Impaired Flexible Reward-Based Decision-Making in Binge Eating Disorder: Evidence from Computational Modeling and Functional Neuroimaging

Andrea MF Reiter*,[1,2], Hans-Jochen Heinze[1,3,4], Florian Schlagenhauf[1,5] and Lorenz Deserno[1,3,5]

[1]Max Planck Fellow Group 'Cognitive and Affective Control of Behavioral Adaptation', Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; [2]Department of Psychology, TU Dresden, Dresden, Germany; [3]Department of Neurology, Otto-von-Guericke University, Magdeburg, Germany; [4]Department of Behavioral Neurology, Leibniz Institute for Neurobiology, Magdeburg, Germany; [5]Department of Psychiatry and Psychotherapy, Campus Charité Mitte, Charité - Universitätsmedizin Berlin, Berlin, Germany

Despite its clinical relevance and the recent recognition as a diagnostic category in the DSM-5, binge eating disorder (BED) has rarely been investigated from a cognitive neuroscientific perspective targeting a more precise neurocognitive profiling of the disorder. BED patients suffer from a lack of behavioral control during recurrent binge eating episodes and thus fail to adapt their behavior in the face of negative consequences, eg, high risk for obesity. To examine impairments in flexible reward-based decision-making, we exposed BED patients (n = 22) and matched healthy individuals (n = 22) to a reward-guided decision-making task during functional resonance imaging (fMRI). Performing fMRI analysis informed via computational modeling of choice behavior, we were able to identify specific signatures of altered decision-making in BED. On the behavioral level, we observed impaired behavioral adaptation in BED, which was due to enhanced switching behavior, a putative deficit in striking a balance between exploration and exploitation appropriately. This was accompanied by diminished activation related to exploratory decisions in the anterior insula/ventro-lateral prefrontal cortex. Moreover, although so-called model-free reward prediction errors remained intact, representation of ventro–medial prefrontal learning signatures, incorporating inference on unchosen options, was reduced in BED, which was associated with successful decision-making in the task. On the basis of a computational psychiatry account, the presented findings contribute to defining a neurocognitive phenotype of BED.

Neuropsychopharmacology (2017) **42**, 628–637; doi:10.1038/npp.2016.95; published online 16 November 2016

INTRODUCTION

A hallmark of binge eating disorder (BED), recently recognized in the DSM-5 (American Psychiatric Association, 2013) as a diagnostic category, is the lack of behavioral control during recurrent binge eating episodes, ie, the feeling that one cannot stop eating. As propounded by the diagnostic criteria, excessive food intake takes place despite psychological and physical consequences, such as feelings of guilt, shame or remorse, and high risk for obesity and associated maladies. To put it differently, patients suffering from BED make disadvantageous decisions and fail to adapt their behavior in the face of negative consequences. The first few behavioral studies are in accordance with the idea of decision-making impairments in BED (Danner *et al*, 2012; Davis *et al*, 2010; Svaldi *et al*, 2010; Voon *et al*, 2015).

Healthy individuals guide their decisions by values learnt via prediction errors (PEs) generated from observed outcomes, indicating that an outcome is better or worse than expected. These PEs are defined as 'model-free' because they neglect the structure of the environment and simply lead to a repetition of previously reinforced actions, but only allow for slow adaptation in dynamic environments (Daw *et al*, 2005; Dayan and Daw, 2008). When decisions made in the past have indeed turned out to be rewarding, individuals can exploit this experience for maximal gain. As is in everyday life, environmental conditions are frequently probabilistic and dynamic, challenging the individual to explore new alternatives at the right time (Cohen *et al*, 2007; Daw *et al*, 2006; Frank *et al*, 2009; Hare, 2014). Interestingly, PEs do not only exist for options actually chosen, but can also process information on alternative choice options: this results in more complex PEs incorporating 'what might have been', that is, inference about unchosen options and their fictitious outcomes (Boorman *et al*, 2009; Bromberg-Martin *et al*, 2010; Glascher *et al*, 2009; Hampton *et al*, 2006; Lohrenz *et al*, 2007). Thus, concurrent tracking of chosen and unchosen decision options enables flexible goal-directed behavior and helps to balance exploration and exploitation.

Indeed, important recent studies have linked BED to impairments in making goal-directed decisions (Voon *et al*, 2015) and to biases towards exploratory behavior

*Correspondence: Dr AMF Reiter, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany.
Tel: +49 341 9940 2674, Fax: +49 341 9940 2221,
E-mail: reiter@cbs.mpg.de

(Morris et al, 2016). However, neural learning signatures of impaired flexible behavioral control in BED remain to be elucidated. A key brain region for flexible goal-directed behavior is the medial prefrontal cortex (mPFC). Studies in healthy participants have emphasized its important role in selecting reward goals and monitoring the value of actions (for review see Rushworth et al, 2011). In particular, the ventro–medial part of the PFC (vmPFC) is deemed responsible for on-the-fly valuation processes, which rely on the incorporation of environmental structure (Glascher et al, 2009; Hampton et al, 2006; Wunderlich et al, 2012). Intriguingly, Hampton et al (2006) and Glascher et al (2009) reported activation in the vmPFC to be associated with prediction error signals incorporating task structure during flexible decision-making. On the basis of this work, we report, to the best of our knowledge, the first task-based functional magnetic resonance imaging (fMRI) study investigating neural learning signatures of impaired flexible decision-making in BED.

To this end, we adopted a computational psychiatry approach (Montague et al, 2012; Stephan and Mathys, 2014) to investigate mechanisms of behavioral adaptation and associated neural signatures in BED by combining computational modeling with fMRI during a dynamic choice task. We aimed to elucidate the hypothesized impairment in flexible behavioral adaption of BED patients via the application of reinforcement learning (RL-) models. Regarding fMRI data, we first studied neural correlates of model-free prediction errors. Second, we investigated flexible behavioral adaptation via PEs incorporating inference about unchosen options and expected these signals, as well as between-group differences, to be associated with BOLD-activation in the vmPFC.

MATERIALS AND METHODS

Participants

Twenty-two BED patients and 22 healthy control subjects (HC) were recruited. All subjects underwent the Structured Clinical Interview for DSM IV, Axis I disorders, SCID-I (First et al, 2001). HC who were included reported no current nor past psychiatric disorders. Patients suffering from BED were diagnosed according to DSM-5 criteria by a psychologist using the German version of the structured Eating Disorder Examination Interview (Hilbert et al, 2004). As body mass index (BMI) is not a diagnostic criterion according to the DSM-5, patients were included irrespective of their BMI (Dingemans and van Furth, 2012). BED and control group did not differ significantly with respect to BMI; the average BMI corresponded to the definition of being overweight in both groups (compare Table 1). Participants who were included did not use any psychotropic medication. Owing to raw data artifacts, fMRI data sets from two participants (n = 1 BED, n = 1 HC) were excluded. For demographic and clinical characteristics, see Table 1. All participants provided written informed consent and were paid on an hourly basis. The local Ethics committee (Medical Faculty of the University of Leipzig) approved the study.

Neuropsychological Testing

Behavioral control has been shown to be linked to general cognitive capacities (Otto et al, 2013, 2015; Schad et al, 2014), which might relate to between-group differences in patient studies (Sebold et al, 2014). Thus, participants underwent neuropsychological testing in an independent session on the following domains: working memory (Digit Span, Wechsler, 1955), cognitive speed (Digit-Symbol-Substitution Test, Wechsler, 1955), reasoning (Matrices Test, Amthauer et al, 1999), verbal IQ (German vocabulary test, Schmidt and Metzler, 1992) visual attention (Reitan Trail Making A, Reitan, 1955) and executive functioning (Reitan Trail Making B, Reitan, 1955) as summarized in Table 1.

Task

During fMRI, participants performed 160 trials of a decision-making task designed to examine flexible behavioral adaptation. Participants decided between two cards distinguishable by two different abstract stimuli (Figure 1a). Maximum response time was 1.5 s. The location of the stimuli (right vs left) was randomized over trials. After the choice of one stimulus by button press, the selected card was highlighted and feedback (monetary win, 10 Eurocents vs monetary loss, crossed 10 Eurocents) was displayed for 0.5 s. During the inter-trial interval, a fixation cross was shown for a variable duration (jittered and exponentially distributed, range 1–12.5 s, mean 3.5 s). On average, trials were 4s long. If no response occurred in time, then no outcome but the message 'too slow' was presented. Mean number of missing trials was 1.14 (SD = 2.06, maximum: 9). No significant group difference on missing trials (mean$_{HC}$ = 1.05, SD$_{HC}$ = 2.08; mean$_{BED}$ 1.23, SD$_{BED}$ = 2.09; p = .56) nor on reaction times (mean$_{HC}$ = 0.650, SD$_{HC}$ = 0.08; mean$_{BED}$ = 0.614, SD$_{BED}$ = 0.08; p = 0.30) was observed.

One of the two cards had a reward probability of 80% and a punishment probability of 20% (vice versa for the other card). In this way, the task implied a simple higher-order structure: reward probabilities of the two decision options were anti-correlated; whenever card A was a good choice, card B would be a bad choice and vice versa. Even though the outcome for the unchosen option was never shown, from the experienced value of one stimulus the hypothetical value of the other stimulus ('what might have happened') could be deduced by inference on the anti-correlated task-structure. This anti-correlated task structure is similar to previous tasks based on adaptive behavioral criteria for reversals (Glascher et al, 2009; Hampton et al, 2006). The task required flexible behavioral adaptation as reward contingencies remained stable for the initial 55 trials (initial block and pre-reversal), then, changed four times after 15 or 20 trials (middle block and reversal) and became stable again for the last 35 trials (last block and post-reversal). Thus, contingency reversals were independent of participants' choices (Figure 1b), which resembles the task used by Behrens et al (2007) but without varying reward probabilities (Behrens et al, 2007).

Prior to the experiment (outside the MRI scanner), participants were instructed to opt for the card with the higher chance of winning. Depending on their choice, they could either win or lose 10 Eurocents per trial and the balance was paid to them at the end of the experiment. Participants were informed that reward probabilities might change over the course of the main experiment. These instruction slides did not provide details of reward probabilities, reversals nor of the task structure. The

Table 1 Group characteristics

	Healthy controls	Binge eating patients	Test statistic
Demographic characteristics			
Age (22/22)	M = 27.8 (SD = 4.54)	M = 29.0 (SD = 9.40)	$t = 1.27, p = 0.57$
Gender (male/female, 22/22)	7/15	6/16	$\chi^2 = 0.11 \; p = 0.74$
Handedness according to EHI (right/both/left, 22/22)	20/0/2	17/2/3	$\chi^2 = 2.44, p = 0.30$
Smokers (22/22)	8	9	$\chi^2 = 0.10. \; p = 0.76$
BMI (22/22)	M = 26.06 (SD = 4.35)	M = 28.27 (SD = 6.58),	$t = 1.31, p = 0.20$
School leaving qualification (none/compulsory basic secondary schooling/ intermediate school certificate/university entrance qualification, 22/22)	0/0/1/21	0/0/5/17	$\chi^2 = 3.09, p = 0.08$
Total years of unemployment (22/22)	M = 0.5 (SD = 1.05)	M = 0.7 (SD = 1.68)	$t = 0.55, p = 0.60$
Clinical characteristics			
Yale Food Addiction Scale (17/22)	M = 0.82 (SD = 0.63)	M = 4.68 (SD = 2.21)	$t = 3.88, p < 0.001$
Food-craving state (22/22)	M = 27.55 (SD = 12.41)	M = 41.00 (SD = 11.18)	$t = 7.78, p < 0.001$
Food-craving trait (17/22)	M = 78.82 (SD = 18.41)	M = 158.14 (SD = 34.10)	$t = 8.65, p < 0.001$
Number of objective Binge eating episodes (last 28 days) (20)		M = 8.25 (SD = 3.42)	-
Age of onset (14)		M = 18.93 (SD = 9.01)	-
Duration of disease (14)	-	M = 8.29 (SD = 7.01)	-
Neuropsychological measurements			
Reasoning (matrices) (21/22)	12.29 (SD = 2.99)	10.36 (SD = 3.55)	$t = 1.92, p = 0.06$
Working Memory (backward digit span) (21/22)	8.38 (SD = 7.52)	7.41 (SD = 2.18)	$t = 1.43, p = 0.16$
Visual attention (Trail Making A) Time (21/22)	21.74 (SD = 6.88)	27.24 (SD = 8.79)	$t = 2.27, p = 0.03$
Visual attention (Trail Making A) Errors (21/22)	0 (SD = 0)	0.05 (SD = 0.22)	-
Complex attention/task Switching (Trail Making B) Time (21/21)	46.70 (SD = 14.34)	69.46 (SD = 56.22)	$t = 1.80, p = 0.08$
Complex attention/task switching (Trail Making B) errors (21/21)	0.14 (SD = 0.36)	0.57 (SD = 1.17)	-
Cognitive speed (digit symbol substitution test) (21/22)	87.33 (SD = 14.54)	82.23 (SD = 13.41)	$t = 1.20, p = 0.24$
Verbal intelligence (German vocabulary test) (21/21)	110.29 (SD = 8.83)	102.24 (SD = 11.80)	$t = 2.50, p = 0.02$

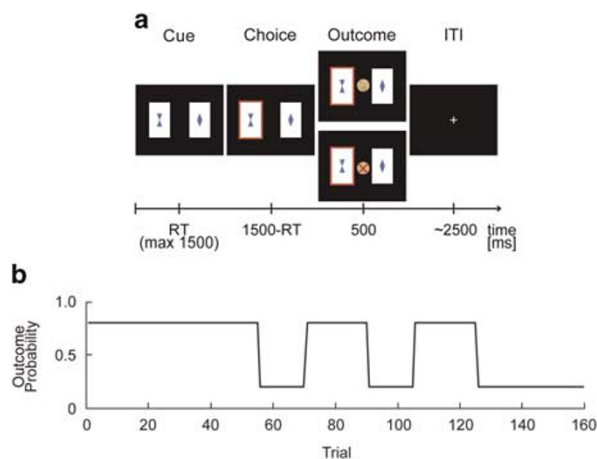Abbreviations: BMI, body mass index; EHI, Edinburgh Handedness Inventory.

Figure 1 Anti-correlated decision-making task. (a) Exemplary trial sequence. Binary choice task: participants were instructed to choose the card that they thought would lead to a monetary reward. After the participant had selected one stimulus this card was highlighted and feedback was displayed. Outcome stimuli were a 10 Eurocent coin (in the case of a win) and a crossed 10 Eurocents coin (in the case of a loss). (b) One of the stimuli had a reward probability of 80% and a punishment probability of 20% (vice versa for the other stimulus). Reward contingencies were stable for the first 55 trials (pre-reversal block) and also for the last 35 trials (post-reversal block). In the intermediate block, reward contingencies changed four times (reversal block).

Table 2 Model selection: exceedance probabilities (XP) for all models

	SU	SU-WL	DU	DU-WL	iDU	iDU-WL	Sutton K1
Overall (n = 44)	0.054	0.001	< 0.001	0.039	0.601	0.303	< 0.001
HC (n = 22)	0.035	0.005	0.003	0.112	0.321	0.522	< 0.001
BED (n = 22)	0.210	0.057	0.030	0.033	0.426	0.150	0.094

Abbreviations: SU, single update; DU, double update; iDU, inidividually weighted double update. WL indicates that the model had separate learning rates for wins and losses. For expected posterior probabilities compare Supplementary Table S2.

instruction session included 20 training trials with a different set of cards and without any reversal.

Behavioral Raw Data Analysis

First, correct choices were defined as choosing the stimulus with 80% reward probability and analyzed using repeated-measures analysis of variance (rm-ANOVA: within-subjects factor phase (pre-reversal, reversal, and post-reversal phase),

between-subjects factor group (BED *vs* HC)). Second, switching behavior as a function of the outcome in the preceding trial was analyzed using rm-ANOVA (within-subject factor outcome (win *vs* loss), between-subjects factor group). Third, we analyzed perseveration in the context of loss, defined as how often participants repeated choices for one stimulus despite two consecutive losses after having chosen this stimulus in the two preceding trials relative to all loss trials. Data analyses were performed using MATLAB R2012, IBM SPSS Statistics for Windows, Version 22 and R 3.2.0 (https://cran.r-project.org/bin/windows/base/old/3.2.0/).

Computational Modeling

We used computational modeling to analyze choices and to examine group differences in decision-making. The tested model space included four variations of RL-models. These models update expectations via PEs, which quantify the mismatch between actual outcome and prediction. Model-free PEs are only computed for chosen stimuli, although PEs can also be computed for the unchosen stimulus (Boorman *et al*, 2009; Lohrenz *et al*, 2007). Accordingly, the first three RL-models applied here differ in the degree of updating values for chosen and unchosen stimuli: (I) a model-free learner updating values for the chosen stimulus only. This neglects the anti-correlated task-structure. We refer to this model as the single-update (SU) model; (II) a learner updating values of chosen and unchosen stimuli to the same extent, thus, using inference of the task structure. We refer to this model as the full double-update (DU) model; (III) a model that individually weights the degree of double-update learning thereby accounting for inter-individual variability regarding this type of inference via the parameter κ. We name this the individually weighted DU (iDU) model. In previous studies, it was suggested that behavior in probabilistic reversal learning tasks might be explained by a RL-model with a dynamic learning rate (estimating learning rates on a trial-by-trial basis, eg, Hauser *et al*, 2014; Krugel *et al*, 2009). To test this, we additionally included the Sutton-K1 model (Sutton, 1992), which updates the learning rate dynamically as a function of the change in prediction errors.

For all the four RL-models, we translate values into actions using a Softmax rule including the parameter β, which estimates how tightly decisions are influenced by the contrast of values between the alternatives. Higher β values indicate that decisions are influenced more by relative value (low decision noise), whereas with lower β estimates, decisions are more stochastic (high decision noise). In previous studies, this was interpreted as reflecting exploitative (low decision noise) *vs* explorative (high decision noise) behavior (Cohen *et al*, 2007; Daw *et al*, 2006). In total, seven models were compared: SU, DU, iDU, each with one learning rate or separate learning rates for rewards and punishments, and the Sutton-K1 model. For equations and model fitting, see Supplementary Information and Supplementary Table S2.

Model Selection

The aim of model selection is to define a model that accounts best for the behavior in each group. Model evidences (Supplementary Information) for each model and participant were subjected to random-effects Bayesian Model Selection (BMS, spm_BMS in SPM8, www.fil.ion.ucl.ac.uk/spm/, Stephan *et al*, 2009) to determine Expected Posterior Probabilities (PP) and Exceedance Probabilities (XP) for each model. XPs describe the probability that PPs of one model exceed that of another model in the comparison set. Bayesian Model Selection was run for all subjects together and for each group separately to account for the possibility that the groups differ in best-fitting models.

Statistical Analysis of the fMRI data

See Supplementary Information for fMRI acquisition and preprocessing. We applied the general linear model approach (SPM8) for an event-related analysis. At the first level, onsets of feedback were entered into the model and modulated parametrically by two trial-by-trial regressors, which were constructed by using each individual's set of best-fitting parameters. The following regressors were computed: (1) model-free PE_{SU}: PEs for chosen values as computed on basis of the SU-Model with $\kappa = 0$; (2) more complex PE_{DU}: PEs for chosen values were computed based on the DU-Model with $\kappa = 1$. We computed the difference between $PE_{DU} - PE_{SU}$ to account for collinearity between the two regressors (for such an implementation also see Daw *et al*, 2011; Deserno *et al*, 2015a, b; Wimmer *et al*, 2012); as chosen values from the DU- and SU-algorithms differ in their degree of correlation, this difference is explicitly quantified for each individual by including the parameter κ in the iDU-algorithm. In other words, DU- and SU learning are nested in the iDU model. Thus, the difference regressor quantifies to which extent inference about alternative choices (and thus regarding the anticorrelated task structure) are incorporated in neural correlates of PEs beyond model-free PEs from the SU algorithm. Throughout the manuscript, this second parametric modulator—the difference regressor—is named PE_{DU}.

Building on the behavioral finding of elevated stochastic behavior in BED, each individual's trial-by-trial choice-probabilities from the decision-model were classified according to whether the actual choice was indeed the one predicted by the model to have the highest choice probability (exploitative) or the one with a lower choice probability (exploratory). Next, we added the onsets of cues to the first-level model of the fMRI data described ==above with binary trial-type (exploitative *vs* exploratory) as the first parametric modulator and the continuous choice probabilities as the second parametric modulator==. Compare the study by Daw *et al* (2006) for the same implementation of an analysis on exploratory *vs* exploitative trials (Daw *et al*, 2006). Onset of outcomes with PE_{SU} and PE_{DU} remained in the model to partial out their influence.

Missing trials were modeled separately. The six realignment parameters, the first temporal derivative of the translational realignment parameters and a further regressor censoring scan-to-scan movement > 1 mm were included in the analysis to account for residual effects of motion.

At the second level, the contrast images of PE_{SU} and PE_{DU} were entered to a full-factorial ANOVA with the type of PEs (PE_{SU}/PE_{DU}) as the within-subject factor and group as the between-subject factor. For contrast images regarding exploration, an independent-sample *t*-test (exploratory *vs* exploitative trials) was calculated. Results were accepted as
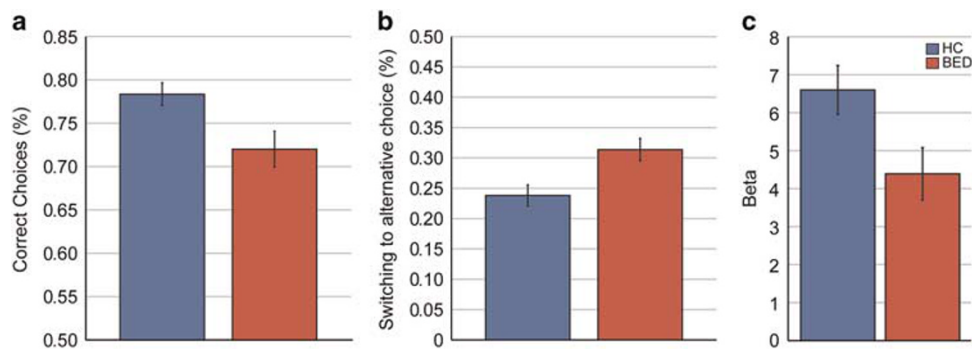
Figure 2 Behavioral results. (a) Raw data results. Correct choices differed significantly between groups ($t = 3.48$, $p = 0.001$, left panel). (b) BED patients showed enhanced switching behavior between the two stimuli ($F = 8.75$, $p = 0.005$). (c). Comparison of modeling parameters revealed that BED patients had a lower decision parameter β. Lower values of β indicate a higher degree of stochastic choices unrelated to the current choice values. Hence, lower values in BED indicate enhanced exploratory choices ($t = 2.51$, $p = 0.016$). BED, binge eating disorder; HC, healthy controls.

significant at $p < 0.05$ using family-wise-error (FWE) correction for the whole brain for task effects across all participants. Correction for multiple comparisons regarding between-group comparisons was performed in our *a priori* region of interest, the vmPFC. Therefore, an anatomical search volume was defined according to criteria described in Rushworth *et al* (2011), comprising the superior medial frontal gyrus and the medial orbitofrontal gyrus based on anatomical labeling (Tzourio-Mazoyer *et al*, 2002), truncated dorsally at MNI $z = +10$ (also compare Bartra *et al*, 2013). See Supplementary Figure S1 for details and a depiction of the anatomical search volume. A significance threshold of p-FWE < 0.05 based on this anatomical search volume was applied. As we had no regional *a priori* hypothesis for the explore-exploit fMRI analysis, we took the entire map of the whole-brain corrected F-contrast for exploration *vs* exploitation across both groups to correct between-group comparisons (Figure 4 and Supplementary Table S7) and accepted results as significant at p-FWE < 0.05 based on this functional search volume.

RESULTS

Neuropsychological Testing

We tested for differences in general cognitive capacities by subjecting results of all neuropsychological tests (Table 1) to a multivariate analysis of variance (MANOVA) with the between-subject factor group. No significant effect of group was observed ($F = 1.52$, $p = 0.19$). For an exploratory analysis of all subscales and their relationship to task performance see Table 1, Supplementary Table S1 and Supplementary Results.

Choice Behavior

Rm-ANOVA on correct choices, including the within-subject factor phase (pre-reversal, reversal, and post-reversal) and the between-subject factor group (BED *vs* HC), showed main effects of *phase* ($F_{(2,84)} = 35.97$, $p < 0.001$) and of *group* ($F_{(1,42)} = 5.72$, $p = 0.02$, Figure 2a), but no significant phase × group interaction ($F_{(1,84)} = 0.17$, $p = 0.79$). Switching behavior as a function of the outcome in the preceding trial was analyzed using rm-ANOVA (within factor outcome

(win *vs* loss), between-subject factor *group* (BED *vs* HC)). This revealed a main effect of outcome ($F_{(2,42)} = 288.93$, $p < 0.001$), and a main effect of group ($F_{(1,42)} = 8.75$, $p = 0.005$, Figure 2b), but no significant outcome × group interaction ($F_{(1,42)} = 0.11$, $p = 0.74$). Thus, irrespective of the outcome in the previous trial, BED patients switched choices more frequently. Further, an independent *t*-test did not indicate any difference between groups in repeating choices for one stimulus despite two consecutive losses after having chosen this stimulus in the two preceding trials ($\text{mean}_{BED} = 0.11 \pm .07$, $\text{mean}_{HC} = 0.10 \pm .07$, $t = 0.18$, $p = 0.86$).

Computational Modeling: Model Selection

BMS across all participants revealed that iDU-models provided the best account for observed choices peaking for iDU with one learning rate (iDU XP $= 0.60$, iDU-WL XP $= 0.30$, Table 2). Thus, we use parameters derived from this model in all subsequent analyses. When running BMS for both groups separately, iDU models clearly outperformed other models for the control group, whereas results were more ambiguous in the BED group indicating pronounced heterogeneity in this group (Table 2).

Computational Modeling: Parameter Comparison

Independent sample *t*-tests with Bonferroni-correction (adjusted p $= 0.017$) to compare the three modeling-derived parameters between groups (decision parameter β, learning rates for chosen values α_c, and learning rates for unchosen values α_{uc}, as product of κ by α_c) revealed a significant group difference for the decision parameter β ($t = 2.51$, $p = 0.016$, Figure 2c), but not for any other parameter (ts < 0.81, ps > 0.42). A lower decision parameter indicates a higher degree of stochastic choices unrelated to the current choice value, ie, lower values in BED indicate noisier decision-making. Importantly, when excluding two patients who were not fit better than chance by the model (for definition see Supplementary Information), the difference remained significant ($t = 2.14$, $p = 0.039$). Thus, the significantly lower decision parameter β did not simply result from very poor fit (random choice behavior) in the patient group.
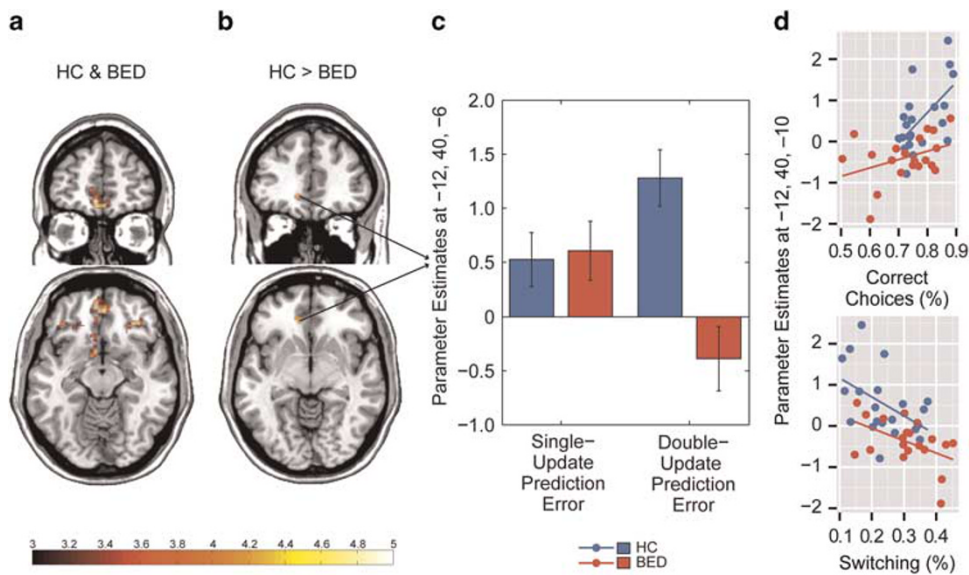
Figure 3 Neural correlates of single-update and double-update prediction error processing. (a) Across both groups, peak conjoint activity elicited by PE_{SU} and PE_{DU} was observed in the ventro-medial prefrontal cortex (vmPFC, X = -6 Y = 52 Z = -12, $t = 4.86$, p-FWE for the whole brain = 0.03, see Supplementary Table S6). (b, c) Comparing PE_{SU} and PE_{DU} between groups revealed significantly reduced activation associated with of PE_{DU} signatures in BED in the medial prefrontal cortex (X = −12, Y = 40, Z = −6, $t = 4.06$, FWE-corrected for vmPFC $p = 0.03$). For display purposes, threshold is set at $p < 0.001$, cluster level $k = 10$. (d) Parameter estimates at the peak-coordinate of the group difference in vmPFC for PE_{DU} were extracted and, for both groups separately, correlated with behavioral performance (percentage of correct choices, percentage of switching). This revealed a significant positive association between activation associated with PE_{DU} and correct choices in both, BED ($r = 0.60$, $p = 0.005$) and HC ($r = 0.53$, $p = 0.02$). The correlation between PE_{DU} related activation and switching was significant ($r = −0.35$, $p = 0.03$) and did not indicate any evidence for an interaction effect with group ($t < 1.41$, $p > 0.17$, R^2 change due to moderator < 0.03). PEs, prediction errors; BED, binge eating patients; DU, double update; HC, healthy controls; SU, single update.

Neural PE Processing: Entire Sample

We aimed to explore neural signatures of simple and more complex PE processing in BED *vs* HC. Thus, we analyzed activation associated with PEs for chosen values as a function of SU- *vs* DU-learning, that is, PEs derived from the SU-Model (PE_{SU}) *vs* PEs derived from the DU-Model (PE_{DU}), see Supplementary Table S6 and Figure 3a for results. We observed activation at p-FWE$_{wholebrain} < 0.05$ associated with PE_{SU} in the bilateral ventral striatum, as well as the ventro–medial prefrontal/orbitofrontal cortex (vmPFC/OFC), amygdala, right hippocampus, right putamen, and posterior cingulum. PE_{DU} co-varied with activation in similar regions, including the bilateral ventral striatum and vmPFC/OFC (p-FWE$_{wholebrain} < 0.05$, Supplementary Table S6). The conjunction of PE_{SU} and PE_{DU} reached significance peaking in the vmPFC/OFC (p-FWE$_{wholebrain} < 0.05$, Supplementary Table S6).

Neural PE Processing: Group Comparison

Regarding model-free PE_{SU} processing, we did not observe significant between-group differences based on the anatomical vmPFC search volume (X = −10, Y = 60, Z = −10, $t = 2.17$, p-FWE$_{vmPFC} = 0.89$). There was no significant group difference in other regions at a liberal threshold (cluster level $k = 10$, $p < 0.001$ uncorrected). To investigate between-group differences in BOLD-activation related to more complex PE signatures, we tested for a type of PE (PE_{SU}/PE_{DU}) × group interaction. On the basis of the anatomical vmPFC search volume, this interaction was significant in vmPFC (X = −12,

Y = 40, Z = −6, $t = 4.00$, p-FWE$_{vmPFC} = 0.04$). As a *post hoc* contrast, we compared PE_{SU} and PE_{DU} between groups and observed significantly reduced activation associated with PE_{DU} in BED in the vmPFC (X = −12, Y = 40, Z = −6, $t = 4.06$, p-FWE$_{vmPFC} = 0.03$, Figure 3b) but no significant differences for the other *post hoc* contrasts (p-FWE$_{vmPFC} > = 0.68$).

Next, we tested for an association of neural activation related to PE_{DU} and choice behavior: parameter estimates at the peak-coordinate of the group difference in vmPFC (X = −12, Y = 40, Z = −6) for PE_{DU} were extracted and correlated with behavioral performance (percentage of correct choices, percentage of switching) for both groups separately. One outlier (z-value of parameter estimates $< − 2.8$) in the BED group was removed beforehand. We found a significant positive association between the neural PE_{DU} signature and correct choices in BED ($r = 0.60$, $p = 0.005$) as well as in HC ($r = 0.53$, $p = 0.02$). Indeed, with group as a covariate, the correlation between PE_{DU} and correct choices was significant ($r = 0.55$, $p < 0.001$). The association of the PE_{DU} signature and switching was significantly negative in HC ($r = − 0.47$, $p = 0.03$), but non-significant in BED ($r = -0.24$, $p = 0.32$). Across both groups, when controlling for group, the negative correlation between PE_{DU} and switching was significant ($r = − 0.35$, $p = 0.03$). No moderation effect of group on the association of neural signature and behavioral performance was found ($t < 1.41$, $p > 0.17$, R^2 change due to moderator < 0.03). These findings suggest that neural activation corresponding to PE_{DU} is related to better behavioral performance and less switching in both healthy individuals and BED (Figure 3c and d).
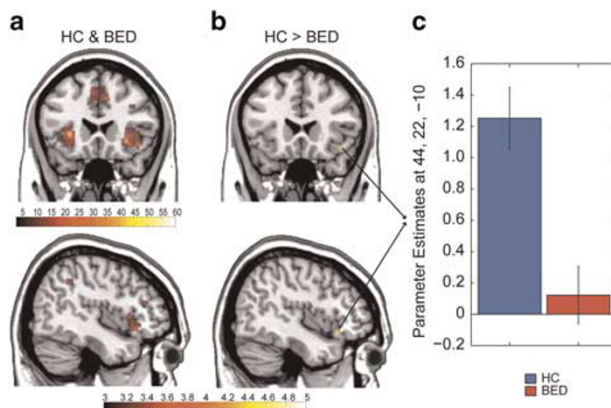
Figure 4 Neural correlates of the exploration–exploitation trade-off. (a). Across both groups, exploratory trials vs exploitative trials were associated with bilateral activation of the anterior insula/ventro–lateral prefrontal cortex (see also Supplementary Table S7). (b,c) Comparing activation in exploratory vs exploitative trials between groups demonstrated that BED patients show significantly diminished activity in the aI/vlPFC during exploratory trials ($X = 44$, $Y = 22$, $Z = -10$, $t = 3.91$, FWE-corrected for the aI/vlPFC, p-FWE < 0.05). For display purposes, threshold is set at $p < 0.001$, cluster level $k = 10$. FWE, family-wise-error.

Neural Correlates of the Exploration-Exploitation Trade-Off

Building on the observation of enhanced exploration in BEDs' choices, we compared activity elicited in exploratory vs exploitative trials using an F-contrast. This revealed peak activation in bilateral anterior insula/ventro–lateral prefrontal cortex (aI/vlPFC) and in the dorso–medial prefrontal cortex (p-FWE$_{wholebrain} < 0.05$, Figure 4a, Supplementary Table S7), due to higher activation for exploratory vs exploitative trials. A between-group effect was revealed in the right aI/vlPFC. BED patients showed significantly lower activation for exploratory trials compared to HC ($X = 44$, $Y = 22$, $Z = -10$, $t = 3.91$, p-FWE$_{main\ effect\ exploration-exploitation} < 0.05$, Figure 4b and c).

For both groups separately, we tested for an association between BOLD activation in response to exploratory trials at the peak coordinate of the between-group difference in the aI/lPFC and behavioral performance (correct choices and switching behavior). Results did not indicate any significant correlation (correlation with correct choices: $r_{HC} = 0.35$ $p_{HC} = 0.12$, $r_{BED} = -0.14$, $p_{BED} = 0.55$; correlation with switching: $r_{HC} = 0.13$ $p_{HC} = 0.58$, r_{BED} .0.17, $p_{BED} = 0.47$).

DISCUSSION

The results of the current study, which combined fMRI and computational modeling of Reinforcement Learning, provide novel insight into the neural correlates of maladaptive decision-making in BED, thereby helping to refine a neurocognitive phenotype of the newly classified disorder. We observed impaired behavioral adaptation in a dynamic environment in BED as compared with healthy controls. Whereas we found compelling evidence that healthy controls used inference on alternative choices to guide decision-making, Bayesian model selection did not reveal convincing evidence that Binge Eating patients employed this type of inference to solve the task. Relatedly, patients showed

reduced BOLD-activation associated with learning signatures incorporating alternative choice options in the vmPFC. Moreover, we found decision-making in BED to be characterized by enhanced switching between choices, indicating a bias towards exploratory decisions during behavioral adaptation in a dynamic environment. Parallel to this behavioral observation, BED was characterized by less aI/vlPFC activation during exploratory decisions.

Reduced Incorporation of Inference on Alternative Options in the Neural Correlates of Learning

According to BMS, HCs convincingly integrated inference on alternative choices into decision-making, to use 'what might have happened' when making decisions. Contrary to this, BMS did not reveal convincing evidence for this type of inference on alternative choices being dominant in BED patients. In accordance with the aforementioned results, BOLD-activation associated with PEs incorporating inference on alternative options was reduced in the vmPFC of BED patients. In healthy individuals, concurrent tracking of multiple decision possibilities and their potential consequences contributes to flexible goal-directed behavior in dynamic environments (Abe and Lee, 2011; Bromberg-Martin *et al*, 2010; Glascher *et al*, 2009; Hampton *et al*, 2006; Lohrenz *et al*, 2007; Takahashi *et al*, 2013). In the present study, vmPFC PE signatures incorporating inference on alternative options were indeed positively associated with successful choices and negatively associated with switching behavior. Thus, the specific reduction in vmPFC signaling could be one common substrate for impaired goal-directed decision-making in BED as reported previously in a behavioral study using a sequential decision-making task (Voon *et al*, 2015). In accordance with this conclusion, the latter study also found an association between impaired goal-directed behavior and reduced gray matter density in vmPFC/mOFC in BED (Voon *et al*, 2015).

Disadvantageous Switching Behavior in BED

Although clinical characteristics and diagnostic criteria suggest impaired mechanisms of flexible behavioral adaptation as crucial to BED, systematic investigations of this impairment are scarce. In this study, we observed deficits in BED in the flexible adaptation of behavior in a changing environment. Our analytic approach, including computational modeling, allowed us to differentially disentangle this deficit: while neither learning rates nor neural correlates of model-free learning differed between groups, patients suffering from BED did not exploit a relatively better option as consistently as controls but showed pronounced switching behavior. This can be regarded as an impaired balance between exploratory and exploitative choice behavior (Cohen *et al*, 2007; Daw *et al*, 2006). Although it is obviously advantageous for an individual to explore alternatives in a changing environment, this type of behavior observed in BED was accompanied by fewer correct choices, confirming that the amount of exploration was indeed suboptimal. Notably, control analyses showed that this was not owing to overall random switching behavior in BED. Accordingly, our interpretation is consistent with a very recent study, which found obese people with BED to be characterized by

enhanced exploration compared to obese people without BED (Morris et al, 2016).

In patients, this behavioral tendency to switch was paralleled by reduced activation during such exploratory choices in the anterior aI/vlPFC, key regions implicated in reversing behavior (Cools et al, 2002; Menon and Uddin, 2010). Thus, less activation in this region might hinder the individual to get back on the right track after an exploratory try that has not resulted in positive benefits. In line with this idea, prior imaging studies have also reported aI/vlPFC activation during uncertainty prediction and when making a risky compared with a safe decision (Paulus et al, 2003; Preuschoff et al, 2008; Singer et al, 2009). This notion is complemented by the interpretation that enhanced anterior aI/vlPFC activation in exploratory trials as observed in healthy controls could reflect a potential warning or uncertainty signal for these trials. In fact, in healthy subjects, uncertainty was proposed to mediate exploratory behavior (Badre et al, 2012; Daw et al, 2005, 2006; Frank et al, 2009; Kakade and Dayan, 2002) and thus, could be hypothesized to underlie switching behavior observed in BED. In this framework, aI/vlPFC activation could guide choices in moments of uncertainty, while computations of uncertainty itself appear to be more likely associated with the frontal pole (Badre et al, 2012). Therefore, the reduction of such aI/vlPFC signaling during exploratory decisions may be thought of as reduced awareness regarding the uncertain (or disadvantageous) character of these decisions. This might bias the individual toward more and suboptimal exploratory decisions (instead of selecting a relatively good option based on accumulated experience).

In summary, the latter aspects of our data suggest that patients learn similar to controls but perform suboptimally owing to enhanced switching paralleled by reduced aI/vlPFC exploration signaling. However, more heterogeneous BMS results and, most importantly, reduced coding of PEs incorporating task structure in patients' vmPFC, are indicative of a specific learning deficit as learning values via PEs incorporating task structure could result in choosing, and staying with, the most valuable option, when appropriate. These lines of reasoning might motivate future studies to systematically dissect where (and when) the observed deficits originate, eg, by testing performance in extinction (Frank et al, 2004; Gold et al, 2012).

Relevance to Addiction Theories

A hallmark of BED, the maintenance of maladaptive behaviors despite negative consequences, closely resembles key criteria of substance addiction and a current debate relates to the classification of BED as an addiction spectrum disorder (Smith and Robbins, 2013; Volkow et al, 2013). However, a noteworthy albeit debated review (Ziauddeen et al, 2012) issues caveats against a premature adoption of the 'food addiction model': the article deems functional imaging attempts to profile BED as insufficient to date and calls for task-dependent measurements based on cognitive-neuroscience models in order to relate behavioral and cognitive phenotypes to neuroimaging findings. The study at hand is one step in this direction. The adopted computational psychiatry approach enables estimation of specific parameters that provide mechanistic accounts of functioning in one or another cognitive domain (Wiecki

et al, 2015) and informs the modeling-based fMRI analysis of neural learning signatures (Stephan et al, 2015).

Limitations

Although general neuropsychological measures did not differ between-groups in a MANOVA and our between-group behavioral findings remained significant when adjusting for neurocognitive functioning, exploratory analyses (Supplementary Results) also revealed group × cognitive speed, as well as group × verbal intelligence interactions due to correlations between the cognitive measures and task performance in the patient group. This invites speculations that relatively better cognitive functioning could protect against or compensate for impaired flexible decision-making in BED. To test this hypothesis, larger samples and longitudinal designs are warranted as the cross-sectional design precludes any conclusions on causality. Future studies may also investigate the extent to which the observed deficits in BED are influenced by certain task specifics. It would be interesting to determine whether findings generalize to similar tasks with mildly correlated reward probabilities (Wimmer et al, 2012), whether changes or drifts in reward probabilities (Behrens et al, 2007; Daw et al, 2005) exacerbate deficits and how explicit presentation of forgone rewards impacts patients' behavior (Chiu et al, 2008; Li and Daw, 2011). In the current study, we identified vmPFC as an a priori region of interest. However, the finding of between-group differences in the vmPFC and aI/vlPFC associated with different decision signatures begs the question as to whether vmPFC activity or vmPFC-vlPFC interactions mediate switching behavior. Interestingly, vmPFC signals were clearly associated with behavioral performance. Lesion studies in animals and their translation to humans, eg via brain stimulation techniques, could bear on the question of which PFC regions mediate the observed alterations in switching behavior.

CONCLUSIONS

In summary, this study provides insight into specific impairments in reward-guided decision-making in BED. That is, a disadvantageous behavioral bias towards switching behavior accompanied by less activation associated with these exploratory trials in the aI/vlPFC, as well as diminished representation of PEs incorporating information about the task structure in the vmPFC. By adopting a computational psychiatry approach combined with modeling-informed fMRI analysis, this study contributes to refining the neurocognitive phenotype of BED as an addition to clinical observations and new diagnostic criteria in the DSM-5.

AUTHOR CONTRIBUTIONS

AMFR, FS, and LD designed the study. AMFR and LD performed the research. AMFR and LD analyzed the data. AMFR and LD wrote the initial version of the manuscript. AMFR, HJH, FS, and LD read and corrected versions of the manuscript.

REFERENCES

Abe H, Lee D (2011). Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron* **70**: 731–741.

American Psychiatric Association (2013). *The Diagnostic and Statistical Manual of Mental Disorder: DSM 5*. American Psychiatric Association: Washington D.C, USA.

Amthauer RBB, Liepmann D, Beauducel A (1999). *Intelligenz-Struktur-Test 2000*. Horgrefe: Göttingen, Germany.

Badre D, Doll BB, Long NM, Frank MJ (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron* **73**: 595–607.

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

Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007). Learning the value of information in an uncertain world. *Nat Neurosci* **10**: 1214–1221.

Boorman ED, Behrens TE, Woolrich MW, Rushworth MF (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* **62**: 733–743.

Bromberg-Martin ES, Matsumoto M, Hong S, Hikosaka O (2010). A pallidus-habenula-dopamine pathway signals inferred stimulus values. *J Neurophysiol* **104**: 1068–1076.

Chiu PH, Lohrenz TM, Montague PR (2008). Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. *Nat Neurosci* **11**: 514–520.

Cohen JD, McClure SM, Yu AJ (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos Trans R Soc Lond B Biol Sci* **362**: 933–942.

Cools R, Clark L, Owen AM, Robbins TW (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci* **22**: 4563–4567.

Danner UN, Ouwehand C, van Haastert NL, Hornsveld H, de Ridder DT (2012). Decision-making impairments in women with binge eating disorder in comparison with obese and normal weight women. *Eur Eat Disord Rev* **20**: e56–e62.

Davis C, Patte K, Curtis C, Reid C (2010). Immediate pleasures and future consequences. A neuropsychological study of binge eating and obesity. *Appetite* **54**: 208–213.

Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**: 1204–1215.

Daw ND, Niv Y, Dayan P (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* **8**: 1704–1711.

Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006). Cortical substrates for exploratory decisions in humans. *Nature* **441**: 876–879.

Dayan P, Daw ND (2008). Decision theory, reinforcement learning, and the brain. *Cogn Affect Behav Neurosci* **8**: 429–453.

Deserno L, Huys QJ, Boehme R, Buchert R, Heinze HJ, Grace AA et al (2015a). Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proc Natl Acad Sci USA* **112**: 1595–1600.

Deserno L, Wilbertz T, Reiter AMF, Horstmann A, Neumann J, Villringer A et al (2015b). Lateral prefrontal model-based signals are reduced in healthy individuals with high trait impulsivity. *Transl Psychiatry* **5**: e659.

Dingemans AE, van Furth EF (2012). Binge Eating Disorder psychopathology in normal weight and obese individuals. *Int J Eat Disord* **45**: 135–138.

First MB, Spitzer RL, Gibbon M, Williams J (2001). *Structured Clinical interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition With Psychotic Screen (SCID-I/P W/ PSY SCREEN)*. New York State Psychiatric Institute: New York, NY, USA.

Frank MJ, Doll BB, Oas-Terpstra J, Moreno F (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat Neurosci* **12**: 1062–1068.

Frank MJ, Seeberger LC, O'Reilly R C (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* **306**: 1940–1943.

Glascher J, Hampton AN, O'Doherty JP (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb Cortex* **19**: 483–495.

Gold JM, Waltz JA, Matveeva TM, Kasanova Z, Strauss GP, Herbener ES et al (2012). Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. *Arch Gen Psychiatry* **69**: 129–138.

Hampton AN, Bossaerts P, O'Doherty JP (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci* **26**: 8360–8367.

Hare T (2014). Neuroscience. Exploiting and exploring the options. *Science* **344**: 1446–1447.

Hauser TU, Iannaccone R, Ball J, Mathys C, Brandeis D, Walitza S et al (2014). Role of the medial prefrontal cortex in impaired decision making in juvenile attention-deficit/hyperactivity disorder. *JAMA Psychiatry* **71**: 1165–1173.

Hilbert A, Tuschen-Caffier B, Ohms M (2004). Eating disorder examination: a German version of the structured eating disorder interview. *Diagnostica* **50**: 98–106.

Kakade S, Dayan P (2002). Dopamine: generalization and bonuses. *Neural Networks* **15**: 549–559.

Krugel LK, Biele G, Mohr PN, Li SC, Heekeren HR (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proc Natl Acad Sci USA* **106**: 17951–17956.

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

Lohrenz T, McCabe K, Camerer CF, Montague PR (2007). Neural signature of fictive learning signals in a sequential investment task. *Proc Natl Acad Sci USA* **104**: 9493–9498.

Menon V, Uddin LQ (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct* **214**: 655–667.

Montague PR, Dolan RJ, Friston KJ, Dayan P (2012). Computational psychiatry. *Trends Cogn Sci* **16**: 72–80.

Morris LS, Baek K, Kundu P, Harrison NA, Frank MJ, Voon V (2016). Biases in the explore-exploit tradeoff in addictions: the

role of avoidance of uncertainty. *Neuropsychopharmacology* **41**: 940–948.

Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND (2013). Working-memory capacity protects model-based learning from stress. *Proc Natl Acad Sci USA* **110**: 20941–20946.

Otto AR, Skatova A, Madlon-Kay S, Daw ND (2015). Cognitive control predicts use of model-based reinforcement learning. *J Cogn Neurosci* **27**: 319–333.

Paulus MP, Rogalsky C, Simmons A, Feinstein JS, Stein MB (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *Neuroimage* **19**: 1439–1448.

Preuschoff K, Quartz SR, Bossaerts P (2008). Human insula activation reflects risk prediction errors as well as risk. *J Neurosci* **28**: 2745–2752.

Reitan RM (1955). The relation of the trail making test to organic brain damage. *J Consult Psychol* **19**: 393–394.

Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron* **70**: 1054–1069.

Schad DJ, Junger E, Sebold M, Garbusow M, Bernhardt N, Javadi AH *et al* (2014). Processing speed enhances model-based over model-free reinforcement learning in the presence of high working memory functioning. *Front Psychol* **5**: 1450.

Schmidt K-H, Metzler P (1992). *Wortschatztest (WST)*. Beltz Test GmbH: Weinheim.

Sebold M, Deserno L, Nebe S, Schad DJ, Garbusow M, Hagele C *et al* (2014). Model-based and model-free decisions in alcohol dependence. *Neuropsychobiology* **70**: 122–131.

Singer T, Critchley HD, Preuschoff K (2009). A common role of insula in feelings, empathy and uncertainty. *Trends Cogn Sci* **13**: 334–340.

Smith DG, Robbins TW (2013). The neurobiological underpinnings of obesity and binge eating: a rationale for adopting the food addiction model. *Biol Psychiatry* **73**: 804–810.

Stephan KE, Iglesias S, Heinzle J, Diaconescu AO (2015). Translational perspectives for computational neuroimaging. *Neuron* **87**: 716–732.

Stephan KE, Mathys C (2014). Computational approaches to psychiatry. *Curr Opin Neurobiol* **25**: 85–92.

Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009). Bayesian model selection for group studies. *Neuroimage* **46**: 1004–1017.

Sutton RS (1992). Gain adaptation beats least squares? *Proceedings of the 7th Yale Workshop on Adaptive and Learning Systems*, pp 161–166.

Svaldi J, Brand M, Tuschen-Caffier B (2010). Decision-making impairments in women with binge eating disorder. *Appetite* **54**: 84–92.

Takahashi YK, Chang CY, Lucantonio F, Haney RZ, Berg BA, Yau HJ *et al* (2013). Neural estimates of imagined outcomes in the orbito-frontal cortex drive behavior and learning. *Neuron* **80**: 507–518.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N *et al* (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**: 273–289.

Volkow ND, Wang GJ, Tomasi D, Baler RD (2013). The addictive dimensionality of obesity. *Biol Psychiatry* **73**: 811–818.

Voon V, Derbyshire K, Ruck C, Irvine MA, Worbe Y, Enander J *et al* (2015). Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry* **20**: 345–352.

Wechsler D (1955). *Wechsler Adult Intelligence Scale Manual*. Psychological Corporation: New York, NY, USA.

Wiecki T, Poland J, Frank MJ (2015). Model-based cognitive neuroscience approaches to computational psychiatry clustering and classification. *Clinical Psychologial Science* **3**: 378–399.

Wimmer GE, Daw ND, Shohamy D (2012). Generalization of value in reinforcement learning by humans. *Eur J Neurosci* **35**: 1092–1104.

Wunderlich K, Dayan P, Dolan RJ (2012). Mapping value based planning and extensively trained choice in the human brain. *Nat Neurosci* **15**: 786–791.

Ziauddeen H, Farooqi IS, Fletcher PC (2012). Obesity and the brain: how convincing is the addiction model? *Nat Rev Neurosci* **13**: 279–286.