incentive presentation and confidence rating as a single time point since the rating moment followed the presentation of the incentive after 900 ms, with regressors time-locked to the onset of incentive presentation. We also included a regressor for the moment of feedback to explain variance in neural responses related to feedback on accuracy and value that was not related to the decision-making process, but this regressor was not of interest for the current analyses. All whole-brain activation maps were thresholded using family-wise error correction (FWE) at cluster level (P FWE_clu < 0.05), with a voxel cluster-defining threshold of $P < 0.001$ uncorrected.

Using GLM 1, with regressors for choice modulated by early certainty, for incentive/rating modulated by incentive and confidence, and for feedback modulated by accuracy, we were able to investigate our contrasts of interest: (i) choice moment modulated by early certainty; (ii) incentive/rating moment modulated by incentive value; and (iii) incentive/rating moment modulated by confidence rating. For details see Appendix S1.

To study the interaction between incentive motivation and metacognitive ability on the neurobiological level we leveraged the factorial design of our task to build GLM 2. We used GLM 2 to explicate the effect of incentive motivation on both the integration of evidence at choice moment, as well as on confidence formation, and compare those between groups. GLM 2 consisted of regressors for each time point (choice and incentive/rating moments) and for each incentive condition, as well as a single regressor at feedback moment, resulting in seven regressors. For all of these events we examined both baseline activity and regression slopes relating to their pmod of interest: signed evidence for choice and confidence for incentive/rating. See Appendix S1 for more details.

Since the results by Hoven et al.[30] suggested that the VMPFC plays an important role in the interaction between incentive motivation and metacognition, we created a functional region of interest (ROI) that represented the confidence-related activity in the VMPFC cluster from our GLM 1 across group results (see Fig. 3d, Table 5). We then extracted individual t statistics within this ROI (i.e. normalized beta estimates[59]) from our contrasts of interest and performed one-sample t tests against 0 to check for positive or negative activation patterns. Then, we compared them between incentive conditions and groups, and studied their interactions using mixed ANOVAs implemented in the afex package. When appropriate, we performed post hoc testing using

Table 2. Demographics

	HCs	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	$F_{2,107} = 3.22$, $P = 0.0438$
				HC vs GD: $t(80) = 2.41$, $P = 0.0181$
				HC vs OCD: $t(81) = 0.70$, $P = 0.487$
				GD vs OCD: $t(53) = 1.79$, $P = 0.0791$
Y-BOCS***	0.25 ± 1.76	1.19 ± 2.60	20.36 ± 6.15	$F_{2,107} = 322.2$, $P < 0.001$
				HC vs GD: $t(80) = -1.01$, $P = 0.0592$
				HC vs OCD: $t(81) = -22.64$, $P < 0.001$
				GD vs OCD: $t(53) = 14.97$, $P < 0.001$
PGSI***	0.05 ± 0.40	14.85 ± 4.80	0.64 ± 1.91	$F_{2,107} = 380.5$, $P < 0.001$
				HC vs GD: $t(80) = -22.84$, $P < 0.001$
				HC vs OCD: $t(81) = -2.20$, $P = 0.0305$
				GD vs OCD: $t(53) = -14.52$, $P < 0.001$
HAMA***	1.09 ± 1.97	3.93 ± 5.88	11.43 ± 6.28	$F_{2,107} = 48.02$, $P < 0.001$
				HC vs GD: $t(80) = -3.24$, $P = 0.0017$
				HC vs OCD: $t(81) = -11.22$, $P < 0.001$
				GD vs OCD: $t(53) = 4.57$, $P < 0.001$
HDRS***	1.31 ± 2.31	5.07 ± 6.24	7.71 ± 4.04	$F_{2,107} = 24.97$, $P < 0.001$
				HC vs GD: $t(80) = -3.97$, $P < 0.001$
				HC vs OCD: $t(81) = -9.19$, $P < 0.001$
				GD vs OCD: $t(53) = 1.87$, $P = 0.0673$
Sex (men/women)***	33/22	24/3	11/17	Chi-square= 14.483, $P < 0.001$

*$P < 0.05$, ***$P < 0.001$.

Statistics for group comparisons are shown, including F and chi-square statistics, d.f., and P-values. Means ± SDs of various demographic variables are shown per group. For sex, counts are displayed.

GD, gambling disorder; HAMA, Hamilton Anxiety Rating Scale; HCs, healthy controls; HDRS, Hamilton Depression Rating Scale; OCD, obsessive–compulsive disorder; PGSI, Problem Gamblers Severity Index; Y-BOCS, Yale-Brown Obsessive Compulsive Scale.

the emmeans package, correcting for multiple comparisons using the Tukey method. Since we also hypothesized that the VS would play a role in the interaction between incentives and metacognition, we performed the same ROI analysis in the VS with a functional ROI that represented the incentive-related activity in the VS cluster from our GLM 1 across group results (see Table 5).

Results

Demographics

IQ and sex distributions differed between groups (IQ: $F_{2,107} = 3.222$, $P = 0.0438$; sex: $X = 14.483$, d.f. $= 2$ $[P < 0.001]$), with higher IQ scores for HCs compared with patients with GD ($t = 2.53$, $P = 0.014$), and with mostly men in the GD group and relatively more women in the OCD group (Table 2). This corresponds to the natural distribution observed in epidemiological studies for OCD and GD, showing a higher prevalence of GD among men and a slightly higher prevalence of OCD in women.[60–63] Age did not differ between groups. For post hoc group differences on questionnaire scores, see Appendix S1.

Behavioral results

To start, we answered our main questions: (i) are there group differences in confidence; and (ii) what is the influence of incentive motivation on confidence. Model 1 showed a main effect of group ($F_{2,112} = 4.7910$, $P = 0.01$) and incentive ($F_{2,112} = 20.9371$, $P < 0.001$) on confidence (Fig. 2, Table S3). We also found a main effect of accuracy ($F_{1,15\,107} = 608.8906$, $P < 0.001$), with individuals showing higher confidence for correct answers. Moreover, there was a significant two-way interaction of group and evidence ($F_{2,15\,099} = 3.5094$, $P = 0.02994$). As expected, we also found a significant interaction between accuracy and evidence, replicating the "X-pattern" signature of evidence integration where confidence increases with increasing evidence when correct, and vice versa ($F_{1,15\,097} = 185.3245$, $P < 0.001$).[64] Interestingly, the

evidence integration effect differed per group, as signaled by a significant three-way interaction between accuracy, evidence, and group ($F_{2,15\,094} = 3.0533$, $P = 0.04723$) (Fig. S3, Table S3; for post hoc tests see Appendix S1). Last, the interaction between incentive and group revealed a trend towards an effect ($F_{4,112} = 2.2821$, $P = 0.06487$).

Post hoc tests indicated a significantly higher confidence in patients with GD versus patients with OCD (GD-OCD $= 6.38 \pm 2.12$, Z ratio $= 3.014$; $P = 0.0073$), and a trend towards higher confidence in patients with GD compared with HCs (GD-HC $= 4.30 \pm 1.84$, Z ratio $= 2.333$; $P = 0.0513$), whereas patients with OCD did not differ from HCs. Moreover, we replicated the parametric effect of incentive value on confidence (loss-neutral $= -1.80 \pm 0.429$, Z ratio $= -4.192$ $[P < 0.001]$; loss-gain $= -3.14 \pm 0.486$, Z ratio $= -6.460$ $[P < 0.001]$; neutral-gain $= -1.34 \pm 0.363$, Z ratio $= -3.683$ $[P < 0.001]$). With regards to the three-way interaction, we found that confidence in patients with GD was less influenced by evidence for correct answers compared with both HCs and patients with OCD (see Appendix S1, Fig. S3). Exploratory post hoc analyses on the group * incentive interaction effect showed that, especially in the context of possible gains, patients with GD were more confident than patients with OCD (GD-OCD $= 8.12 \pm 2.24$, Z ratio $= 3.621$ $[P < 0.001]$) and HCs (GD-HC $= 5.83 \pm 1.95$, Z ratio $= 2.989$ $[P = 0.0079]$), with no differences between HCs and patients with OCD in any incentive condition (Table 3).

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

Since sex and IQ were significantly different between the groups, we aimed to control for these variables by adding them as fixed effects, resulting in model 4. The main effect of group did not remain
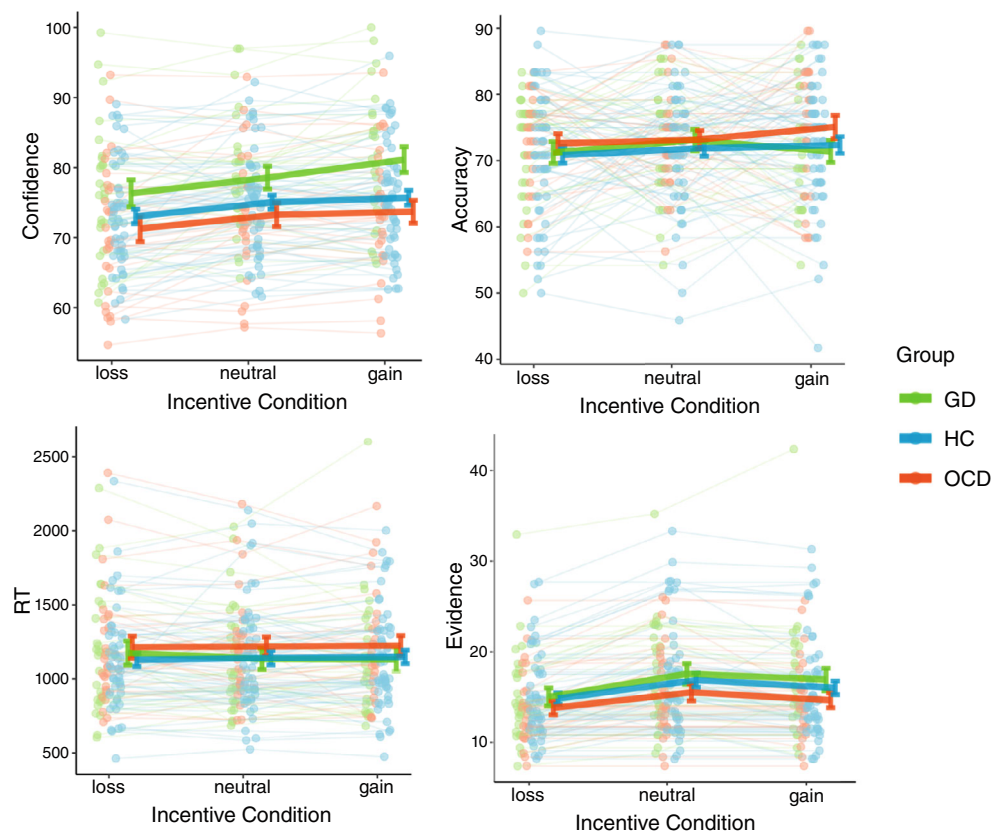
Fig. 2 Behavioral results. Individual-averaged confidence, accuracy, reaction times, and evidence as a function of incentive condition (loss, neutral, and gain) per group. Green dots and lines represent patients with gambling disorder (GD), blue dots and lines represent healthy controls (HCs), and red dots and lines represent patients with obsessive–compulsive disorder (OCD). Dots represent individuals, and lines highlight within-subject variation across conditions. Error bars represent sample mean $\pm$ SEM per group.

significant but showed a trend towards an effect ($F_{2,112} = 2.7465$, $P = 0.06846$), while the main effect of incentive did remain significant ($F_{2,112} = 20.9326$, $P < 0.001$). We found no evidence for a significant effect of sex ($F_{1,110} = 2.8776$, $P = 0.09264$) or IQ ($F_{1,109} = 0.0291$, $P = 0.86489$). The interaction effect between group and incentive remained nonsignificant at trend level ($F_{4,112} = 2.2898$, $P = 0.06412$). The significant three-way interaction between accuracy, evidence, and group persisted ($F_{2,15\,094} = 3.0871$, $P = 0.04566$). Importantly, when performing a chi-square test on the log-likelihood values of the models without and with the demographic variables to compare model fit, the model without demographics showed a better model fit ($X^2 = 2.7018$, d.f. $= 2$; $P = 0.259$), thereby favoring this simpler model. Additionally, to investigate how confidence was differently affected by sex in our HCs, we performed a two-sample t test, which showed that men were generally more confident than women (men: 76.51 ± 1.04; women: 71.70 ± 0.77) ($t_{52} = 2.6518$, $P = 0.01057$). However, both sex and IQ did not show a significant influence on confidence level in model 4.

Next to confidence, we also examined calibration and metacognitive sensitivity (see Appendix S1). In short, we showed that patients with GD were more overconfident than patients with OCD, without an effect of incentive condition. No differences in metacognitive sensitivity were found between groups or incentive conditions.

fMRI Results GLM 1

We analyzed functional neuroimaging data to test for differences in brain activity between groups for our contrasts of interest: (i) choice moment modulated by early certainty; (ii) rating/incentive moment modulated by incentive value; and (iii) rating/incentive moment modulated by confidence. The results from the fMRI group analysis revealed no significant differences between the groups for any of our contrasts.

Next, we grouped all participants together and performed one-sample t tests on our contrasts of interest to examine the results among groups (cluster-generating voxel threshold $P < 0.001$ uncorrected; clusterwise correction for multiple comparisons $P_{FWE} < 0.05$). During choice, early certainty positively correlated with activation in the precuneus, VMPFC, bilateral VS and putamen, and bilateral visual areas (Fig. 3a). The dorsal anterior cingulate cortex, bilateral dorsomedial and dorsolateral PFC, bilateral insula, thalamus, middle frontal gyrus, bilateral sensorimotor cortex, and superior and inferior parietal lobe related negatively to early certainty (Fig. 3a).

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

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

Interaction between metacognition and incentives in VMPFC (GLM 2)

Our recent study suggested an important role of the VMPFC in the interaction between incentive-processing and metacognitive signals.[30] To investigate how this interaction takes effect in and differs between our clinical groups, we performed an ROI analysis by leveraging our factorial design. We extracted VMPFC activations for both time points (choice and rating), all incentives (loss, neutral and gain), and

Table 3. Results of linear mixed-effects models

Model 1	Confidence
Incentive	$F(2.00, 112.34) = 20.94$, $P < 0.001$
Group	$F(2.00, 112.51) = 4.79$, $P = 0.010$
Accuracy	$F(1.00, 15\,107.05) = 608.89$, $P < 0.001$
Evidence	$F(1.00, 15\,104.05) = 0.04$, $P = 0.848$
Incentive:group	$F(4.00, 112.10) = 2.28$, $P = 0.065$
Accuracy:evidence	$F(1.00, 15\,097.33) = 185.32$, $P < 0.001$
Group:accuracy	$F(2.00, 15\,106.28) = 2.27$, $P = 0.103$
Group:evidence	$F(2.00, 15\,099.41) = 3.51$, $P = 0.030$
Group:accuracy: evidence	$F(2.00, 15\,094.35) = 3.05$, $P = 0.047$
Model 4	Confidence
Incentive	$F(2.00, 112.34) = 20.93$, $P < 0.001$
Group	$F(2.00, 112.50) = 2.75$, $P = 0.068$
Sex	$F(1.00, 110.26) = 2.88$, $P = 0.093$
IQ	$F(1.00, 109.80) = 0.03$, $P = 0.865$
Accuracy	$F(1.00, 15\,107.01) = 609.14$, $P < 0.001$
Evidence	$F(1.00, 15\,104.51) = 0.04$, $P = 0.845$
Incentive:group	$F(4.00, 112.11) = 2.29$, $P = 0.064$
Accuracy:evidence	$F(1.00, 15\,097.16) = 185.42$, $P < 0.001$
Group:accuracy	$F(2.00, 15\,106.06) = 2.30$, $P = 0.100$
Group:evidence	$F(2.00, 15\,098.91) = 3.45$, $P = 0.032$
Group:accuracy: evidence	$F(2.00, 15\,094.15) = 3.09$, $P = 0.046$

Shown are the results of model 1 (without demographics) and model 4 (with demographics) acquired using type 3 F tests with Satterthwaite approximation for d.f. using the afex package. Shown are F values, with corresponding d.f. and P-values.

Table 4. Results of control models

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: reaction time ~ 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	$F_{4,219} = 0.4269$, $P = 0.7891$

Shown here are the results of model 2 and model 3 linear mixed-effects models, acquired using type 3 F tests with Satterthwaite approximation for d.f. using the afex package. Shown are F values, with corresponding d.f. and P-values.
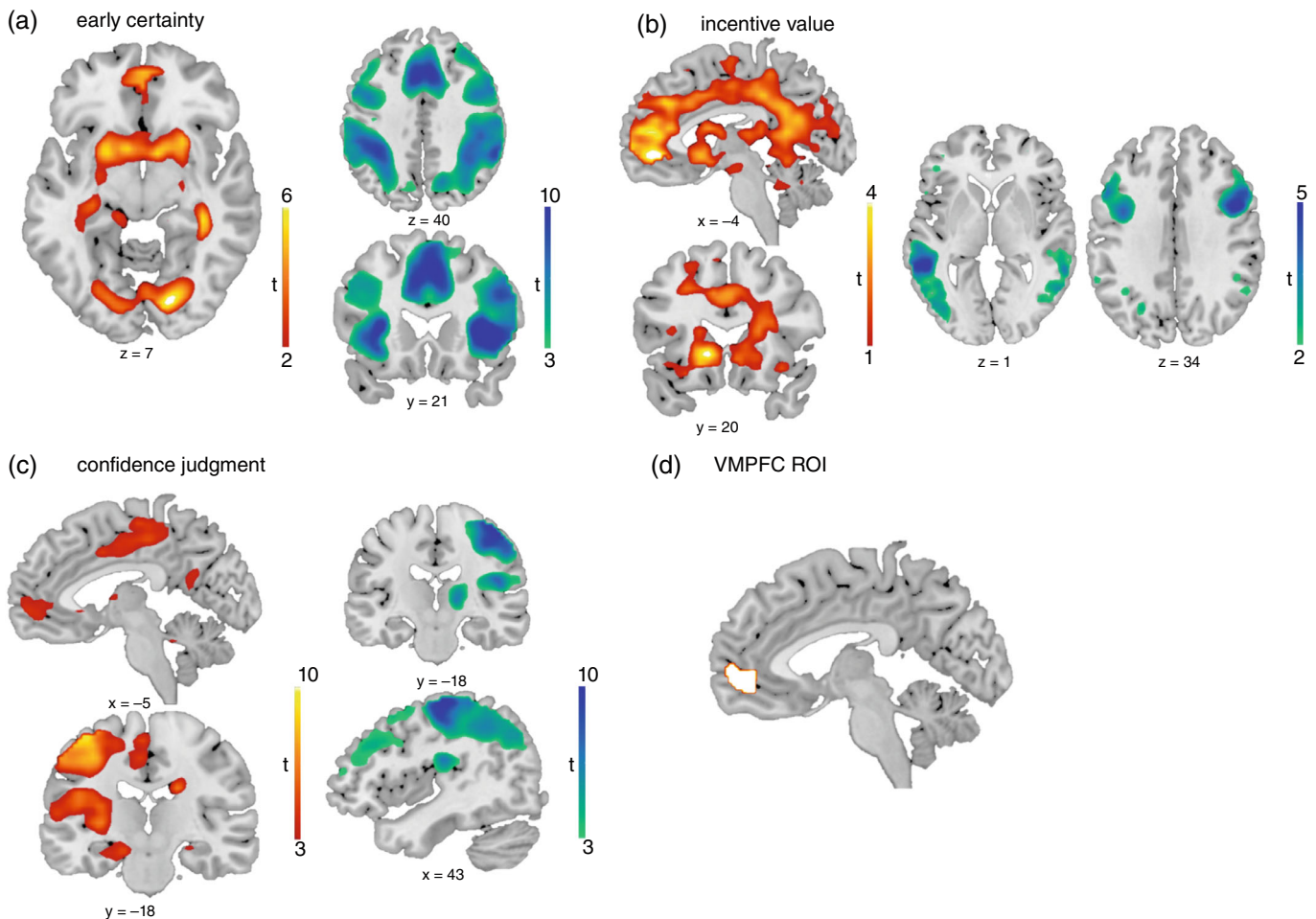
Fig. 3 Whole-brain statistical bold-oxygen-level-dependent activity among groups. Red/yellow areas represent areas with a positive relationship to the variable of interest, while green/blue areas represent areas that have a negative relationship. (a) Areas correlating significantly with early certainty at choice moment. Shown are positive activations in ventromedial prefrontal cortex (VMPFC), ventral striatum (VS), and visual cortices. Negative activations in dorsal anterior cingulate cortex, dorsolateral prefrontal cortices, insula, and parietal cortices. (b) Areas correlating significantly with incentive value at incentive/rating moment. Shown are positive activations in VMPFC, anterior cingulate cortex, and VS. 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 VMPFC, motor cortex, and putamen. Negative clusters in motor cortex and dorsolateral prefrontal cortex. All clusters survived $P < 0.05$ family-wise error correction cluster correction. Voxel-wise cluster-defining threshold was set at $P < 0.001$, uncorrected. For whole-brain activation see Table 5. (d) Region of interest (ROI) of the VMPFC used for general linear model 2 analyses.

all groups (HC, OCD, and GD), for both baseline activity and a regression slope with: (i) signed evidence; and (ii) confidence judgments (see Fig. 3d for the ROI).

First, one-sample t tests showed that, overall, VMPFC baseline activations were negative at choice and rating moment (choice: $t_{100} = -3.611$, $P < 0.001$; baseline: $t_{100} = -4.9287$, $P < 0.001$). The correlations between VMPFC activity and both signed evidence at choice moment and confidence at rating moment, however, were significantly positive (choice: $t_{100} = 3.057$, $P = 0.003$; baseline: $t_{100} = 3.7399$, $P < 0.001$) (Fig. 4). This implies that the VMPFC represents both confidence judgments and signed evidence (i.e. interaction between accuracy and evidence: increased VMPFC activity with increased evidence when correct and vice versa).

Then, we investigated whether there were effects of incentive condition and group around this general signal. As expected, at choice moment there were no effects of incentive condition on VMPFC baseline activity, nor on its correlation with the signed evidence signal (i.e. slope) (Fig. 4, Table 6). Despite the behavioral group effect on evidence integration, we did not find a group nor an interaction effect on both baseline VMPFC activity and the correlation with signed evidence. At rating moment, however, incentive

condition had a significant effect on both the baseline VMPFC activity, as well as its correlation with confidence. Post hoc testing showed that the baseline VMPFC activity was higher during gain versus loss (t_{196}: -3.874, $P < 0.001$) and during gain versus neutral ($t_{196} = -3.228$, $P < 0.001$), but no differences between neutral and loss conditions were found ($t_{196} = -0.646$, $P = 0.7948$). The correlation of VMPFC activity with confidence was significantly higher (i.e. increased slope) in gain versus neutral ($t_{196} = -3.053$, $P = 0.0072$), while no differences between gain and loss, or between neutral and loss were found. Moreover, there was a significant group effect on VMPFC baseline activity during rating moment. The post hoc tests revealed that patients with OCD had significantly decreased activity compared with HCs, averaged over incentive conditions ($t_{98} = -2.515$, $P = 0.0358$). No interaction effects between group and incentive were found on baseline activity or its correlation with confidence at rating moment.

Similar analyses using an ROI of the VS were performed (see Appendix S1), with similar results: VS activity correlated with signed evidence, but no incentive, group, or interaction effects were found at choice moment. Similarly, the correlation of VS activity with confidence was significantly higher in gain versus neutral, with no group difference at rating moment.

Table 5. Whole-brain activation tables

Effect	Brain region	k	Peak z score	P (FWE cluster-corrected)	Peak MNI x	y	z	Hemisphere
Early certainty +	Precuneus	2180	6.66	<0.001	−6	−34	11	LR
	Ventromedial PFC							
	VS							
	Putamen							
	Lingual gyrus (visual cortex)	154	6.39	<0.001	18	−81	−4	R
	Lingual gyrus (visual cortex)	54	4.49	0.045	−21	−79	−4	L
Early certainty −	Dorsal anterior cingulate	13 299	Inf (>8)	<0.001	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	<0.001	−30	−91	−4	L
	Middle temporal gyrus				−48	−67	−1	
	Lateral occipitotemporal gyrus				−45	−61	−10	
	Right cerebellum	144	6.64	<0.001	33	−55	−31	R
Incentive value +	VS	74	4.75	0.004	−12	11	−4	L
	VMPFC	212	4.53	<0.001	−3	44	−4	LR
	Dorsomedial PFC				−9	50	−4	
					0	35	14	
	Superior temporal gyrus	48	4.25	0.026	−45	−16	−1	L
					−39	−22	5	
Incentive value −	Precentral gyrus stretching into premotor cortex and dorsolateral PFC	283	5.81	<0.001	39	11	26	R
					45	5	32	
					48	14	29	
	Middle temporal gyrus	277	5.26	<0.001	−54	−43	2	L
	Superior temporal gyrus				−51	−52	11	
					−48	−25	−7	
	Lateral occipitotemporal gyrus	183	5.06	<0.001	−45	−61	−13	L
	Medial occipitotemporal gyrus				−24	−73	−7	
					−24	−82	−10	
	Middle frontal gyrus	299	4.93	<0.001	−45	2	53	L
	Inferior frontal gyrus				−39	17	23	
					−54	17	14	
	Lateral occipitotemporal gyrus	116	4.90	<0.001	42	−58	−13	R
					45	−49	−13	
	Middle temporal gyrus	47	3.74	0.029	57	−46	11	R
					60	−46	2	
					57	−61	2	
Confidence +	Middle occipitotemporal gyrus	1947	Inf (>8)	<0.001	12	−73	−10	R
	Lateral occipitotemporal gyrus				21	−70	−7	
	Cerebellum				15	−52	−16	
	Motor cortex (precentral gyrus)	993	Inf (>8)	<0.001	−36	−25	65	L
					−36	−19	47	
					−54	−16	47	
	Putamen	968	5.91	<0.001	−30	−19	2	L
	Rolandic operculum				−45	−16	20	
					−30	−22	14	
	Occipital lobe	65	4.58	0.011	42	−67	5	R

Table 5. (*Continued*)

Effect	Brain region	k	Peak z score	P (FWE cluster-corrected)	Peak MNI x	y	z	Hemisphere
	VMPFC	92	4.39	0.002	−3	56	−4	LR
					−12	47	8	
					−19	41	−1	
Confidence −	Lingual gyrus (visual cortex)	1144	Inf (>8)	<0.001	−9	−79	−7	L
	Cerebellum				−15	−52	−22	
					−24	−67	−28	
	Motor cortex (precentral gyrus)	2421	Inf (>8)	<0.001	45	−16	59	R
	Stretching into dorsolateral PFC				42	−37	62	
					39	−52	41	
	Supramarginal gyrus	262	6.92	<0.001	45	−19	20	R
	Thalamus				15	−22	2	
	Superior parietal lobe	168	5.09	<0.001	−33	−58	41	L
	Inferior parietal lobe				−39	−52	47	
					−39	−43	41	
	Middle frontal gyrus (dorsolateral PFC)	71	4.49	0.007	−45	32	32	R
					−45	23	35	

Brain activations (whole-brain analyses) showing activity related to early certainty at choice moment, as well as activity related to incentive and confidence at incentive/rating moment. All whole-brain activation maps were thresholded using family-wise error correction (FWE) for multiple correction at cluster level (*P* FWE_clu <0.05), with a voxel cluster-defining threshold of *P* < 0.001, uncorrected. Activity that positively correlated with given variable is denoted by +, whereas negative correlations are denoted by −.

Inf, infinite; L, Left; PFC, prefrontal cortex; R, Right; VMPFC, ventromedial prefrontal cortex; VS, ventral striatum.
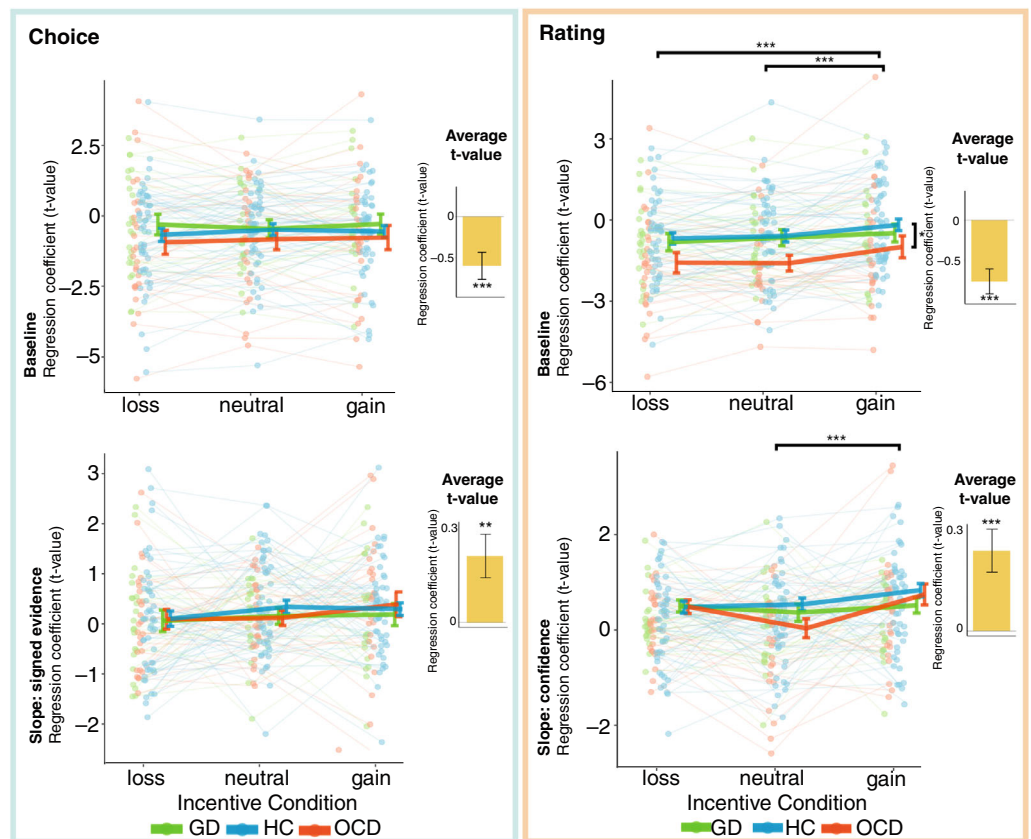

Fig. 4 Ventromedial prefrontal cortex region of interest (ROI) analysis. *t* Values corresponding to baseline and regression slopes were extracted for all three groups and three incentive conditions at two time points of interest: choice and incentive/rating moment. Green dots and lines represent patients with gambling disorder (GD), blue dots and lines represent healthy controls (HCs), and red dots and lines represent patients with obsessive–compulsive disorder (OCD). Dots represent individual *t* statistics, and error bars represent sample mean ± 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).

Table 6. Results of VMPFC ROI analysis

	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 "Signed evidence"	$F(1.99, 195) = 1.15, P = 0.320$	$F(2, 98) = 0.20, P = 0.819$	$F(3.98, 195) = 0.31, P = 0.869$
Rating baseline	$F(1.91, 186.81) = 8.61, P < \mathbf{0.001}$	$F(2, 98) = 3.24, P = \mathbf{0.044}$	$F(3.81, 186.81) = 0.44, P = 0.771$
Rating slope "Confidence judgment"	$F(1.92, 187.68) = 4.67, P = \mathbf{0.012}$	$F(2, 98) = 0.99, P = 0.375$	$F(3.83, 187.68) = 1.29, P = 0.277$

Shown are the results of the mixed ANOVAs of t statistics in the ventromedial prefrontal cortex (VMPFC) region of interest (ROI) using the afex package. Shown are the main effects of incentive condition, group, and their interaction effect on the choice and rating time points, focusing on both the baseline activity as well as the slope of signed evidence and confidence judgments, respectively. F values with corresponding d.f. and P-values are reported.

Discussion

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

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

Notably, our patient groups seemed to be situated on opposite sides of the confidence spectrum, with patients with GD being more confident than patients with OCD. However, this effect was partly driven by sex and IQ differences between groups. The GD group consisted mostly of men, whereas the OCD group had a more mixed composition, mirroring the prevalence distribution of these disorders.[69–72] Consistent with our findings of increased confidence in male HCs, recent studies have shown that men are more confident than women, despite equal performance.[73,74] Therefore, the effect of sex might have explained some variance in our data but does not fully explain the group differences, since we do find a trend towards a group effect. The importance of taking into account sex and gender as factors in both neuroscience and psychiatry research is increasingly recognized and acted upon,[75] since sex differences play a role in the incidence, treatment, and manifestation of psychopathology.[76,77] The precise role of sex and gender in metacognition deserves more attention and should be characterized further in future research.

Our data show no convincing evidence for an exaggerated decrease/increase in confidence during loss/gain anticipation in OCD/GD, respectively. However, the group * incentive interaction approached significance, with increased confidence in patients with GD compared with both patients with OCD and HCs, specifically in the gain condition. This finding is in agreement with the literature demonstrating increased reward sensitivity in GD.[78,79] Confidence in patients with OCD has been mostly studied using metamemory paradigms, and abnormalities were most profound in OCD-relevant contexts.[80–85] Earlier studies probing confidence in GD are sparse, and while they all did show an effect of overconfidence in (sub)clinical problem gamblers, none of the studies actively controlled for performance differences, making it difficult to draw strong conclusions about confidence biases.[16,17,86]

Since confidence in GD and OCD did not differ from the healthy population, we cannot technically speak of confidence "abnormalities" in GD and OCD. Future work is necessary to study the link between compulsivity and confidence more directly. One interesting method is transdiagnostic research to study metacognition in psychiatry. Transdiagnostic research methods are useful, since (meta)cognition might relate more closely to symptoms than diagnoses, because of high levels of comorbidity and heterogeneity of symptoms within disorders. Indeed, a transdiagnostic factor of "anxious-depression" was negatively related to confidence, whereas "compulsive behavior and intrusive thoughts" were positively related to confidence and showed decoupling of confidence and behavior by diminished utilization of perceptual evidence for confidence judgments.[87] This latter result is in line with our findings of diminished evidence integration into confidence judgments in patients with GD.

The brain areas we found to be related to confidence and incentive processing converge with earlier work. Confidence was found to be positively related to the VMPFC via automatic processing at the choice moment.[20,46,47,55] Early certainty processing was also positively related to activity in the VS and precuneus.[39,49,51] We also observed a widespread network of areas negatively related to early certainty, containing the dACC, dorsolateral PFC, insula, inferior parietal lobe, and midfrontal gyrus, a network repeatedly associated with uncertainty and metacognitive processes.[39,44,45,51] Also, well-known relationships between reward processing and activity in both VS and VMPFC[21,22] were replicated. Moreover, we found negative relationships between incentive value and BOLD activity in the central executive network (i.e. lateral PFC and middle frontal gyrus), as well as superior temporal gyrus.[88,89] Confidence was found to be related to VMPFC activity, not only at the choice moment but also during rating.[20,46,47] Overall, our fMRI findings closely resemble activation patterns previously shown in healthy populations.

We also replicated the effect of incentive condition on VMPFC baseline activity and on the correlation of VMPFC activity with confidence, which was highest in gain conditions and also found in the VS.[30] While we found aberrant evidence integration in patients with GD on a behavioral level, we did not find any group differences in evidence processing on the neurobiological level. Interestingly, patients with OCD showed decreased baseline VMPFC activity during the incentive/rating moment, which fits with earlier work showing neurobiological deficits in a "ventral motivational circuit" including

the VMPFC.[90,91] However, we did not find any interactions with incentive condition in the VMPFC activity related to either signed evidence or confidence.

In sum, contrary to our hypotheses, we did not find neurobiological deficits directly related to confidence or to the effects of incentive on confidence in our clinical samples. This might not be surprising, given that the behavioral group effects were small (and disappeared when controlling for demographics), which limited our ability a priori to find impairments in neural circuits mediating confidence processes. Because, to our knowledge, the present study represents the first attempt in investigating the joint neural basis of metacognitive and reward processes in both GD and OCD, further study – e.g. looking into transdiagnostic variations of symptoms – might be more powerful in detecting clinically useful neurocognitive signatures of those processes than the present clinical case-control comparisons.[92]

Acknowledgments

Disclosure statement

None of the authors have any conflicts of interest to declare.

Author contributions

Conception and design of the study: ML, RJvH, JL; acquisition of the data: MH, NSdB; analysis of data: MH; drafting of the article and figures: MH; and review and editing of article and figures: AG, DD, ML, NSdB, RJvH, and JL.

References

1. Luigjes J, Lorenzetti V, de Haan S *et al*. Defining compulsive behavior. *Neuropsychol. Rev.* 2019; **29**: 4–13.
2. Stein DJ. Obsessive–compulsive disorder. *Lancet* 2002; **360**: 397–405.
3. van Timmeren T, Daams JG, van Holst RJ, Goudriaan AE. Compulsivity-related neurocognitive performance deficits in gambling disorder: A systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* 2018; **84**: 204–217.
4. Chamberlain SR, Blackwell AD, Fineberg NA, Robbins TW, Sahakian BJ. The neuropsychology of obsessive compulsive disorder: The importance of failures in cognitive and behavioural inhibition as candidate endophenotypic markers. *Neurosci. Biobehav. Rev.* 2005; **29**: 399–419.
5. Figee M, Pattij T, Willuhn I *et al*. Compulsivity in obsessive–compulsive disorder and addictions. *Eur. Neuropsychopharmacol.* 2016; **26**: 856–868.
6. Fleming SM, Dolan RJ, Frith CD. Metacognition: Computation, biology and function. *Philos. Trans. R. Soc. B Biol. Sci.* 2012; **367**: 1280–1286.
7. Pouget A, Drugowitsch J, Kepecs A. Confidence and certainty: Distinct probabilistic quantities for different goals. *Nat. Neurosci.* 2016; **19**: 366–374.
8. Meyniel F, Sigman M, Mainen ZF. Perspective confidence as Bayesian probability: From neural origins to behavior. *Neuron* 2015; **88**: 78–92.
9. Folke T, Jacobsen C, Fleming SM, De Martino B. Explicit representation of confidence informs future value-based decisions. *Nat. Hum. Behav.* 2017; **1**: 1–8.
10. Samaha J, Switzky M, Postle BR. Confidence boosts serial dependence in orientation estimation. *J. Vis.* 2019; **19**: 25.
11. Hoven M, Lebreton M, Engelmann JB, Denys D, Luigjes J, van Holst RJ. Abnormalities of confidence in psychiatry: An overview and future perspectives. *Transl. Psychiatry* 2019; **9**: 1–18.
12. Fortune EE, Goodie AS. Cognitive distortions as a component and treatment focus of pathological gambling: A review. *Psychol. Addict. Behav.* 2012; **26**: 298–310.
13. Goodie AS, Fortune EE. Measuring cognitive distortions in pathological gambling: Review and meta-analyses. *Psychol. Addict. Behav.* 2013; **27**: 730–743.
14. Nestadt G, Kamath V, Maher BS *et al*. Doubt and the decision-making process in obsessive–compulsive disorder. *Med. Hypotheses* 2016; **96**: 1–4.
15. Samuels J, Bienvenu OJ, Krasnow J *et al*. An investigation of doubt in obsessive–compulsive disorder. *Compr. Psychiatry* 2017; **75**: 117–124.
16. Goodie AS. The role of perceived control and overconfidence in pathological gambling. *J. Gambl. Stud.* 2005; **21**: 481–502.
17. Lakey CE, Goodie AS, Campbell WK. Frequent card playing and pathological gambling: The utility of the Georgia gambling task and Iowa gambling task for predicting pathology. *J. Gambl. Stud.* 2007; **23**: 285–297.
18. Pessoa L, Engelmann JB. Embedding reward signals into perception and cognition. *Front. Neurosci.* 2010; **4**. https://doi.org/10.3389/fnins.2010.00017.
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.
20. Lopez-Persem A, Bastin J, Petton M *et al*. Four core properties of the human brain valuation system demonstrated in intracranial signals. *Nat. Neurosci.* 2020; **23**: 664–675.
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.
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.
23. Pessiglione M, Lebreton M. From the Reward Circuit to the Valuation System: How the Brain Motivates Behavior. In: Gendolla G, Tops M, Koole S. eds. *Handbook of Biobehavioral Approaches to Self-Regulation*. Springer; 2015; 157–173.
24. Clark L, Boileau I, Zack M. Neuroimaging of reward mechanisms in gambling disorder: An integrative review. *Mol. Psychiatry* 2019; **24**: 674–693.
25. Shephard E, Stern ER, van den Heuvel OA *et al*. Toward a neurocircuit-based taxonomy to guide treatment of obsessive–compulsive disorder. *Mol. Psychiatry* 2021; **26**: 4583–4604.
26. Allen M, Frank D, Samuel Schwarzkopf D *et al*. Unexpected arousal modulates the influence of sensory noise on confidence. *Elife* 2016; **5**: 1–17.
27. Koellinger P, Treffers T. Joy leads to overconfidence, and a simple countermeasure. *PLoS One* 2015; **10**: 1–22.
28. Massoni S. Emotion as a boost to metacognition: How worry enhances the quality of confidence. *Conscious. Cogn.* 2014; **29**: 189–198.
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.
30. Hoven M, Brunner G, De Boer NS *et al*. Motivational signals disrupt metacognitive signals in the human ventromedial prefrontal cortex. *Commun. Biol.* 2022; **5**: 1–13.
31. Lebreton M, Langdon S, Slieker MJ *et al*. Two sides of the same coin: Monetary incentives concurrently improve and bias confidence judgments. *Sci. Adv.* 2018; **4**: eaaq0668.
32. Lebreton M, Bacily K, Palminteri S, Engelmann JB. Contextual influence on confidence judgments in human reinforcement learning. *PLoS Comput. Biol.* 2019; **15**: e1006973.
33. Bey K, Lennertz L, Riesel A *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.
34. Bey K, Weinhold L, Grützmann R *et al*. The polygenic risk for obsessive–compulsive disorder is associated with the personality trait harm avoidance. *Acta Psychiatr. Scand.* 2020; **142**: 326–336.
35. Summerfeldt LJ, Kloosterman PH, Antony MM, Swinson RP. Examining an obsessive–compulsive core dimensions model: Structural validity of harm avoidance and incompleteness. *J. Obsessive Compuls. Relat. Disord.* 2014; **3**: 83–94.
36. Baird B, Smallwood J, Gorgolewski KJ, Margulies DS. Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *J. Neurosci.* 2013; **33**: 16657–16665.
37. Allen M, Glen JC, Müllensiefen D *et al*. Metacognitive ability correlates with hippocampal and prefrontal microstructure. *Neuroimage* 2017; **149**: 415–423.

38. Hilgenstock R, Weiss T, Witte OW. You'd better think twice: Post-decision perceptual confidence. *Neuroimage* 2014; **99**: 323–331.

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

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

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

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

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

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

45. Morales J, Lau H, Fleming SM. Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *J. Neurosci.* 2018; **38**: 2360–2317.

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

47. Lebreton M, Abitbol R, Daunizeau J, Pessiglione M. Automatic integration of confidence in the brain valuation signal. *Nat. Neurosci.* 2015; **18**: 1159–1167.

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

49. Rouault M, Mcwilliams A, Allen MG, Fleming SM. Human metacognition across domains: Insights from individual differences and neuroimaging. *Personal. Neurosci.* 2018; **1**: 17.

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

51. Hebart MN, Schriever Y, Donner TH, Haynes J-D. The relationship between perceptual decision variables and confidence in the human brain. *Cereb. Cortex* 2016; **26**: 118–130.

52. Haber SN, Knutson B. The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology* 2010; **35**: 4–26.

53. Haber SN, Behrens TEJ. The neural network underlying incentive-based learning: Implications for interpreting circuit disruptions in psychiatric disorders. *Neuron* 2014; **83**: 1019–1039.

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

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

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

57. Singmann H, Bolker B, Westfall J, Aust F, Ben-Shachar MS. afex: Analysis of factorial experiments. R package version 0.13–145. 2015.

58. Lenth R, Singmann H, Love J, Buerkner P, Herve M. Emmeans: Estimated marginal means, aka least-squares means. *R. Packag.* 2018; **1**: 3. https://doi.org/10.1080/00031305.1980.10483031.

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

60. Black DW, Shaw M. The epidemiology of gambling disorder. *Gambl. Disord.* 2019; **1**: 29–48.

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

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

63. Mathes BM, Morabito DM, Schmidt NB. Epidemiological and clinical gender differences in OCD. *Curr. Psychiatry Rep.* 2019; **21**: 1–7.

64. Sanders JI, Hangya BB, Kepecs A. Signatures of a statistical computation in the human sense of confidence. *Neuron* 2016; **90**: 499–506.

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

66. Ledgerwood DM, Dyshniku F, McCarthy JE, Ostojic-Aitkens D, Forfitt J, Rumble SC. Gambling-related cognitive distortions in residential treatment for gambling disorder. *J. Gambl. Stud.* 2020; **36**: 669–683.

67. Mallorquí-Bagué N, Vintró-Alcaraz C, Verdejo-García A *et al*. Impulsivity and cognitive distortions in different clinical phenotypes of gambling disorder: Profiles and longitudinal prediction of treatment outcomes. *Eur. Psychiatry* 2019; **61**: 9–16.

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

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

70. Subramaniam M, Wang P, Soh P *et al*. Prevalence and determinants of gambling disorder among older adults: A systematic review. *Addict. Behav.* 2015; **41**: 199–209.

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

72. Swedo SE, Rapoport JL, Leonard H, Lenane M, Cheslow D. Obsessive–compulsive disorder in children and adolescents: Clinical phenomenology of 70 consecutive cases. *Arch. Gen. Psychiatry* 1989; **46**: 335–341.

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

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

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

76. Cosgrove KP, Mazure CM, Staley JK. Evolving knowledge of sex differences in brain structure, function, and chemistry. *Biol. Psychiatry* 2007; **62**: 847–855.

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

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

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

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

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

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

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

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

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

86. Brevers D, Cleeremans A, Bechara A *et al*. Impaired metacognitive capacities in individuals with problem gambling. *J. Gambl. Stud.* 2014; **30**: 141–152.

87. Seow TXF, Gillan CM. Transdiagnostic phenotyping reveals a host of metacognitive deficits implicated in compulsivity. *Sci. Rep.* 2020; **10**: 1–11.

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

89. Wilson RP, Colizzi M, Bossong MG *et al*. The neural substrate of reward anticipation in health: A meta-analysis of fMRI findings in the monetary incentive delay task. *Neuropsychol. Rev.* 2018; **28**: 496–506.

90. Thorsen AL, Hagland P, Radua J *et al*. Emotional processing in obsessive–compulsive disorder: A systematic review and meta-analysis of 25 functional neuroimaging studies. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 2018; **3**: 563–571.

91. Stein DJ, Costa DLC, Lochner C *et al*. Obsessive–compulsive disorder. *Nat. Rev. Dis. Prim.* 2019; **51**: 1–21.

92. Parkes L, Tiego J, Aquino K *et al*. Transdiagnostic variations in impulsivity and compulsivity in obsessive–compulsive disorder and gambling disorder correlate with effective connectivity in cortical-striatal-thalamic-cortical circuits. *Neuroimage* 2019; **202**: 1–10. https://doi.org/10.1016/j.neuroimage.2019.116070.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Supporting Information.