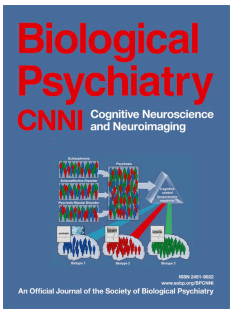


Neurocomputational mechanisms underlying differential reinforcement learning from wins and losses in obesity with and without binge eating

Maria Waltmann, Nadine Herzog, Andrea M.F. Reiter, Arno Villringer, Annette Horstmann, Lorenz Deserno

PII: S2451-9022(24)00160-5

DOI: https://doi.org/10.1016/j.bpsc.2024.06.002

Reference: BPSC 1237

To appear in: *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*

Received Date: 17 October 2023

Revised Date: 24 May 2024

Accepted Date: 9 June 2024

Please cite this article as: Waltmann M., Herzog N., Reiter A.M.F., Villringer A., Horstmann A. & Deserno L., Neurocomputational mechanisms underlying differential reinforcement learning from wins and losses in obesity with and without binge eating, *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* (2024), doi: https://doi.org/10.1016/j.bpsc.2024.06.002.

Neurocomputational mechanisms underlying differential reinforcement learning from wins and losses in obesity with and without binge eating

Running title: Distinct reinforcement learning in obesity and BED

Maria Waltmann[a,b]*, Nadine Herzog[b], Andrea M.F. Reiter [a,c,d], Arno Villringer[b,e], Annette Horstmann[b,f], Lorenz Deserno[a,b,g]

[a] Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Centre of Mental Health, University of Würzburg, Würzburg, Germany

[b] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[c] CRC-940 Volition and Cognitive Control, Faculty of Psychology, Technical University of Dresden, Dresden, Germany

[d] Department of Psychology, Julius-Maximilians-University of Würzburg, Würzburg, Germany

[e] MindBrainBody Institute, Berlin School of Mind and Brain, Charité-Universitätsmedizin Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[f] Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland

[g] Department of Psychiatry and Psychotherapy, Technical University of Dresden, Dresden, Germany

* Corresponding author

E-Mail: waltmann_m@ukw.de

1. Abstract

Background – Binge Eating Disorder (BED) is thought of as a disorder of cognitive control but evidence regarding its neurocognitive mechanisms is inconclusive. Key limitations in prior research are a lack of consistent separation between effects of BED and obesity, and a disregard for self-report evidence suggesting that neurocognitive alterations may emerge primarily in loss- or harm-avoidance contexts.

Methods – Addressing these gaps, this longitudinal study investigated behavioral flexibility and its underlying neuro-computational processes in reward-seeking and loss-avoidance contexts. Obese participants with BED (BED), without BED (OB), and healthy normal-weight participants (NW) (N_{total}=96) performed a probabilistic reversal learning task during functional imaging, with different blocks focused on obtaining wins or avoiding losses. They were reinvited for a 6-months follow-up.

Results – Analyses informed by computational models of reinforcement learning showed that unlike BED, OB performed worse in the win than the loss condition. Computationally, this was explained by differential learning sensitivities in the win vs loss conditions between groups. In the brain, this was echoed in differential neural learning signals in the ventromedial prefrontal cortex (vmPFC) per condition. The differences were subtle, but scaled with BED symptoms, such that more severe BED symptoms were associated with increasing bias towards improved learning from wins vs losses. Across conditions, OB switched more between choice options than NW. This was reflected in diminished representation of choice certainty in the vmPFC.

Conclusions – Our study highlights the importance of distinguishing between obesity with and without BED to identify unique neuro-computational alterations underlying different styles of maladaptive eating behavior.

2. Introduction

Binge Eating Disorder (BED) is a common psychiatric condition (1,2) characterized by repetitive, subjectively uncontrollable overeating. It causes significant distress and is linked to a number of serious comorbidities such as depression, anxiety, and obesity (2–6). BED is recognized as an important public health issue (e.g., 2) but the neurocognitive drivers of binge eating – as distinct from the excessive food intake without loss of control that characterizes obesity – remain poorly understood.

BED can be thought of as a disorder of cognitive-behavioral control (7–10). Consistent with this view, self-report evidence suggests that impulsivity and compulsivity are enhanced in patients (e.g., 11,12). However, the experimental evidence is less clear. Thus, research investigating e.g., delay discounting, risky decision-making, or set-shifting abilities in BED has yielded mixed results (reviewed in e.g., 7,8,10). Methodological limitations may account for this inconsistency (8). Many experimental tasks lack adequate reliability (13,14), reducing their power to detect differences between individuals with and without BED. Further, many studies employ either a normal-weight or an obese control group (e.g., 15,16, but see 17), but both are necessary to capture differential effects of excess weight and BED. Most studies are cross-sectional, precluding investigations of within-subject changes associated with symptoms. Finally, previous research may not have sufficiently engaged with literature showing enhanced “negative urgency” in BED (12,18,19), which hints that patients may show more impulsive behavior in harm-avoidance contexts.

Addressing these limitations, this study examined behavioral flexibility – an important aspect of cognitive-behavioral control – and its neural correlates in obese individuals with BED (BED), without BED (OB), and normal-weight (NW) individuals in a longitudinal design. We employed a reversal learning task that has previously been used to investigate BED in conjunction with functional magnetic resonance imaging (fMRI) (20) and is known to produce reliable metrics (21). To capture potentially different behavior in reward-seeking vs. loss-avoidance contexts, we introduced separate win and loss conditions. To capture within-subject changes in binge-eating symptoms and behavioral flexibility, we reinvited participants for a 6-month follow-up. To obtain more mechanistic insights, we employed computational modelling using reinforcement learning models to inform our behavioral and MRI analyses.

In line with previous work, we hypothesized that both BED and OB would perform worse than NW due to enhanced switching between options (20,22,23). We expected this to be computationally accounted for by greater choice stochasticity (noise in the decision-making process) (20,23,24) and neurally reflected as reduced coding of learning signals – (counterfactual) prediction errors and relative expected values – in the medial prefrontal cortex, as reported previously (20,25). In light of earlier studies (22,26), we further speculated that motivational context might differentially affect BED vs. OB participants. Specifically, considering the role of negative urgency in disordered eating (12,18,19), we hypothesized that BED participants would perform worse than OB in the loss-avoidance condition.

3. Methods

3.1. Participants and procedure.

For this sub-study of a larger project (see ref. 25), we initially enrolled 129 participants between 16 and 49 years (43 each of NW, OB, and BED) matched for age, gender, and body mass index (BMI) in the case of OB and BED (for in- and exclusion criteria, see supplement). The presence of full-blown or subclinical BED was ascertained using the Eating Disorder Examination Interview (27) (details in supplement). Thirteen participants had contra-indications for MRI scanning and completed the experiment outside the scanner. We excluded these subjects and their matches from the analysis reported here (final n=96). However, we report the analysis on the full sample in the supplement and note discrepancies.

As part of the study protocol (https://osf.io/fyn6q), participants performed a probabilistic reversal learning task during fMRI (25). A minimum of 6 months after their first visit (T1), participants were re-invited for a follow-up session (T2) in which they repeated the task without MRI measurement. The interval was chosen to allow for change in binge-eating symptoms, however, due to restrictions during the Covid-19 pandemic, many participants were re-assessed after a longer period (max 28 months, median = 7.85 months). Participants provided written informed consent and were financially compensated for their time (parental consent and Amazon voucher for minors). The Leipzig University ethics committee granted ethical approval (385/17-ek). For sample characteristics, see Table 1.

3.2. Task.

We employed a modified probabilistic reversal learning task to assess reinforcement learning and behavioral flexibility (20,25,28–30). Our version (25) has two blocks of 140 trials in which participants make repeated choices between two cards. The cards have different probabilities of yielding a win (+10 cents) vs. a neutral (± 0 cents) outcome (80% and 20%) in the win block, and of yielding a loss (− 10 cents) vs. a neutral outcome (± 0 cents) in the loss block (order counterbalanced) (Fig. 1 – A). Five times in each block, the outcome contingencies reverse, and participants have to re-learn them. Neutral outcomes represent negative feedback (no win) in the win condition, and positive feedback (no loss) in the loss condition. This allows us to differentiate asymmetric learning from valenced feedback (positive vs. negative) from asymmetric learning for reward-seeking vs. loss-avoidance. For further details on the task and procedure, see supplement.

3.3. Analysis

3.3.1. Task performance

We used trial-by-trial logistic mixed-effects models with maximal random effects (31) using the *fitglme* function in MATLAB R2023a to estimate accuracy (probability of choosing the better card), choice switching (probability of choosing another card than in the previous trial), and perseveration (probability of choosing the same card after it has been punished twice). As predictors, we included group (NW, OB, BED, with OB as reference category), condition (win vs. loss), and previous feedback (positive vs. negative) for choice switching. Using OB as reference allowed us to test the two comparisons central to our design: the difference between BED and OB, reflecting effects of loss-of-control eating separate from excess weight; and the difference between NW and OB, reflecting effects of excess weight separate from loss-of-control eating. We further differentiated between pre-reversal trials, i.e., the trials leading up to each reversal (115 trials in total per block), and post-reversal trials, i.e., the 5 trials directly following each reversal (25 trials per block) (for details and descriptive statistics, see supplement).

3.3.2. Computational modelling

To assess processes underlying behavior, we fit 15 computational models (full descriptions in supplement). According to the winning model based on integrated BICs

(Fig. S1), agents learn the expected value of each card by using trial-by-trial prediction errors (i.e., the difference between expected value and actual outcome, PEs) to update the value of both the chosen (Eq. 1) and the unchosen option (Eq. 2). The latter update requires inference on the outcome of the counterfactual choice (Fig. 1 – B).

$$Q_{chosen,trial+1} = Q_{chosen,trial} + \alpha\left(\rho * reward - Q_{chosen,trial}\right) \quad (1), \text{ where}$$

$$\alpha = \alpha_+ \text{ and } \rho = \rho_+ \quad \forall \text{ reward} > 0$$

$$\alpha = \alpha_- \text{ and } \rho = \rho_- \quad \forall \text{ reward} < 0$$

$$Q_{unchosen,trial+1} = Q_{chosen,trial} + \kappa\alpha\left(-(\rho * reward) - Q_{unchosen,trial}\right) \quad (2), \text{ where}$$

$$\alpha = \alpha_- \text{ and } \rho = \rho_+ \quad \forall \text{ reward} > 0$$

$$\alpha = \alpha_+ \text{ and } \rho = \rho_- \quad \forall \text{ reward} < 0$$

Action selection is performed by a softmax rule:

$$p(a_i) = \frac{exp\left(Q_{a_i}\right)}{\sum_{j=1}^{K} exp\left(Q_{a_j}\right)}$$

The model has separate learning rates (α) for positive and negative feedback, and a weight (κ) on the learning rate for updates of the unchosen option. The reinforcement sensitivity parameter (ρ) determines the maximum difference between expected values and thus poses a lower bound to choice stochasticity. The model allows for different sensitivity to positive and negative feedback, resulting in asymmetric stay-switch behavior (e.g., with higher positive ρ, the tendency to stay after positive feedback would be stronger than the tendency to switch after negative feedback). The model showed overall good fit and recoverability (see supplement). We compared fitted parameters from the winning model between groups using linear mixed-effects models (using *fitlme* in MATLAB R2023a).

3.3.3. Effects of binge-eating frequency.

To investigate the effects of binge-eating frequency (BEF) on task performance, we repeated all generalized linear mixed-effects models in the BED group only, with average BEF across sessions and change in BEF as predictors. Change was included as the difference from the average in each session (i.e., if BEF was 6 at T1 and 4 at

T2, the average would be 5, and change would be +1 at T1 and -1 at T2). This allows for the separation of within- and between subject effects related to symptom severity. Crucially, the regression coefficient for within-subject changes reflects *change* in symptoms associated with a *change* in BEF. We removed one participant with implausibly high BEF (>3 SD from mean, i.e., 80 binge-eating episodes/month).

3.3.4. Post-hoc tests, sensitivity and exploratory analyses

For all models, we used simple effects analyses to unpack interactions. We ascertained that the results were not driven by group differences in depression or anxiety in sensitivity analyses. We further explored effects of BMI in OB and BED, and of the UPPS-scales *urgency* and *lack of perseverance* across groups. The results are reported in the supplement.

3.4. fMRI

For scanning sequences and preprocessing steps, see supplement. We applied event-related analyses using the general linear model implemented in SPM12, with feedback onsets, cue onsets, missing trials, and the 6 movement parameters as regressors. We added parametric modulators informed by computational modelling as described previously (25). Thus, we added single and double update prediction errors as modulators of feedback onsets, and choice probability (relative expected value of the chosen option) as modulator of cue onsets (for details, see supplement). Data from the win and loss blocks were analyzed in one model, with each block modeled as a separate session. The regressors were convolved with the canonical hemodynamic response function.

For 2nd level analyses, we estimated random-effects ANOVAs on the contrast images of the parametric modulators. The ANOVAs included a within-subject condition factor (win vs. loss) and a between-subject group factor. Hence, we estimated models predicting choice probability from group and condition, and PE-coding from SU vs DU, group and condition. For group comparisons, we focused on the ventromedial prefrontal cortex (vmPFC) based on previous work (20). For small-volume-correction, we used an ROI defined as a 4mm sphere (encompassing 33 voxels) around the peak vmPFC voxel associated with valuation ([2 46 -8]) identified in a meta-analysis (32). Results were considered significant at $p_{FWE\_SVC}<.05$, where family-wise error correction was applied to the peak level.

4. Results

4.1. Task performance

Context-independent effects. There were no straightforward performance differences between groups. However, BED and NW had larger differences in accuracy between pre- and post-reversal trials than OB (BED-OB x trial-type: beta=-0.25 t(44717)=-2.4, p=.02; NW-OB x trial-type: beta=-0.25, t(44717)=-2.49, p=.01; Fig. 2–A). Simple effects analyses suggest that the difference between BED and OB was mainly driven by worse performance of BED after reversals (BED-OB pre-reversal: beta=0.29, t(44717)=1.53, p=.13; BED-OB post-reversal: beta=-0.21, t(44717)=-2.49, p=.01). The difference between OB and NW appeared to be driven by both relatively worse performance before (NW-OB: beta=.34, t(44717)=-1.84, p=.07 pre-reversal) and better performance after reversals (NW-OB: beta=-.17, t(44717)=-2.05, p=.04 post-reversal) in OB. BED and NW did not differ.

The difference in accuracy between pre- and post-reversal trials tracks successful learning during pre-reversal trials, suggesting that OB learn less efficiently than NW, with BED landing in between. Our analysis of choice switching corroborated this: BED and OB did not differ; however, OB had a smaller difference in choice switching before and after reversals than NW (NW-OB x trial-type: beta=0.13, t(43943)=2.26, p=.02, Fig. 2-B). Simple effects analyses suggested that this was driven by enhanced switching before (NW-OB: beta=-.41, t(43943)=-1.96, p=.05) but not after reversals (NW-OB: beta=-0.15, t(43943)=-0.79, p=.43) in OB. BED and NW did not differ. Excessive switching between options, especially before reversals, reflects inefficient learning (correlation between pre-reversal choice switching and accuracy: r=-.91, p<.001).

There were no context-independent differences between BED, OB, or NW in terms of perseveration.

Context-dependent effects. Motivational context had a different impact on accuracy in BED compared to OB (BED-OB x condition: beta=0.12, t(44717)=2.42, p=.02, Fig. 3–A). Thus, BED had similar accuracy in the win and loss conditions (condition in BED: beta=0.02, t(44717)=0.5, p=.62), whilst OB performed worse in the win than the loss condition (condition in OB: beta=-0.10, t(44717)=2.94, p=.003). BED and NW did not differ.

There were no context-dependent differences between BED, OB, or NW in terms of choice switching. However, the group differences in accuracy were mirrored in perseveration (BED-OB x condition: beta=-0.28, t(9008)=-2.53, p=.01, Fig. 3–C). Simple effects analyses showed that BED perseverated to a similar extent in the win and loss conditions (condition in BED: beta=-0.07, t(9008)=-0.88, p=.38), while OB perseverated more in the win than the loss condition (condition in OB: beta=0.21, t(9008)=2.66, p=.008). BED and NW did not differ.

Like excessive switching, perseveration signals ineffective learning (correlation between perseveration and accuracy: r=-.19, p=.03). The results therefore suggest that, in contrast to BED, OB may learn more effectively in the loss condition due to reduced perseveration.

Sensitivity analyses. When participants without MRI measurement were included in the analysis (N=129), the context-independent effects were no longer significant. However, the context-dependent group differences between BED and OB participants in accuracy and perseveration remained (accuracy: BED-OB x condition: beta=0.09, t(61370)= 2.11, p=.04; perseveration: BED-OB x condition: beta=-0.19, t(12309)= -1.98, p=.05) (see supplement for details).

4.2. Effects of binge-eating frequency.

Next, we investigated how average binge-eating frequency (BEF) and longitudinal change in BEF were associated with task performance. Consistent with the group-level findings, results showed a context-dependent effect of average BEF on accuracy (average BEF x condition: beta=0.14, t(13014)=1.99, p=.05). As Fig. 3 – B shows, the difference in accuracy between conditions (win>loss) increased with increasing average BEF. There was no effect of longitudinal change in BEF on accuracy.

There were no effects of average BEF on choice switching. However, there was an interaction between longitudinal change in BEF and condition (beta=-0.36, t(12807)=-2.09, p=.04, Fig. 3 – D). Thus, participants switched less between options in the win condition when they reported higher BEF relative to baseline (change in BEF in win condition: beta=-1.26, t(12807)=-3.38, p<.001; change in BEF in loss condition: beta=-0.53, t(12807)=-1.14, p=.25).

Together, the findings suggest that worse BED symptoms may be associated with a bias towards improved learning from wins vs losses, possibly due to diminishing choice switching in the win condition as symptoms worsen.

In addition, there was complex three-way interaction of change in BEF, previous feedback, and trial-type on choice switching (for details, see supplement). There were no significant effects of average BEF or change thereof on perseveration.

4.3. Computational modelling.

Context-independent effects. Contrary to expectations, there were no significant group differences in reinforcement sensitivities (ρ). However, BED participants had more asymmetric learning rates (α) and double-update learning rates (α*κ) for positive and negative feedback than OB (α – BED-OB x feedback: beta=0.04, t(372)=2.29, p=.02; α*κ – BED-OB x feedback: beta=0.02, t(372)=2.53, p=.01; Fig. 2–C). Simple effects analyses showed a trend for slower double-update learning after negative but not positive feedback in BED (BED-OB in negative feedback: beta=-0.05, t(372)=-1.93, p=.06; BED-OB in positive feedback: beta=-0.01, t(372)=-0.52, p=.60). NW also had more asymmetric double-update learning rates (α*κ) for positive and negative feedback than OB (NW-OB x feedback: beta=0.02, t(372)=2.07, p=.04; Fig. 2–C). This seemed to be due to slower double-update learning after positive and faster double-update learning after negative feedback in OB, although the group effect was not significant for either positive or negative feedback.

Asymmetric learning rates for positive and negative feedback make agents more resistant to uninformative (stochastic) negative feedback. Thus, reduced asymmetry between learning rates leads to less sharply distinguished expected values and therefore to choice switching (correlation between double-update learning rate asymmetry and choice switching: r=-.36, p<.001). Thus, the findings suggest that reduced asymmetry of double-update learning rates for positive and negative feedback may account for our finding of enhanced choice switching in OB.

Context-dependent effects. Across groups, participants had higher reinforcement sensitivities (ρ), i.e., behaved less noisily, in the loss condition (condition: beta=-0.03, t(380)=-2.21, p=.03). However, there were no group differences.

BED and OB groups significantly differed in terms of their learning rates (α) in the win and loss conditions (BED-OB x condition: beta=0.02, t(372)=2.09, p=.04). BED, but not OB participants, had lower learning rates in the loss than in the win condition (condition in BED: beta=0.1, t(372)=2.41, p=.02; condition in OB: beta=-0.004, t(372)=-0.54, p=.58). This pattern did not quite match our observations at the behavioral level. We therefore reasoned that a lower learning rate in the loss condition in BED might compensate for higher reinforcement sensitivity in the loss condition across groups. The result would be higher "learning sensitivity" in the loss condition in OB compared to BED, possibly accounting for improved performance in this condition.

We therefore calculated the product of learning rate and reinforcement sensitivity (ρ*α, the "learning sensitivity"). BED had higher learning sensitivity for wins than losses, while OB had higher learning sensitivity for losses than wins (BED-OB x condition: beta=0.09, t(372)=2.06, p=.04, Fig 3–E). Critically, learning sensitivity (ρ*α) was highly correlated with accuracy (r=.82, p<.001). Differences in learning sensitivities in the win and loss conditions may thus account for context-dependent differences in accuracy between BED and OB.

There were no context-dependent differences in double-update learning rate between BED, OB, or NW. There were no associations between BEF and model parameters.

Sensitivity analyses. The results did not change when we excluded 7 participants fit at chance level. When participants without MRI measurement were included during fitting (N=129), the difference between BED and OB participants in learning sensitivity for wins and losses, and learning rates for positive and negative feedback were still significant (for details, see supplement). The other group differences were no longer significant.

4.4. fMRI results

Next, we explored how our findings of group differences at the behavioral and computational level were reflected in the coding of model-derived learning signals in the brain (for group level results, see supplement).

Context-independent effects. The two obese groups showed less activation associated with choice probability in the vmPFC than NW (2/46/-12, t=2.83, $p\text{-FWE}_{\text{small-volume}}$=.014, Fig. 2 – D and E). Since choice probability can be understood as reflecting

confidence in the upcoming choice, this dovetails with our findings at the behavioral and computational level: OB showed enhanced switching, which we have argued may be a consequence of less sharply distinguished expected values due to less asymmetric learning from positive and negative feedback. There were no other context-independent group differences.

Context-dependent effects. There was a marginal interaction between DU vs SU, condition, and group with respect to PE-coding in the vmPFC (2/42/-8, t=2.16, p-FWE$_{small\text{-}volume}$=.063, Fig. 3 – F and G). Activation in response to DU-PEs was stronger in the win condition in BED, and stronger in the loss condition in in OB. That is, the neural learning signal incorporating counterfactual inference was more pronounced in the win than the loss condition in BED participants and vice versa in OB. Importantly, this signal also correlated with average BEF in BED, such that more binge-eating episodes were associated with larger differences in DU-signal between the win and loss conditions (r=.32, p=.04, Fig. 3 – H). This echoes our findings of higher learning sensitivity in reward-seeking vs. loss-avoidance contexts in BED, and more asymmetric learning from wins vs losses in participants with more frequent binge-eating episodes. There were no other context-dependent group differences.

5. Discussion

In this study, we used a probabilistic reversal learning task, computational modelling, and fMRI in a longitudinal design to investigate shared and distinct neurocognitive mechanisms of altered decision-making in BED and OB compared to NW.

We demonstrate subtle differences between BED and OB with regard to learning in different motivational contexts. Thus, unlike BED, OB performed better when learning for loss-avoidance (loss condition) than reward-seeking (win condition), putatively thanks to less perseveration. This is broadly in line with our hypotheses regarding BED but somewhat at odds with prior reports of difficulty with learning from losses in obesity (33,34). However, the samples used in (33,34) were not screened for BED and may have been confounded – something the study reported here was designed to avoid. For our data, computational modelling suggested that the condition-specific performance difference may reflect relatively enhanced learning sensitivity (product of reinforcement sensitivity and learning rate) in the loss condition in OB, and relatively

reduced learning sensitivity in the loss condition in BED. Consistent with this, a neural learning signal incorporating counterfactual inference in the vmPFC was comparatively stronger in the loss condition in OB and comparatively weaker in the loss condition in BED. This effect was only marginally significant. However, reduced coding of counterfactual prediction errors in the vmPFC has previously been shown to characterize BED (20); and the signal was correlated with binge-eating frequency, with greater differences between conditions (win>loss) associated with higher frequencies. This mirrored the association between binge-eating frequency and the difference in performance between the win and loss conditions, where higher frequencies were also associated with greater differences between conditions. In addition, and consistent with this, worsening symptoms were associated with less choice switching in the win but not the loss condition. In sum, the results suggests that BED and OB may be characterized by different neurocognitive learning biases, with better learning from wins than losses in BED, and better learning from losses than wins in OB.

Independent of motivational context, OB showed more choice-switching, particularly before reversals, leading to worse performance in OB compared to NW before, and better performance after reversals compared to BED and NW. Computationally, switching was accounted for by less asymmetric counterfactual learning rates for positive and negative feedback in OB compared to BED and NW. A lack of asymmetry makes agents more sensitive to uninformative feedback, and thus leads to less sharply distinguished expected values and enhanced choice switching. At the neural level, obese participants (with and without BED) showed reduced coding of choice probability – a reflection of the difference between expected values – in the vmPFC. The effects were small but resonate with previous reports of enhanced switching in OB (23). The absence of discernible differences between BED and OB in pre-reversal accuracy, switching behavior, and fMRI observations gesture towards a general effect of obesity, irrespective of BED. However, given group differences in post-reversal accuracy and model parameters which diverge from this trend, we hesitate to draw definitive conclusions.

Together, our results indicate that obesity in the context of BED may vary qualitatively from obesity without loss-of-control eating. Indeed, there may be a bias towards worse learning from losses than wins in BED, and vice versa in OB. This fits with the clinical picture of BED, where patients repeat actions they know will make them feel bad. It

also chimes with demonstrably enhanced negative urgency in this group (18,19,35), which may disturb learning and decision-making for loss-avoidance. Indeed, there is evidence that inhibitory control and risk-taking in BED are affected by negative mood (36,37). It would be interesting to test, by experimentally manipulating stress or mood before task performance in the lab, or via more ecological, smartphone-based approaches, whether this extends into the reinforcement learning realm.

Our results are intriguing but not without limitations. The effects we show are subtle, complex, and not always easily unpacked. Importantly, while we saw specific differences between BED and OB participants, the BED and NW groups did not significantly differ from one another. This raises the question whether the BED-OB differences are driven by BED-specific alterations or a normalization of obesity-associated alterations. While our analysis of the effects of binge-eating severity and the absence of context-specific differences between OB and NW suggest the former, further research in large clinical multicenter samples will be necessary to replicate and clarify group differences, and to properly disentangle the dimensional effects of BMI and binge-eating severity. Our data yield important leads but the groups were too small to produce sufficiently dependable dimensional estimates, let alone investigate potential non-linear effects, which may be especially interesting with regard to obesity without BED (38). Likewise, our evidence is confined to the monetary domain, however, food rewards may be processed differently.

In conclusion, our data suggest that reinforcement learning in obesity with and without BED may be subject to qualitatively different neurocomputational learning biases. Thus, individuals with BED may have a bias towards worse learning from losses than wins, and obese individuals without BED may have a bias towards worse learning from wins than losses. Obesity without BED was further associated with reinforcement learning difficulties due to enhanced choice switching. Our findings highlight the importance of distinguishing between obesity with and without BED.

6. Acknowledgements

This work was supported by the IFB Adiposity Diseases, Federal Ministry of Education and Research (BMBF), Germany, GN: 01EO1501, the German Research Foundation (DFG) as part of Collaborative Research Centre 265 “Losing and Regaining Control

over drug intake" (402170461, Project A02), and the Max Planck Society. Neither funding source had any role in study design, data collection, analysis or the interpretation of the data. AR further acknowledges support from the German Research Foundation (DFG RE 4449/1-1, SFB 940-3/B7, RTG-2660) and by a 2020 BBRF Young Investigator Grant.

LD and AH designed the study and acquired funding; NH and MW acquired the data; MW and LD analyzed the data and wrote the original draft; MW, NH, AH, AV, AR, LD reviewed and edited the manuscript.

We thank Miriam Huml, Eva Burmeister, and Lisa-Marie Okhof for their help with data collection and management.

All data (except subject-level MRI data) and analysis code is available via https://osf.io/b9tyf/.

7. Disclosures

MW, NH, AH, AV, AR and LD have no biomedical financial interests or potential conflicts of interest to disclose.

8. References

1. Hudson JI, Hiripi E, Pope HGJ, Kessler RC, Pope Jr. HG, Kessler RC (2007): The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biol Psychiatry* 61: 348–358.

2. Kessler RC, Berglund PA, Chiu WT, Deitz AC, Hudson JI, Shahly V, *et al.* (2013): The Prevalence and Correlates of Binge Eating Disorder in the World Health Organization World Mental Health Surveys. *Biol Psychiatry* 73: 904–914.

3. Grilo CM, White MA, Barnes RD, Masheb RM (2013): Psychiatric disorder co-morbidity and correlates in an ethnically diverse sample of obese patients with binge eating disorder in primary care settings. *Compr Psychiatry* 54: 209–216.

4. Thornton LM, Watson HJ, Jangmo A, Welch E, Wiklund C, von Hausswolff-Juhlin Y, *et al.* (2017): Binge-eating disorder in the Swedish national registers: Somatic comorbidity. *Int J Eat Disord* 50: 58–65.

5. Udo T, Grilo CM (2018): Prevalence and Correlates of DSM-5-Defined Eating Disorders in a Nationally Representative Sample of U.S. Adults. *Biol Psychiatry* 84: 345–354.

6. Welch E, Jangmo A, Thornton LM, Norring C, von Hausswolff-Juhlin Y, Herman BK, *et al.* (2016): Treatment-seeking patients with binge-eating disorder in the Swedish national registers: clinical course and psychiatric comorbidity. *BMC Psychiatry* 16: 163.

7. Cury MEG, Berberian A, Scarpato BS, Kerr-Gaffney J, Santos FH, Claudino AM (2020): Scrutinizing Domains of Executive Function in Binge Eating Disorder: A Systematic Review and Meta-Analysis. *Front Psychiatry* 11: 288.

8. Waltmann M, Herzog N, Horstmann A, Deserno L (2021): Loss of control over eating: A systematic review of task based research into impulsive and compulsive processes in binge eating. *Neurosci Biobehav Rev* 129: 330–350.

9. Kessler RM, Hutson PH, Herman BK, Potenza MN (2016): The neurobiological basis of binge-eating disorder. *Neurosci Biobehav Rev* 63: 223–238.

10. Prunell-Castañé A, Jurado MÁ, García-García I (2021): Clinical binge eating, but

not uncontrolled eating, is associated with differences in executive functions: Evidence from meta-analytic findings. *Addict Behav Reports* 13: 100337.

11. Deal LS, Wirth RJ, Gasior M, Herman BK, McElroy SL (2015): Validation of the yale-brown obsessive compulsive scale modified for binge eating. *Int J Eat Disord* 48: 994–1004.

12. Magel CA, von Ranson KM (2021): Negative urgency combined with negative emotionality is linked to eating disorder psychopathology in community women with and without binge eating. *Int J Eat Disord* 54: 821–830.

13. Enkavi AZ, Eisenberg IW, Bissett PG, Mazza GL, MacKinnon DP, Marsch LA, Poldrack RA (2019): Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proc Natl Acad Sci* 116: 5472 LP – 5477.

14. Hedge C, Powell G, Sumner P (2018): The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behav Res Methods* 50: 1166–1186.

15. Aloi M, Rania M, Caroleo M, Bruni A, Palmieri A, Cauteruccio MA, *et al.* (2015): Decision making, central coherence and set-shifting: a comparison between Binge Eating Disorder, Anorexia Nervosa and Healthy Controls. *BMC Psychiatry* 15: 6.

16. Manasse SM, Forman EM, Ruocco AC, Butryn ML, Juarascio AS, Fitzpatrick KK (2015): Do executive functioning deficits underpin binge eating disorder? A comparison of overweight women with and without binge eating pathology. *Int J Eat Disord* 48: 677–683.

17. Preuss H, Leister L, Pinnow M, Legenbauer T (2019): Inhibitory control pathway to disinhibited eating: A matter of perspective? *Appetite* 141: 104297.

18. Racine SE, Martin SJ (2016): Exploring divergent trajectories: Disorder-specific moderators of the association between negative urgency and dysregulated eating. *Appetite* 103: 45–53.

19. Ralph-Nearman C, Stewart JL, Jones KA (2020): The role of negative urgency in risky alcohol drinking and binge-eating in United Kingdom male and female students. *Addict Behav Reports* 11: 100274.

20. Reiter AMF, Heinze HJ, Schlagenhauf F, Deserno L (2017): Impaired Flexible Reward-Based Decision-Making in Binge Eating Disorder: Evidence from Computational Modeling and Functional Neuroimaging. *Neuropsychopharmacology* 42: 628–637.

21. Waltmann M, Schlagenhauf F, Deserno L (2022): Sufficient reliability of the behavioral and computational readouts of a probabilistic reversal learning task. *Behav Res Methods*. https://doi.org/10.3758/s13428-021-01739-7

22. Banca P, Harrison NA, Voon V (2016): Compulsivity Across the Pathological Misuse of Drug and Non-Drug Rewards. *Front Behav Neurosci* 10: 154.

23. Mathar D, Neumann J, Villringer A, Horstmann A (2017): Failing to learn from negative prediction errors: Obesity is associated with alterations in a fundamental neural learning mechanism. *Cortex* 95: 222–237.

24. 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.

25. Waltmann M, Herzog N, Reiter AMF, Villringer A, Horstmann A, Deserno L (2023): Diminished reinforcement sensitivity in adolescence is associated with enhanced response switching and reduced coding of choice probability in the medial frontal pole. *Dev Cogn Neurosci* 60: 101226.

26. 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.

27. Hilbert A, Tuschen-Caffier B, Ohms M (2004): Eating Disorder Examination: Deutschsprachige Version des strukturierten Essstörungsinterviews. *Diagnostica* 50: 98–106.

28. Boehme R, Deserno L, Gleich T, Katthagen T, Pankow A, Behr J, *et al.* (2015): Aberrant Salience Is Related to Reduced Reinforcement Learning Signals and Elevated Dopamine Synthesis Capacity in Healthy Adults. *J Neurosci* 35: 10103–10111.

29. Reiter AMF, Deserno L, Kallert T, Heinze H-J, Heinz A, Schlagenhauf F (2016): Behavioral and Neural Signatures of Reduced Updating of Alternative Options in Alcohol-Dependent Patients during Flexible Decision-Making. *J Neurosci* 36: 10935–10948.

30. Deserno L, Boehme R, Mathys C, Katthagen T, Kaminski J, Stephan KE, *et al.* (2020): Volatility Estimates Increase Choice Switching and Relate to Prefrontal Activity in Schizophrenia. *Biol Psychiatry Cogn Neurosci Neuroimaging* 5: 173–183.

31. Barr DJ, Levy R, Scheepers C, Tily HJ (2013): Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang* 68: 10.1016/j.jml.2012.11.001.

32. 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.

33. Kastner L, Kube J, Villringer A, Neumann J (2017): Cardiac Concomitants of Feedback and Prediction Error Processing in Reinforcement Learning . *Frontiers in Neuroscience* , vol. 11. Retrieved from https://www.frontiersin.org/articles/10.3389/fnins.2017.00598

34. Coppin G, Nolan-Poupart S, Jones-Gotman M, Small DM (2014): Working memory and reward association learning impairments in obesity. *Neuropsychologia* 65: 146–155.

35. Kenny TE, Singleton C, Carter JC (2019): An examination of emotion-related facets of impulsivity in binge eating disorder. *Eat Behav* 32: 74–77.

36. Leehr EJ, Schag K, Dresler T, Grosse-Wentrup M, Hautzinger M, Fallgatter AJ, *et al.* (2018): Food specific inhibitory control under negative mood in binge-eating disorder: Evidence from a multimethod approach*. *Int J Eat Disord* 51: 112–123.

37. Danner UN, Evers C, Sternheim L, van Meer F, van Elburg AA, Geerets TAM, *et al.* (2013): Influence of negative affect on choice behavior in individuals with binge eating pathology. *Psychiatry Res* 207: 100–106.

38. Horstmann A (2015): Argument for a non-linear relationship between severity of human obesity and dopaminergic tone Obesity Etiology and Pathophysiology Argument for a non-linear relationship between severity of human obesity and dopaminergic tone. https://doi.org/10.1111/obr.12303

39. Reitan RM (1955): The relation of the Trail Making Test to organic brain damage. *J Consult Psychol* 19: 393–394.

40. Wechsler D (2008): Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). *San Antonio, TX NCS Pearson* 22: 498.

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

42. Hilbert A, de Zwaan M, Braehler E (2012): How Frequent Are Eating Disturbances in the Population? Norms of the Eating Disorder Examination-Questionnaire ((D. Tomé, editor)). *PLoS One* 7: e29125.

43. Meule A, Vögele C, Kübler A (2011): Psychometrische Evaluation der deutschen Barratt Impulsiveness Scale – Kurzversion (BIS-15). *Diagnostica* 57: 126–133.

44. Whiteside SP, Lynam DR (2001): The Five Factor Model and impulsivity: using a structural model of personality to understand impulsivity. *Pers Individ Dif* 30: 669–689.

45. Hilbert A, Baldofski S, Zenger M, Löwe B, Kersting A, Braehler E (2014): Weight Bias Internalization Scale: Psychometric Properties and Population Norms. *PLoS One* 9: 1–7.

46. Schulte EM, Gearhardt AN (2017): Development of the Modified Yale Food Addiction Scale Version 2.0. *Eur Eat Disord Rev* 25: 302–308.

47. Cepeda-Benito A, Gleaves DH, Williams TL, Erath SA (2000): The development and validation of the state and trait food-cravings questionnaires. *Behav Ther* 31: 151–173.

48. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961): An inventory for measuring depression. *Arch Gen Psychiatry* 4: 561–571.

49. Spielberger CD, Gorsuch RL, Lushene RE (1970): Manual for the State-Trait

Anxiety Inventory.

9. Tables

Table 1. Demographics and sample characterization

	NW	OB	BED	p
N	32	32	32	
Age	29.28 (±6.13)	30.26 (±6.09)	29.82 (±6.90)	0.83
BMI	22.30 (±2.04)	35.54 (±3.57)	35.47 (±4.67)	0.94
Follow-up Interval (years)	0.80 (±0.50)	0.90 (±0.38)	0.76 (±0.29)	0.55
Drop-out	21.88 %	37.50 %	34.38 %	0.37
Gender	75.00 %	75.00 %	75.00 %	1
Years of education (full-time)	17.77 (±3.92)	16.58 (±5.11)	17.76 (±3.48)	0.44
TMT-A	19.69 (±4.43)	20.52 (±5.19)	19.81 (±4.95)	0.77
TMT-B	41.86 (±10.64)	43.10 (±15.32)	37.24 (±8.41)	0.12
Digit Span Forward	6.69 (±1.26)	6.25 (±0.95)	6.45 (±1.09)	0.29
Digit Span Backwards	5.34 (±1.36)	4.81 (±1.26)	5.26 (±1.09)	0.19
Digit-Symbol-Substitution Task	83.41 (±11.57)	81.16 (±15.19)	78.03 (±12.24)	0.26
Verbal IQ (Wortschatztest)	109.78 (±8.88)	102.84 (±9.64)	105.31 (±5.83)	<0.001
EDE-Q Binge episodes (last 28 days)	0.32 (±1.14)	0.48 (±1.86)	6.84 (±5.04)	<0.001
EDE-Q total	0.79 (±0.95)	1.67 (±1.31)	2.48 (±0.78)	<0.001
EDE-Q restraint	0.74 (±0.86)	1.41 (±1.22)	1.58 (±1.01)	<0.001
EDE-Q Eating Concern	0.22 (±0.25)	0.62 (±0.71)	1.85 (±0.90)	<0.001
EDE-Q Weight Concern	1.87 (±2.24)	4.12 (±3.17)	6.06 (±1.74)	<0.001

EDE-Q Shape Concern	0.34 (±1.13)	0.53 (±1.36)	0.44 (±0.92)	0.81
BIS 15	29.71 (±7.53)	30.03 (±6.19)	35.26 (±6.92)	<0.001
UPPS Urgency	24.94 (±5.38)	26.52 (±5.32)	34.32 (±5.28)	<0.001
UPPS Premeditation (-)	22.06 (±4.68)	21.45 (±4.43)	23.45 (±5.42)	0.26
UPPS Perseverance (-)	19.19 (±5.79)	18.29 (±3.36)	22.03 (±4.35)	0.01
UPPS Sensation Seeking	31.65 (±7.32)	31.81 (±7.25)	30.94 (±6.79)	0.88
WBIS	22.27 (±12.32)	39.34 (±14.60)	50.52 (±12.88)	<0.001
YFAS	0.16 (±0.45)	0.75 (±1.57)	3.94 (±2.57)	<0.001
FCQ	10.13 (±3.53)	11.69 (±3.84)	16.11 (±3.59)	<0.001
BDI	4.31 (±5.51)	7.84 (±5.66)	15.79 (±8.99)	<0.001
STAI (Trait)	37.23 (±11.18)	38.97 (±9.76)	50.61 (±10.43)	<0.001

N.B. p-values reflect one-way ANOVAs except for BMI, which reflects a t-test between OB and BED subjects. TMT – Trail Making Test (39), Digit Span Task (40), Digit Symbol Substitution Task (40), Wortschatztest (41), EDE-Q – Eating Disorder Examination Questionnaire (42), BIS-15 – Barratt Impulsiveness Scale – Short Version (43), UPPS – UPPS Impulsive Behavior Scale (44), WBIS – Weight Bias Internalization Scale (45), YFAS – Yale Food Addiction Scale, modified version (46), FCQ – Food Craving Questionnaire (47), BDI – Beck Depression Inventory (48), STAI – State Trait Anxiety Inventory (49)

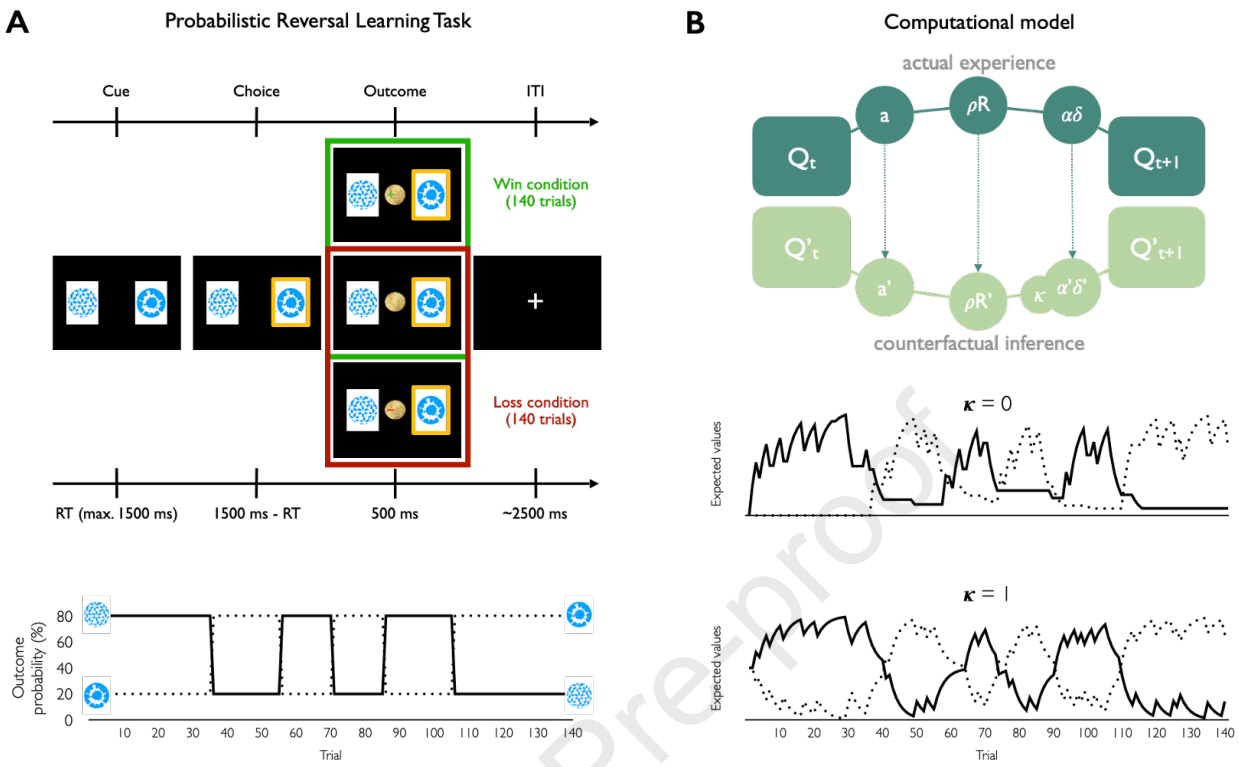
10. Figure legends

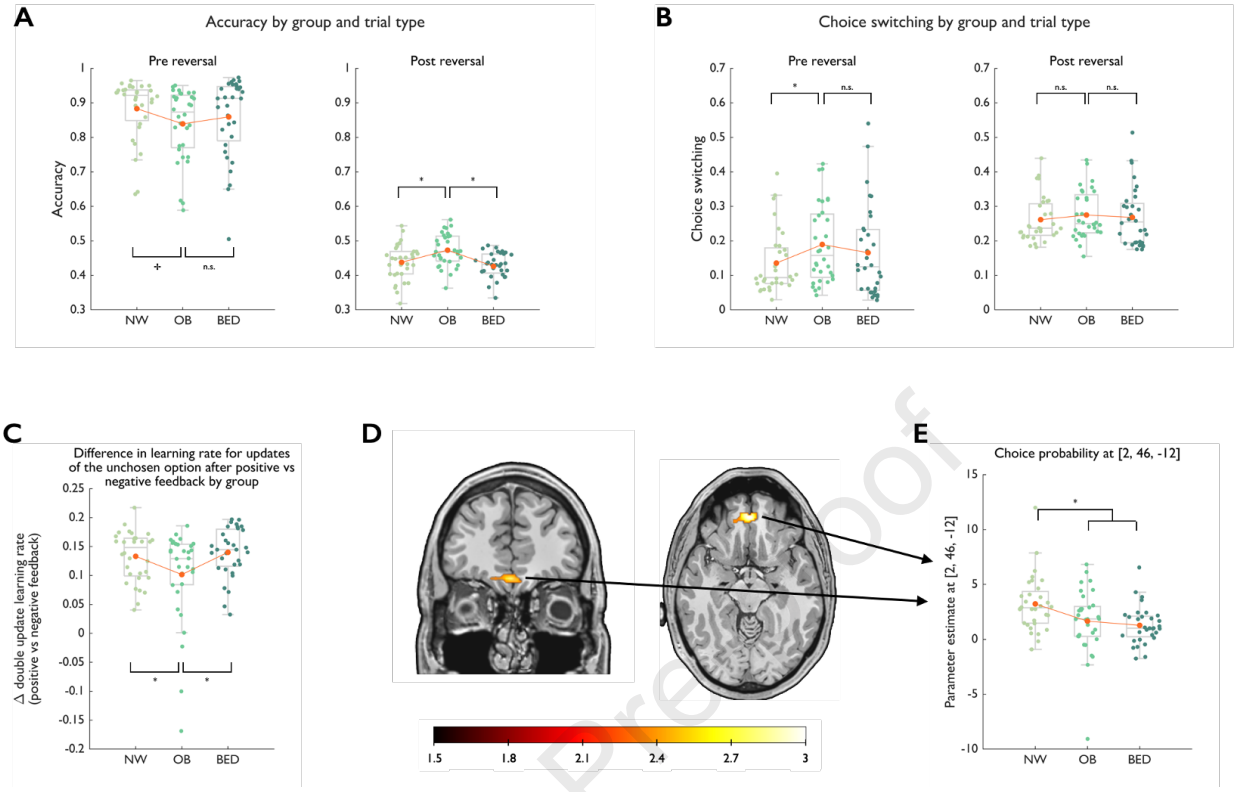
Fig. 1. – A. Schematic of the probabilistic reversal learning task (PRLT). Participants make 140 binary choices between two abstract stimuli (cards) with different probabilities of rewards, neutral outcomes, or losses. The goal is to gain as much and lose as little money as possible, depending on condition (win or loss). In the win condition, a positive outcome means gaining 10 cents, while a neutral response (±0 cents) represents a negative outcome. In the loss condition, a neutral response (±0 cents) is a positive outcome, while a negative outcome means losing 10 cents. In each trial, the stimuli are presented for a maximum of 1500ms or until the participant

responds. A frame then appears around the chosen card and remains visible for the 1500ms minus the response time. Feedback is indicated through pictures of coins: a 10-cents coin for wins, 0-cents coin for neutral outcomes, and minus 10-cents coin for losses. Trials end with a variable intertrial interval (mean 2500ms) during which participants see a fixation cross. The lower panel shows the reward contingencies. In the initial 35 trials, the stimuli have win/loss probabilities of 20% and 80% respectively. The contingencies reverse five times throughout the task (after the 35th, 55th, 70th, 85th, and 105th trial), requiring participants to adapt their behavior to maximize gains and avoid losses. The order of conditions was randomized. – B. Upper panel. Schematic of the winning computational model. Agents learn the expected value (Q) of each card based on their actions (a), the rewards (R) they receive at each trial. More specifically, agents use prediction errors (δ), the difference between expected values and actual outcomes to update the values of both the chosen and the unchosen option. Notably, the latter update depends on inference on the outcome of the counterfactual choice. The learning rate (α) determines how much recent feedback is prioritized over older feedback, the reinforcement sensitivity (ρ) determines choice stochasticity, and the double update weight (κ) scales the learning rate for counterfactual updates. The model has separate learning rates and reinforcement sensitivities for positive and negative feedback. Action selection is performed by a simple softmax rule. Middle panel. Development of expected values when $\kappa = 0$, i.e., when no counterfactual inference takes place. Lower panel. Development of expected values when $\kappa = 1$, i.e., when inferred counterfactual feedback is incorporated in the same way as actual feedback.

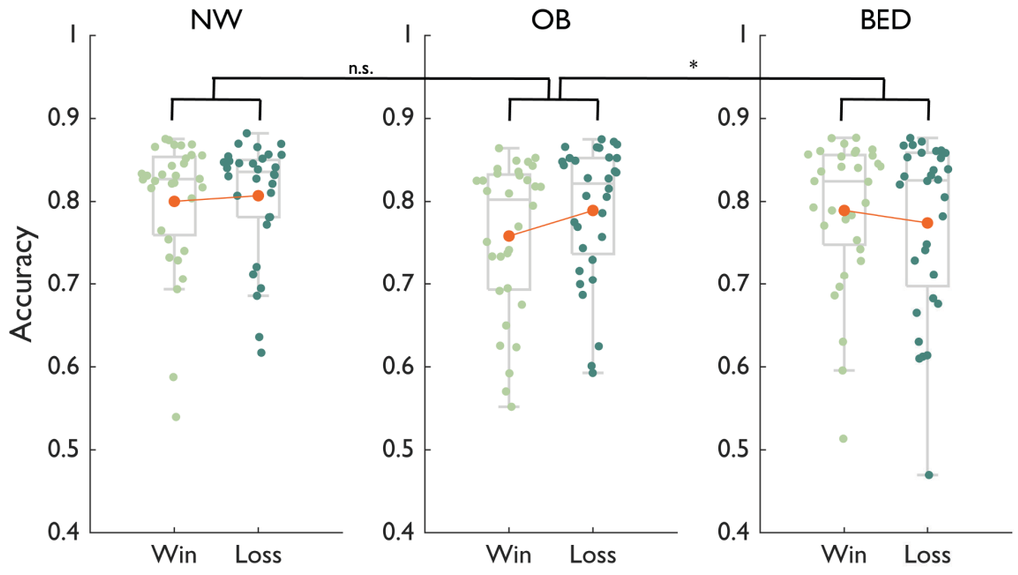
Fig. 2. – A. Accuracy in pre and post reversal trials by group. B. Choice switching in pre and post reversal trials by group. C. Difference in double update (counterfactual) learning rate for positive and negative feedback by group. D. vmPFC cluster reflecting the NW > (OB, BED) contrast on BOLD response to choice probability (t=.2.83, $p_{FWE\ svc}$ at [2, 46, -12] = .01). E. Individual parameter estimates at [2, 46, -12] by group. Individual dots represent predicted values from (generalized) linear mixed-effects models, grey boxplots reflect their distribution, yellow dots and lines indicate group means. OB – obese without binge eating disorder; BED – obese with binge eating disorder; * – p<.05; ✢ – p<.1

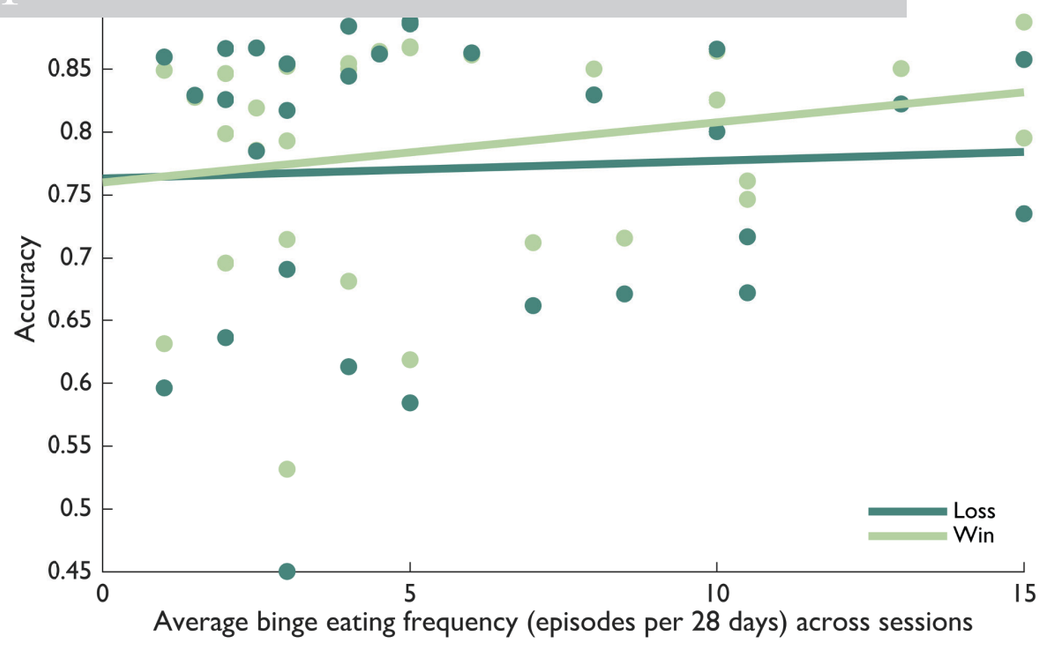
Fig. 3. – A. Accuracy by condition and group. B. Accuracy by condition and average binge-eating frequency (episodes per 28 days) within the BED group. C. Perseveration by condition and group. D. Change in choice switching by condition and change in binge-eating frequency across sessions. E. “Learning sensitivity”, the product of learning rate and reinforcement sensitivity, by condition and group. F. vmPFC cluster reflecting the single vs. double update * condition * OB vs. BED contrast on BOLD response to prediction errors (t=2.16, $p_{FWE\ svc}$ at [2, 42, -8] = .06). G. Individual parameter estimates at [2, 42, -8] by group. H. Correlation between Individual parameter estimates at [2, 42, -8] and average binge-eating frequency across sessions in the BED group. Individual dots represent predicted values from (generalized) linear mixed-effects models, grey boxplots reflect their distribution, yellow dots and lines indicate group means. OB – obese without binge eating disorder; BED – obese with binge eating disorder; * – p<.05; ✢ – p<.1



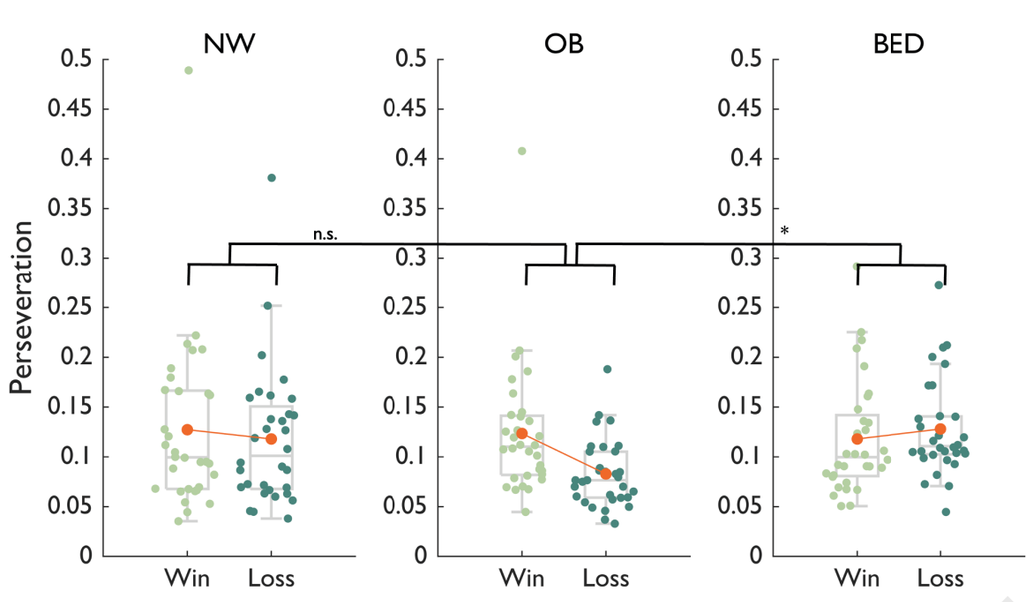


A

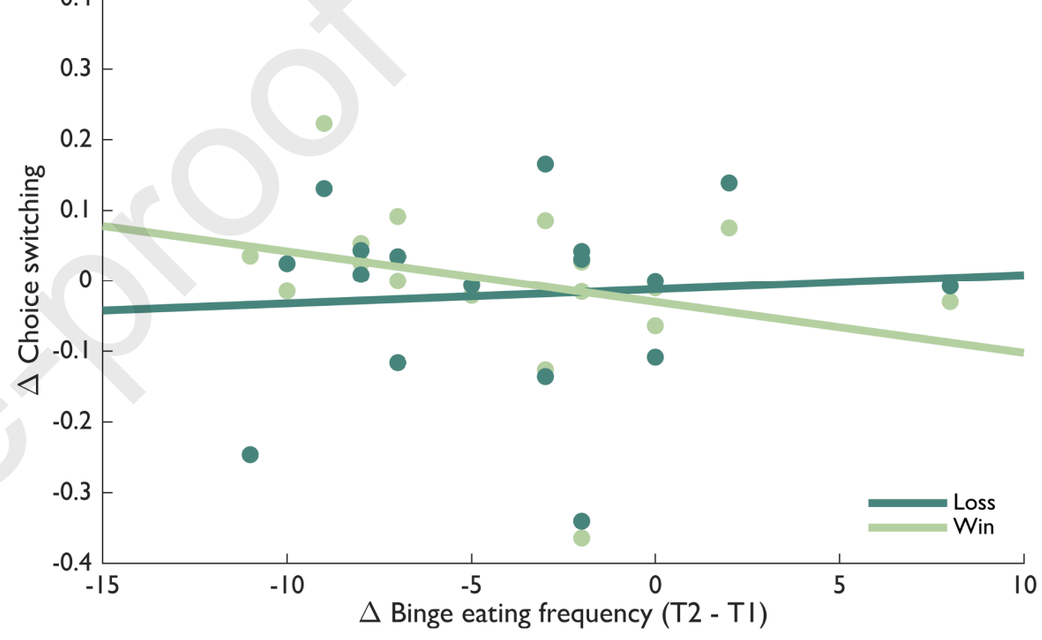




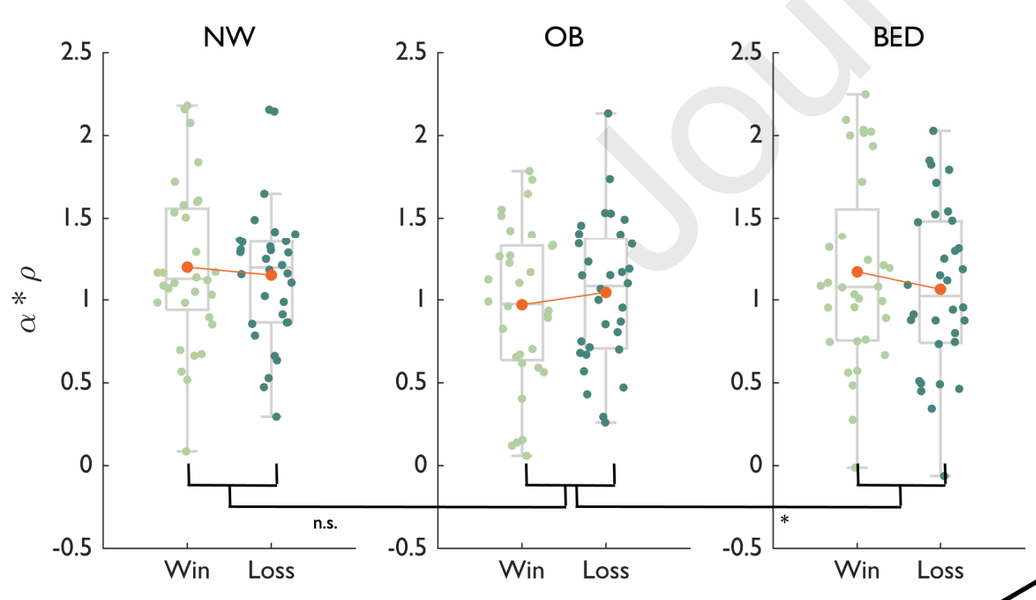
C



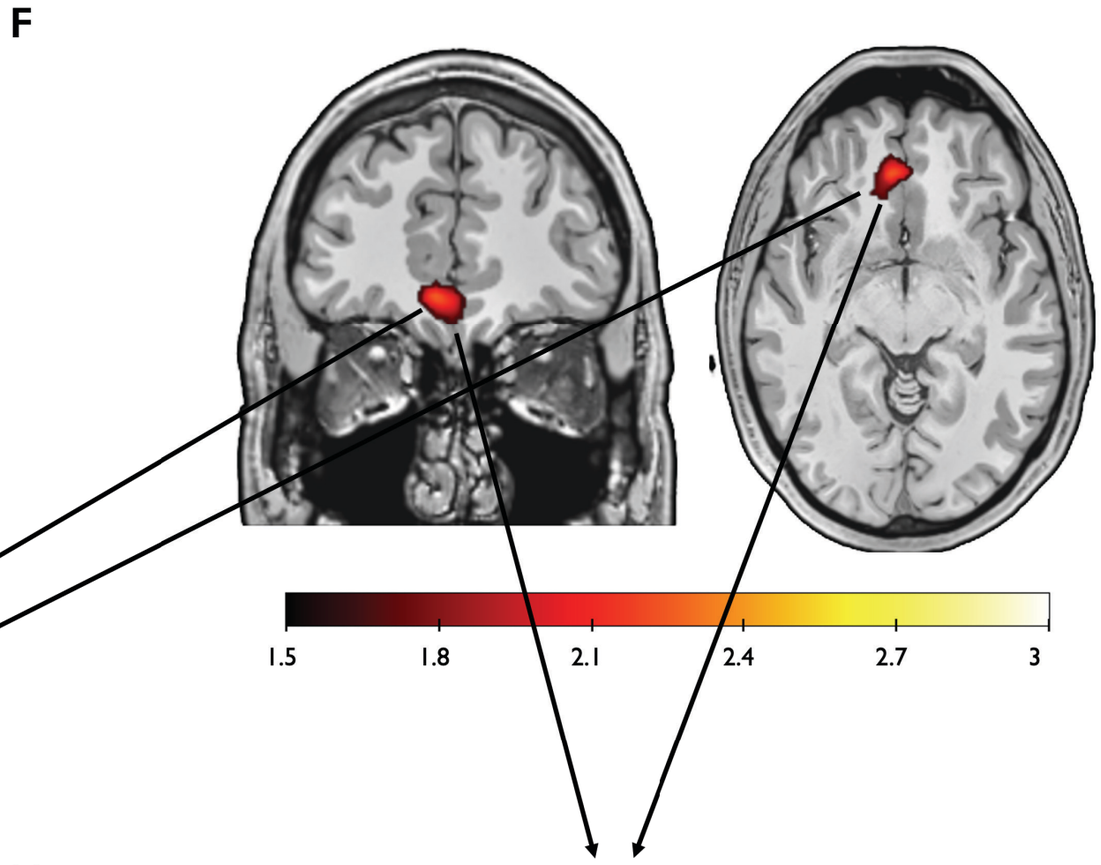
D



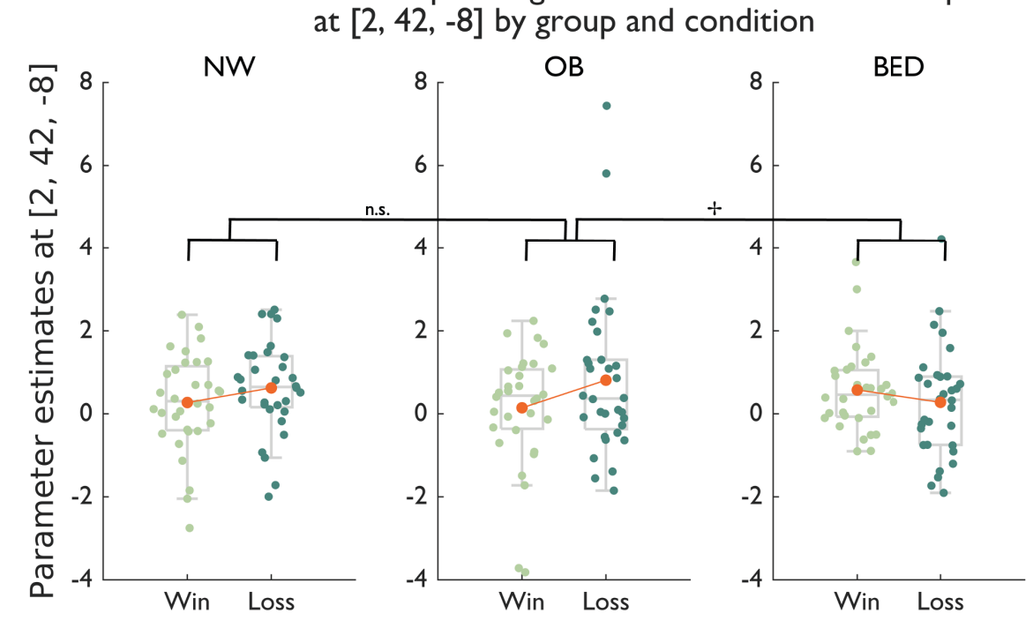
E



F



G



H