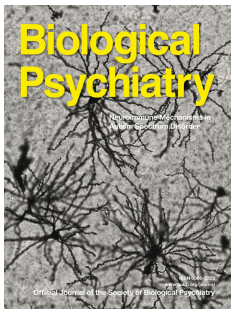
Accepted Manuscript



Altered medial frontal feedback learning signals in anorexia nervosa

Fabio Bernardoni, PhD, Daniel Geisler, MSc, Joseph A. King, PhD, Amir-Homayoun Javadi, PhD, Franziska Ritschel, MSc, Julia Murr, MD, Andrea M.F. Reiter, PhD, Veit Rössner, MD, Michael N. Smolka, MD, Stefan Kiebel, PhD, Stefan Ehrlich, MD

PII: S0006-3223(17)31899-1

DOI: 10.1016/j.biopsych.2017.07.024

Reference: BPS 13298

To appear in: *Biological Psychiatry*

Received Date: 10 January 2017

Revised Date: 12 June 2017

Accepted Date: 5 July 2017

Please cite this article as: Bernardoni F., Geisler D., King J.A., Javadi A.-H., Ritschel F., Murr J., Reiter A.M.F., Rössner V., Smolka M.N., Kiebel S. & Ehrlich S., Altered medial frontal feedback learning signals in anorexia nervosa, *Biological Psychiatry* (2017), doi: 10.1016/j.biopsych.2017.07.024.

Altered medial frontal feedback learning signals in anorexia nervosa

Fabio Bernardoni PhD,[1,2] Daniel Geisler MSc,[1,2] Joseph A. King PhD,[1,2] Amir-Homayoun Javadi PhD,[6] Franziska Ritschel MSc,[1,2] Julia Murr MD, [3] Andrea M.F. Reiter PhD, [4,5] Veit Rössner MD,[1] Michael N. Smolka MD,[7] Stefan Kiebel PhD, [4,7] Stefan Ehrlich* MD [1,2]

[1]Division of Psychological and Social Medicine and Developmental Neuroscience, Faculty of Medicine, Technische Universität Dresden, Dresden, Germany

[2]Eating Disorder Treatment and Research Center, Department of Child and Adolescent Psychiatry, Faculty of Medicine, Technische Universität Dresden, Dresden, Germany

[3]Department of Psychosomatic Therapy, Faculty of Medicine, Technische Universität Dresden, Dresden, Germany

[4]Department of Psychology, Institute of General Psychology, Biopsychology and Methods of Psychology, Technische Universität Dresden, Dresden, Germany

[5]Lifepan Developmental Neuroscience, Technische Universität Dresden

[6]School of Psychology, University of Kent, Canterbury, UK

[7]Department of Psychiatry and Neuroimaging, Technische Universität Dresden, Dresden, Germany

Keywords: probabilistic reversal learning, fMRI, anterior cingulate cortex, computational modelling, hierarchical models, Bayesian inference

Word count of the abstract: 154

Word count of the manuscript: 3992

Number of figures: 3

Number of tables: 3

Number of supplementary materials: 1

Corresponding author: Stefan Ehrlich, MD, Technische Universität Dresden, Faculty of Medicine, University Hospital C. G. Carus, Dresden, Division of Psychological and Social Medicine and Developmental Neuroscience, Fetscherstraße 74, 01307 Dresden, Germany, Phone: +49 (0)351 458-15095, Fax: +49 (0)351 458 -5754, Email: Stefan.Ehrlich@uniklinikum-dresden.de

Abstract

Background

In their relentless pursuit of thinness, individuals with anorexia nervosa (AN) engage in maladaptive behaviors (restrictive food choices, over-exercising) which may originate in altered decision-making and learning.

Methods

In this fMRI study we employed computational modelling to elucidate the neural correlates of feedback learning and value-based decision making in 36 female AN patients and 36 age-matched healthy volunteers (12-24 years). Participants performed a decision task which required adaptation to changing reward contingencies. Data were analyzed within a hierarchical Gaussian filter model, which captures inter-individual variability in learning under uncertainty.

Results

Behaviorally, patients displayed an increased learning rate specifically after punishments. At the neural level, hemodynamic correlates for learning rate, expected value and prediction error did not differ between the groups. However, activity in the posterior medial frontal cortex was elevated in AN following punishment.

Conclusion

Our findings suggest that the neural underpinning of feedback learning is selectively altered for punishment in AN.

Introduction

Anorexia nervosa (AN) is an eating disorder characterized by a relentless pursuit of thinness, mostly by self-starvation. Repeated maladaptive eating behaviors (1, 2) and extreme therapy resistance (3) in this enigmatic illness may originate from alterations in reinforcement learning such as increased sensitivity to reward or punishment and associated impairments in decision-making (4, 5). Aberrant reward-based learning in AN may reflect an entrenched "habit" of restrictive food choice (6, 7). Similarly, it has been proposed that primary rewards (food) become conditioned as punishing, and aversive stimuli (hunger) as rewarding in the brain reward system of individuals with AN (8). However, the precise mechanisms underlying response to and learning from reward and punshiment in AN are still poorly understood.

AN is consistently associated with low reward reactivity and high punishment sensitivity on clinical scales although important differences between subtypes (restrictive vs. binge-purging) may exist (9–13). Most laboratory evidence for altered feedback learning and value-based decision making in AN comes from impaired perfomance in the Iowa Gambling Task (IGT; 14, 15) - a paradigm used to measure choice behavior in the context of outcome (reward vs. punishment) uncertainty. However, reward processing is multifaceted and the typically reported IGT "net score" provides little insight into which aspect(s) might be altered in AN. Suggesting that AN patients may be particularly hypersensitive to punishment, patients have been also found to make less risky choices than healthy controls (HC) in another decision-making paradigm, the Balloon Analogue Risk Task (13). Further evidence comes from neuroimaging studies which found altered reward processing in response to disorder-related stimuli like food or taste (16–18) and secondary reinforcers (19–23). For example, neural response to punishment (monetary loss) has been found to be elevated in acutely ill adolescents in corticostriatal regions involved in valuation and action selection (21). Alteration in motivational and executive corticostriatal circuitry may also be associated with an impaired ability to flexibly adapt to change (24) and an apparently excessive amount of self-control (5, 25).

To gain a new perspective on feedback learning and decision-making in AN, we here apply the methods of computational psychiatry (26) which associate neurobiological signals with defined mechanistic steps, such as those needed to estimate the amount of reward associated with alternative behavioral options based on previous feedback. Compared to conventional analysis methods, this approach avoids i) associating neurobiological signals with subjective reports of patients (which depends on their ability to self-reflect and adequately verbalize mood states or experiences) and ii) the limitations of purely descriptive measures, such as error rates.

Intuitively, we expect healthy subjects to place greater importance on unexpected feedback in a changing environment, but to nearly disregard it in a stable one. The latter guards against

switching away from the preferred option in the presence of environmental noise, i.e. when the differences between expected and received rewards (also called reward prediction errors (27, 28)) are not due to a real change of contingencies. To probe these mechanisms in AN, we employed a reversal learning task in which the preferable choice was rewarded probabilistically (in 80% of all choices) and changed only after a learning criterion was achieved; thereby requiring participants to learn from feedback and adapt to changing reward contingencies. To analyze behavior, we compared a hierarchical Gaussian filter (HGF) model (29) with more classical reinforcement learning models (30). In the HGF model, the weight given to prediction errors is encoded in an adaptive subject-specific learning rate which is high for large environmental uncertainty, and low for small uncertainty.

Previous studies in healthy individuals (31–33) and other patient populations (34) have linked specific model parameters to activation in specific brain regions, e.g. posterior medial frontal cortex (pMFC) for learning rate, ventromedial prefrontal cortex (vmPFC) for expected (subjective) value of a choice option and ventral striatum (VS) for prediction error. Given evidence of hypersensitivity to punishment in AN (9–12, 21, 35, 36), we hypothesized that patients' decision-making would be more affected by punishments (monetary loss) relative to HC and that learning from such negative feedback would be linked to altered activation in the pMFC. The pMFC spans the dorsal anterior cingulate cortex (dACC) and pre-supplementary motor area (pre-SMA) and is broadly implicated in reward-based decision-making and signaling the need for adjustments when behavioral goals are threatened such as when losses occur (35–37).

Methods and Materials

Participants and Procedure

72 females participated in this study: 36 acutely underweight AN (12-23 years old) and 36 pairwise age-matched HC (12-24 years old). Case-control age-matching was carried out resulting in a maximum difference of 1.7 years between the individuals within one pair (Supplemental Methods). AN participants were recruited from specialized eating disorder programs and underwent MRI within 96 hours after admission to behaviorally-oriented nutritional rehabilitation programs. Please refer to

Supplemental Methods for additional information on inclusion and exclusion criteria and clinical assessments. Clinical variables are reported in Table 1.

This study was approved by the Institutional Ethics Review Board and all participants (and their guardians if underage) gave written informed consent.

One AN participant (and her age-matched partner) had to be excluded due to low performance (Supplemental Methods and Figure S1).

Experimental paradigm

We used a probabilistic reversal learning task adapted from Hampton et al., (33) (Figure 1) which includes probabilistic positive and negative monetary feedback and contingency changes according to a learning criterion (see below). In each of the 120 trials participants had to choose one of two symbols, referred to as option A and B. One symbol was designated as correct and led to monetary reward (+20cents) with a probability of 80% and to punishment (-20cents) in 20% of the cases (probabilistic errors). The choice of the 'wrong' symbol led to punishment and reward with inverted probabilities. With a probability of 25% the contingency reversed (change of the 'correct' symbol to the previously 'wrong' symbol) after at least four consecutive correct decisions since the last contingency switch.

Computational Modeling

Our computational model followed the meta-Bayesian 'observing the observer' approach (40). Accordingly, an active decision-making agent makes inferences about the hidden "state of affairs" based on the feedback associated with each option (here: the expected values of option A and B on each trial), using a so-called 'perceptual model'. Subsequently, an 'observational model' predicted the ensuing behavioral responses.

We compared the performance of three perceptual models. In addition to (i) the widely used Rescorla-Wagner model with constant learning rate, we considered two alternative models: (ii) a HGF (29) because it allowed us to quantify different forms of perceptual uncertainty perceived by the agent and (iii) a Rescorla-Wagner model with an adaptive learning rate (41). Since Bayesian Model Selection (42) revealed that the HGF fitted behavior best across HC and AN patients as well as for both groups separately (Protected Exceedance Probability>.996), it was also chosen to fit the fMRI data (Supplemental Methods and Table S1).

The HGF (29) used is a Bayesian learning model that allows for individual differences through subject-specific parameters: the *meta-volatility* (θ, 27) and the *tonic log-volatility* (ω). The *meta-volatility* determines how fast the environmental volatility is assumed to change, while the *tonic log-volatility* is a constant component of the log-volatility, and therefore has a modulating effect on the learning rate. The update equations for the expected values of each option are similar to those in basic Reinforcement Learning Models:

$$prediction(k) = prediction(k-1) + learning\ rate(k)\ \times prediction\ error(k).$$

As in previous studies (31, 33, 41, 44), we used prediction errors ($\delta^{(k)}$), implied learning rates ($\alpha^{(k)}$), and expected values of the chosen option $v^{(k)}$ as parametric modulators in the fMRI analysis.

The probability of an option to be chosen was a softmax function of its inferred expected value relative to the other option, which introduces another subject specific parameter, the *decision noise* ($1/\beta$; Figure 1).

For a precise definition of the models and their update equations, see Supplemental Methods. For the implementation and inversion of the HGF, we used the Translational Algorithms for Psychiatry-Advancing Science (TAPAS) package (http://www.translationalneuromodeling.org/tapas/) with v4.10 of the HGF toolbox (using standard priors for the free model parameters).

110

111 ***Statistical Analysis***

112 *Behavioral Measures*

113 We subjected eight measures to t-tests with group as independent factor: (i) The total
114 amount of money won, (ii) the number of misses (invalid trials), (iii) the ratio of correct responses,
115 (iv) the rate of contingency switches, (v) the log-model-evidence (LME) associated with the inversion
116 of the HGF for each subject, and the trial-independent subject-specific parameters of the
117 computational model, i.e. (vi) *log-decision noise* $\log(1/\beta)$, (vii) *tonic log-volatility* ω and (viii) *log-*
118 *meta-volatility* $\log(\theta)$.

119 The trial-dependent parameters (expected value $v^{(k)}$, prediction error $\delta^{(k)}$ and learning rate
120 $\alpha^{(k)}$) and the reaction times (RT) were treated each within a $2 \times 2 \times 2$ linear mixed model (after a
121 logit and log transform respectively; Supplemental Methods) with response (correct/wrong) and
122 feedback (rewarded/punished) as within-subject factors and group (HC/AN) as between-subject
123 factor. Post hoc t-tests were corrected for multiple comparisons using a Bonferroni-correction.

124 ***MRI Data acquisition***

125 Structural and functional images were acquired between 8 and 9 am after an overnight fast
126 using standard sequences with a 3 T whole-body MRI scanner (TRIO; Siemens, Erlangen, Germany)
127 equipped with a standard head coil (details in Supplemental Methods).

128 ***MRI Data Preprocessing***

129 Functional and structural images were processed using the SPM8 toolbox
130 (http://www.fil.ion.ucl.ac.uk/spm/) within the Nipype framework (45). Preprocessing steps included
131 correcting for slice timing and motion, normalization, smoothing, and noise reduction using CompCor
132 (46). For more details and information regarding image quality control see Supplemental Methods.

133 ***MRI Data Analysis***

134 *First level analysis*

135 In our main analysis, we implemented three different GLMs. All three models included a
136 binary and a parametric modulation regressor of interest (trial-dependent parameter of the HGF),

each associated with an event lasting for 1 second and convolved with a canonical hemodynamic response function, as in previous studies applying computational modelling in a probabilistic reversal learning task (32, 41, 44). In particular, we modulated the (GLM 1) response event (assumed to start one second before the button press) with the expected value of the chosen option $v^{(k)}$, (GLM 2) the learning event (starting at feedback) with the implied learning rate $\alpha^{(k)}$ (31, 41), and (GLM 3) the feedback event (starting at feedback) separately for rewarded and punished trials with the absolute value of the prediction error ($|\delta^{(k)}|$; 25). Follow-up analysis considered a fourth GLM with two binary regressors of interest (and no parametric modulator), starting at feedback and lasting for 1 second, separating the rewarded and the punished trials. Additional nuisance regressors in all four models were the event of stimulus presentation (lasting 0 seconds), six realignment parameters, six principal noise components from the CompCor analysis, and one regressor for each motion or intensity outlier volume.

Second level analysis

To verify that the task elicited the expected activation patterns, we first conducted whole-brain one-sample t-tests on the regression weights of the parametric modulators of the first level GLMs. To test for group differences, we then conducted independent samples t-tests on activation regressors and parametric modulators. We also implemented a whole-brain 2×2 mixed factorial ANOVA with group (AN/HC) as between- and feedback (punished/rewarded) as within-subjects factors on the 1st level coefficients from our follow-up GLM using GLMFlex (http://mrtools.mgh.harvard.edu), which allows for the estimation of partitioned errors terms.

We report results as significant at a family-wise error rate FWE level whole-brain corrected using random field theory (47) with a false-positive rate $\alpha < 0.05$. In the case of non-significant whole-brain results in any of the three *a priori* defined ROIs (Supplemental Methods and Figure S2) corresponding to the vmPFC ($v_{A,B}^{(k)}$), VS ($\delta^{(k)}$), and pMFC ($\alpha^{(k)}$), we computed small volume corrected (SVC) voxel-wise thresholds (FWE-SVC<.05).

Results

Sample Characteristics

There were no significant differences in age, IQ, or handedness score between the pairwise matched groups of AN and HC. However, as expected, AN had lower body mass index (BMI), higher eating disorder symptom and depression scores (Table 1). Differences in the Behavioral Inhibition Scale (BIS) or Junior Temperament and Character Inventory subscale 'harm avoidance' (HA) were not significant in the sample with neuroimaging data. However, in a larger sample with questionaire data, that included the one used for the present study, AN patients had a significantly higher BIS and HA (Supplemental Results).

Behavioral and Modeling Data

The results of the ANOVA on behavioral measures and on trial independent model parameters (and of the Mann-Whitney test on ω) are summarized in Table 2. There were no group differences for the number of correct answers and contingency reversals, for the total win and the number of misses. The LME and the subject-specific model parameters (inverse log-decision noise $\log(\beta)$, *tonic log-volatility* ω and *log-meta-volatility* $\log(\theta)$) also did not differ between the groups.

The results of the 2(HC/AN)×2(rewarded/punished)×2(correct/wrong) mixed model on the trial dependent model parameters and the reaction times are summarized in Table 3 (see also Table S5). The expected main effects and interactions of feedback and response on the learning rate, the prediction error and the expected value were reproduced [(44, 48); Supplemental Results]. Most importantly, a group×feedback interaction indicating a higher learning rate on punished trials in AN was found [$F(1,8262.6)=6.6$, $p=0.010$; Figure 2]. This effect was not influenced by age (Supplemental Results, Table S4). Further explorative analyses indicated that increased learning rate after punishment in AN is not driven by HA or extreme underweight (Supplemental Results, Table S6).

Imaging Data

In line with previous studies (31), BOLD activity in the pMFC correlated with the changing (time-dependent) learning rate $\alpha^{(k)}$ (Figures 3A, S5). Also as in previous studies (32, 33), activation in the vmPFC correlated with the changing expected value $v^{(k)}$ (Figure S3). Furthermore, BOLD activation in the VS correlated with the changing prediction error $|\delta^{(k)}|$ separately in rewarded and punished trials [Figure S3, (32, 33, 41, 44)]. Together, these findings corroborate our task and

193 analytical approach. Other significant activations are reported in Table S4. No group differences were
194 found at FWE or FWE-SVC level.
195 More important regarding our hypotheses, given (i) the behavioral findings indicative of an increased
196 learning rate in AN on punished trials (Figure 2), (ii) previous evidence of elevated sensitivity to
197 punishment in AN (9, 12), and (iii) the linear correlation between learning rate and BOLD activity in
198 pMFC as in previous studies (31, 41), we predicted altered activation in AN in the region associated
199 with learning rate, specifically after punishments. To test this hypothesis, we calculated a 2(group)
200 x2(feedback) ANOVA. Critically, while no group difference in the pMFC was revealed on win trials, the
201 BOLD response was elevated in this region in AN on punished trials. This group difference overlapped
202 the cluster in which BOLD activity correlated with learning rate (Figures 3B, S4, Table S8; see also
203 Figure S5). To investigate possible causal relationships, we conducted mediation analysis using the
204 SPSS PROCESS toolbox (49). However, no mediation effects of the learning rate on the pMFC
205 activation or vice versa were detected (Supplemental Results, Table S9). Moreover, exploratory
206 analysis revealed no correlation between pMFC activation and BMI-SDS, BDI-II, EDI-2 or HA scores in
207 AN (FWE-SVC).

Discussion

We used computational modelling in combination with fMRI to provide insight into the neural mechanisms underlying decision-making and feedback learning in young, acutely ill AN patients. Bayesian Model Comparison (Methods) demonstrated better fit between a recently developed HGF model (29) and the behavioral data for both the AN and HC groups than more classical reinforcement learning models (30). However, AN patients were characterized by an increased learning rate on punished trials; possibly indicating hypersensitivity to punishment which has been observed clinically and empirically in AN (10, 12, 35). This finding suggests that when AN patients experience negative feedback, they question their beliefs to a greater degree than HC. On a neural level, time-dependent parameters of feedback learning correlated with BOLD activity in the same brain regions in both groups. In particular, consistent with previous model-based fMRI studies of decision-making and feedback-learning in healthy participants (31, 41), we found a significant correlation between learning rate and BOLD activation in the pMFC, a region involved in outcome evaluation and initiating adaptive adjustments accordingly (31, 38, 50). Most importantly, mirroring the behavioral group difference, BOLD activation was increased in this region in AN after punishment.

Our finding of increased pMFC activation after punishment in AN converges with recent evidence attributing a role of this region to the pathophysiology of the disorder. For example adolescent AN patients exhibited an elevated neural response to punishment in the “cognitive” zone of the dACC relative to HC in a monetary guessing task. (21). Conversely, Zastrow et al. (24) found decreased pMFC activation specifically on “shift” trials of a target detection task in AN. Altered pMFC activity has also been reported during temporal reward discounting (19, 51) and during inhibitory processing (52). Moreover, a recent resting-state functional connectivity study (53), found reduced connectivity between pMFC and the executive control network in adolescent AN. While these studies suggest altered pMFC functioning in AN, the direction of group differences vary and the possible interpretations range from altered conflict monitoring, excessive cognitive control and increased neural efficiency. Structurally, volume reductions in the ACC (including portions of the pMFC) in acutely ill AN have been related to deficits in perceptual organization and conceptual reasoning, while the degree of normalization during treatment was linked to clinical outcome (54). Using SPECT, reduced regional cerebral blood flow in the dACC extending into the pre-SMA was observed during the acute phase of the illness and after weight recovery (55). Our study gives additional support for

functional pMFC alterations in acutely ill AN using a novel approach that had been applied successfully in other disorders before (42–44). Taken together, our behavioral and imaging findings suggest thatthe elevated pMFC response in AN may help to explain the abnormally rapid learning rate following punishment.

Restrictive food choice and extreme resistance to treatment are just two examples of altered decision-making in AN. While previous laboratory investigations (14, 15) were relatively limited in their ability to isolate specific alterations, a recent cognitive modelling study of IGT performance found a "recency bias" in AN captured by a learning/memory parameter (58). Although the model did not uncover a group difference in a feedback sensitivity parameter, the finding that patients tended to base their decisions on recent experience is commensurate with our finding of increased learning rate in AN. The current evidence of altered decision-making in response to negative feedback is in line with notion of altered reinforcement learning in AN (1–5, 8) and, considered in light of similar recent findings (13), is suggestive of a particular sensitivity to punishment. Decision-making may be intact, however, in paradigms that don't include negative feedback, at least in adolescents (19, 59). Nonetheless, these findings were made in predominately restrictive AN and future studies are needed to clarify potential subtype differences in reward and punishment sensitivity (10, 11). Furthermore, given the presumption that AN is characterized by altered general reward-related decision-making (4, 8, 19) and the lack of group differences in this respect in both the current study and other recent ones (21, 51), future research is also needed to clarify under which conditions the neural substrates of reward processing are aberrant in AN.

While our study was not designed to clarify whether altered decision-making causes AN or is a temporary effect of acute illness, correlation between punishment sensitivity and attachment insecurity has been reported (60). This suggests that, together with attachment style, a decision-making strategy geared toward loss avoidance may develop early in life. Speculatively, oversensitivity to negative feedback may contribute to the onset of AN. For example, negative comments from peers regarding physical appearance might be given exaggerated importance as an effect of an increased learning rate, and consequently, predispose (future) AN patients to change their nutritional habits and activity levels to lose weight (61). Indeed, it has been found that increased HA persists after recovery in AN, raising the possibility that such a trait exists premorbidly (62, 63).

At the neurobiological level, PET imaging studies found associations between HA and 5-HT functioning in various eating disorders (62). Interestingly, a low 5-HT state, probably due to reduced

tryptophan intake because of food restriction (63–65) has been suggested for acute AN (62). In healthy participants (66), it was found that acute tryptophan depletion (ATD), a method for transiently reducing cerebral 5-HT levels, was associated with increased BOLD responses in a region of the dorsomedial PFC overlapping the pMFC during a probabilistic reversal learning task, especially after punishment. Given the role of 5-HT in altered neural mechanisms during feedback learning and evidence suggesting normal or even increased 5-HT levels in recovered AN (62, 67), future studies in weight-recovered AN targeting the pMFC during feedback learning are of great interest.

At a more qualitative level, our model-based approach suggests that learning and decision-making activate the same brain regions similarly in both AN and HC. This finding fits neatly with our model comparison: by using different computational models of feedback learning, we found that the behavior of both groups was better explained by the Bayesian HGF model than Rescorla-Wagner models (either with fixed or flexible learning rate) suggesting that, equally to controls, AN patients place differential importance on prediction errors depending on their perception of environmental volatility. Note that for other psychiatric disorders such as binge eating disorder (57), schizophrenia (68) or alcoholism (69), Bayesian Model Selection indicated that patients' behavior was guided by different (typically less efficient) decision-making strategies. For example, in adolescent ADHD, patients choice behavior was better explained by a Rescorla-Wagner model with constant learning rate whereas for HC the HGF provided a better fit (56). Previous computational modeling studies in AN (16, 70) used a temporal difference model with a fixed learning rate (28) to derive prediction error measures in passive taste reward learning tasks, but model parameters and model comparison data were not reported in these studies.

Our study has to be seen in the light of the following limitations: First, we focused on young (mostly adolescent) patients with acute AN. While this has the advantage of minimizing secondary effects of prolonged malnutrition on cognition, it provides no indication whether parameters such as the learning rate can be seen as biological markers. Therefore, studies measuring patients longitudinally after weight restoration or complete recovery are needed. However, although patients were in a state of undernutrition, they did not show reduced performance and the behavioral results were not driven by particularly underweight patients (Supplemental Results, Table S6). Second, although we compared three computational models of behavior and identified one with best fit for both groups (suggesting that the general strategies employed in AN are normal), there may be better models that lead to different conclusions. Third, although our sample size was large relative to most

fMRI studies in AN and the employed task had a comparable number of trials as in similar clinical studies (21), the power of our study to detect all relevant between-group effects (e.g. reward-related) may be limited and future studies with more observations in larger samples are needed. Fourth, the group difference in self-reported HA was not significant in the present study, presumably because of lack of statistical power (Supplemental Results), and the expected correlation between HA and learning rate after punishment was not found (Supplemental Results). Therefore, alternative explanations of increased learning rate in AN inlcuding impaired memory (58) and uncertainty regarding present beliefs are also plausible. However, an increased learning rate specifically after punishments indicates that an exaggerated importance is placed to negative feedback, despite uncertainty due to the probabilistic nature of contingencies.

Computational approaches focusing on learning mechanisms appear to be particularly promising with respect to the detection of basic mechanisms contributing to the development and maintenance of mental disorders. Altered decision-making has been linked to treatment outcome in AN (71) and quantification of individual differences in learning mechanisms have the potential to guide the development of new therapeutic strategies that directly aim at the modification of such behavior patterns. Given the present results in patients with acute AN, a stronger focus on increasing self-confidence (72) and the ability to tolerate criticism might foster therapeutic success.

Acknowledgments and Disclosures

This work was supported by the Deutsche Forschungsgemeinschaft (EH 367/5-1, SFB 940/1, SM 80/5-2, SM 80/7-1, SM 80/7-2) and the Swiss Anorexia Nervosa Foundation. The authors would like to express their gratitude to Laura Soltwedel, Benjamin Roschinski, Juliane Petermann, Luisa Flohr, Eva Seeger, Lea Scheuvens, Juliane Hantke, Stefanie Huber, Isabelle Hennig, Richard Vettermann, Sabine Clas, Marion Breier, Johannes Zwipp, and Constanze Nicklisch for their assistance with participant recruitment and data collection and thank all participants for their time and cooperation. We thank the Center for Information Services and High Performance Computing (ZIH) at TU Dresden for generous allocations of computer time.

In the last two years, Dr. Roessner has received payment for consulting and writing activities from Lilly, Novartis, and Shire Pharmaceuticals, lecture honoraria from Lilly, Novartis, Shire Pharmaceuticals, and Medice Pharma, and support for research from Shire and Novartis. He has carried out (and is currently carrying out) clinical trials in cooperation with the Novartis, Shire, and Otsuka companies. All other authors reported no biomedical financial interests or potential conflicts of interest.

References

1. Compan V, Walsh BT, Kaye W, Geliebter A (2015): How Does the Brain Implement Adaptive Decision Making to Eat? *J Neurosci*. 35: 13868–13878.

2. Foerde K, Steinglass JE, Shohamy D, Walsh BT (2015): Neural mechanisms supporting maladaptive food choices in anorexia nervosa. *Nat Neurosci*. 18: 1571–1573.

3. Steinglass JE, Foerde K (2015): How does anorexia nervosa become resistant to change? *Manag Sev Endur Anorex Nerv*. Routledge.

4. O'Hara CB, Campbell IC, Schmidt U (2015): A reward-centred model of anorexia nervosa: a focussed narrative review of the neurological and psychophysiological literature. *Neurosci Biobehav Rev*. 52: 131–152.

5. Kaye WH, Wierenga CE, Bailer UF, Simmons AN, Bischoff-Grethe A (2013): Nothing Tastes as Good as Skinny Feels: The Neurobiology of Anorexia Nervosa. *Trends Neurosci*. 36. doi: 10.1016/j.tins.2013.01.003.

6. Walsh BT (2013): The enigmatic persistence of anorexia nervosa. *Am J Psychiatry*. 170: 477–484.

7. Steinglass JE, Walsh BT (2016): Neurobiological model of the persistence of anorexia nervosa. *J Eat Disord*. 4: 19.

8. Keating C (2010): Theoretical perspective on anorexia nervosa: the conflict of reward. *Neurosci Biobehav Rev*. 34: 73–79.

9. Harrison A, Sternheim L, O'Hara C, Oldershaw A, Schmidt U (2016): Do reward and punishment sensitivity change after treatment for anorexia nervosa? *Personal Individ Differ*. 96: 40–46.

10. Glashouwer KA, Bloot L, Veenstra EM, Franken IHA, de Jong PJ (2014): Heightened sensitivity to punishment and reward in anorexia nervosa. *Appetite*. 75: 97–102.

11. Wierenga CE, Ely A, Bischoff-Grethe A, Bailer UF, Simmons AN, Kaye WH (2014): Are Extremes of Consumption in Eating Disorders Related to an Altered Balance between Reward and Inhibition? *Front Behav Neurosci*. 8: 410.

12. Jappe LM, Frank GKW, Shott ME, Rollin MDH, Pryor T, Hagman JO, *et al.* (2011): Heightened Sensitivity to Reward and Punishment in Anorexia Nervosa. *Int J Eat Disord*. 44: 317–324.

13. Adoue C, Jaussent I, Olié E, Beziat S, Van den Eynde F, Courtet P, Guillaume S (2015): A further assessment of decision-making in anorexia nervosa. *Eur Psychiatry J Assoc Eur Psychiatr*. 30: 121–127.

14. Guillaume S, Gorwood P, Jollant F, Van den Eynde F, Courtet P, Richard-Devantoy S (2015): Impaired decision-making in symptomatic anorexia and bulimia nervosa patients: a meta-analysis. *Psychol Med*. 45: 3377–3391.

15. Wu M, Brockmeyer T, Hartmann M, Skunde M, Herzog W, Friederich H-C (2016): Reward-related decision making in eating and weight disorders: A systematic review and meta-analysis of the evidence from neuropsychological studies. *Neurosci Biobehav Rev*. 61: 177–196.

16. Frank GKW, Reynolds JR, Shott ME, Jappe L, Yang TT, Tregellas JR, O'Reilly RC (2012): Anorexia Nervosa and Obesity are Associated with Opposite Brain Reward Response. *Neuropsychopharmacology*. 37: 2031–2046.

17. Cowdrey FA, Park RJ, Harmer CJ, McCabe C (2011): Increased neural processing of rewarding and aversive food stimuli in recovered anorexia nervosa. *Biol Psychiatry*. 70: 736–743.

18. Kaye WH, Wierenga CE, Bailer UF, Simmons AN, Wagner A, Bischoff-Grethe A (2013): Does a shared neurobiology for foods and drugs of abuse contribute to extremes of food ingestion in anorexia and bulimia nervosa? *Biol Psychiatry*. 73: 836–842.

19. Decker JH, Figner B, Steinglass JE (2015): On Weight and Waiting: Delay Discounting in Anorexia Nervosa Pretreatment and Posttreatment. *Biol Psychiatry*. 78: 606–614.

20. Wagner A, Aizenstein H, Venkatraman VK, Fudge J, May JC, Mazurkewicz L, *et al.* (2007): Altered Reward Processing in Women Recovered From Anorexia Nervosa. *Am J Psychiatry*. 164: 1842–1849.

21. Bischoff-Grethe A, McCurdy D, Grenesko-Stevens E, Irvine LEZ, Wagner A, Yau W-YW, *et al.* (2013): Altered brain response to reward and punishment in adolescents with Anorexia nervosa. *Psychiatry Res*. 214: 331–340.

22. Wierenga CE, Bischoff-Grethe A, Melrose AJ, Irvine Z, Torres L, Bailer UF, *et al.* (2015): Hunger does not motivate reward in women remitted from anorexia nervosa. *Biol Psychiatry*. 77: 642–652.

23. Via E, Soriano-Mas C, Sánchez I, Forcano L, Harrison BJ, Davey CG, *et al.* (2015): Abnormal Social Reward Responses in Anorexia Nervosa: An fMRI Study. *PLoS ONE*. 10. doi: 10.1371/journal.pone.0133539.

24. Zastrow A, Kaiser S, Stippich C, Walther S, Herzog W, Tchanturia K, *et al.* (2009): Neural correlates of impaired cognitive-behavioral flexibility in anorexia nervosa. *Am J Psychiatry*. 166: 608–616.

25. Ehrlich S, Geisler D, Ritschel F, King JA, Seidel M, Boehm I, *et al.* (2015): Elevated cognitive control over reward processing in recovered female patients with anorexia nervosa. *J Psychiatry Neurosci JPN*. 40: 140249.

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

27. Bush RR, Mosteller F (2006): A Mathematical Model for Simple Learning. In: Fienberg SE, Hoaglin DC, editors. *Sel Pap Frederick Most*, Springer Series in Statistics. Springer New York, pp 221–234.

28. Sutton RS, Barto AG (1998): *Introduction to Reinforcement Learning*, 1st ed. Cambridge, MA, USA: MIT Press.

29. Mathys C, Daunizeau J, Friston KJ, Stephan KE (2011): A bayesian foundation for individual learning under uncertainty. *Front Hum Neurosci*. 5: 39.

30. Rescorla R, Wagner A (1972): A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black A, Prokasy W, editors. *Class Cond II Curr Res Theory*. Appleton-Century-Crofts, pp 64–99.

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

32. Gläscher 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 N Y N 1991*. 19: 483–495.

33. 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 Off J Soc Neurosci*. 26: 8360–8367.

34. Deserno L, Schlagenhauf F, Heinz A (2016): Striatal dopamine, reward, and decision making in schizophrenia. *Dialogues Clin Neurosci*. 18: 77–89.

35. Harrison A, O'Brien N, Lopez C, Treasure J (2010): Sensitivity to reward and punishment in eating disorders. *Psychiatry Res*. 177: 1–11.

36. Geisler D, Ritschel F, King JA, Bernardoni F, Seidel M, Boehm I, *et al.* (2017): Increased anterior cingulate cortex response precedes behavioural adaptation in anorexia nervosa. *Sci Rep*. 7: 42066.

37. Jocham G, Neumann J, Klein TA, Danielmeier C, Ullsperger M (2009): Adaptive coding of action values in the human rostral cingulate zone. *J Neurosci Off J Soc Neurosci*. 29: 7489–7496.

38. Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004): The role of the medial frontal cortex in cognitive control. *Science*. 306: 443–447.

39. Ullsperger M, Danielmeier C, Jocham G (2014): Neurophysiology of performance monitoring and adaptive behavior. *Physiol Rev*. 94: 35–79.

40. Daunizeau J, Ouden HEM den, Pessiglione M, Kiebel SJ, Stephan KE, Friston KJ (2010): Observing the Observer (I): Meta-Bayesian Models of Learning and Decision-Making. *PLOS ONE*. 5: e15554.

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

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

43. Diaconescu AO, Mathys C, Weber LAE, Daunizeau J, Kasper L, Lomakina EI, *et al.* (2014): Inferring on the Intentions of Others by Hierarchical Bayesian Learning. *PLOS Comput Biol*. 10: e1003810.

44. Javadi AH, Schmidt DHK, Smolka MN (2014): Adolescents Adapt More Slowly than Adults to Varying Reward Contingencies. *J Cogn Neurosci*. 26: 2670–2681.

45. Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, Ghosh SS (2011): Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Front Neuroinformatics*. 5: 13.

46. Behzadi Y, Restom K, Liau J, Liu TT (2007): A Component Based Noise Correction Method (CompCor) for BOLD and Perfusion Based fMRI. *NeuroImage*. 37: 90–101.

47. Worsley KJ, Evans AC, Marrett S, Neelin P (1992): A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab Off J Int Soc Cereb Blood Flow Metab*. 12: 900–918.

48. Chase HW, Kumar P, Eickhoff SB, Dombrovski AY (2015): Reinforcement learning models and their neural correlates: An activation likelihood estimation meta-analysis. *Cogn Affect Behav Neurosci*. 15: 435–459.

49. Hayes AF (2013): *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach*. Methodology in the social sciences. New York, NY: Guilford Press.

50. Rushworth MFS, Behrens TEJ (2008): Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci*. 11: 389–397.

51. King JA, Geisler D, Bernardoni F, Ritschel F, Böhm I, Seidel M, *et al.* (2016): Altered Neural Efficiency of Decision Making During Temporal Reward Discounting in Anorexia Nervosa. *J Am Acad Child Adolesc Psychiatry*. 55: 972–979.

52. Wierenga C, Bischoff-Grethe A, Melrose AJ, Grenesko-Stevens E, Irvine Z, Wagner A, *et al.* (2014): Altered BOLD Response during Inhibitory and Error Processing in Adolescents with Anorexia Nervosa. (C. Soriano-Mas, editor) *PLoS ONE*. 9: e92017.

53. Gaudio S, Piervincenzi C, Beomonte Zobel B, Romana Montecchi F, Riva G, Carducci F, Cosimo Quattrocchi C (2015): Altered resting state functional connectivity of anterior cingulate cortex in drug naïve adolescents at the earliest stages of anorexia nervosa. *Sci Rep*. 5. doi: 10.1038/srep10818.

54. McCormick LM, Keel PK, Brumm MC, Bowers W, Swayze V, Andersen A, Andreasen N (2008): Implications of starvation-induced change in right dorsal anterior cingulate volume in anorexia nervosa. *Int J Eat Disord*. 41: 602–610.

55. Kojima S, Nagai N, Nakabeppu Y, Muranaga T, Deguchi D, Nakajo M, *et al.* (2005): Comparison of regional cerebral blood flow in patients with anorexia nervosa before and after weight gain. *Psychiatry Res*. 140: 251–258.

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

57. Reiter A, Heinze H-J, Schlagenhauf F, Deserno L (2016): Impaired Flexible Reward-Based Decision-Making in Binge Eating Disorder: Evidence from Computational Modeling and Functional Neuroimaging. *Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol*. . doi: 10.1038/npp.2016.95.

58. Chan TWS, Ahn W-Y, Bates JE, Busemeyer JR, Guillaume S, Redgrave GW, *et al.* (2014): Differential impairments underlying decision making in anorexia nervosa and bulimia nervosa: a cognitive modeling analysis. *Int J Eat Disord*. 47: 157–167.

59. Ritschel F, King JA, Geisler D, Flohr L, Neidel F, Boehm I, *et al.* (2015): Temporal delay discounting in acutely ill and weight-recovered patients with anorexia nervosa. *Psychol Med*. 45: 1229–1239.

60. Keating C, Castle DJ, Newton R, Huang C, Rossell SL (2016): Attachment Insecurity Predicts Punishment Sensitivity in Anorexia Nervosa. *J Nerv Ment Dis*. . doi: 10.1097/NMD.0000000000000569.

61. Fabian LJ, Thompson JK (1989): Body image and eating disturbance in young females. *Int J Eat Disord*. 8: 63–74.

62. Kaye W (2008): Neurobiology of anorexia and bulimia nervosa. *Physiol Behav*. 94: 121–135.

63. Wagner A, Barbarich-Marsteller NC, Frank GK, Bailer UF, Wonderlich SA, Crosby RD, *et al.* (2006): Personality traits after recovery from eating disorders: do subtypes differ? *Int J Eat Disord*. 39: 276–284.

64. Ehrlich S, Franke L, Schneider N, Salbach-Andrae H, Schott R, Craciun EM, *et al.* (2009): Aromatic amino acids in weight-recovered females with anorexia nervosa. *Int J Eat Disord*. 42: 166–172.

65. Fernstrom JD, Wurtman RJ (1972): Brain serotonin content: physiological regulation by plasma neutral amino acids. *Science*. 178: 414–416.

66. Evers EAT, Cools R, Clark L, van der Veen FM, Jolles J, Sahakian BJ, Robbins TW (2005): Serotonergic modulation of prefrontal cortex during negative feedback in probabilistic reversal learning. *Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol*. 30: 1138–1147.

67. Frank GK, Kaye WH, Meltzer CC, Price JC, Greer P, McConaha C, Skovira K (2002): Reduced 5-HT2A receptor binding after recovery from anorexia nervosa. *Biol Psychiatry*. 52: 896–906.

68. Schlagenhauf F, Huys QJM, Deserno L, Rapp MA, Beck A, Heinze H-J, *et al.* (2014): Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *NeuroImage*. 89: 171–180.

69. Reiter A, 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 Off J Soc Neurosci*. 36: 10935–10948.

70. DeGuzman M, Shott ME, Yang TT, Riederer J, Frank GKW (2017): Association of Elevated Reward Prediction Error Response With Weight Gain in Adolescent Anorexia Nervosa. *Am J Psychiatry*. appiajp201616060671.

71. Cavedini P, Zorzi C, Bassi T, Gorini A, Baraldi C, Ubbiali A, Bellodi L (2006): Decision-making functioning as a predictor of treatment outcome in anorexia nervosa. *Psychiatry Res*. 145: 179–187.

72. Wild B, Friederich H-C, Zipfel S, Resmark G, Giel K, Teufel M, *et al.* (2016): Predictors of outcomes in outpatients with anorexia nervosa - Results from the ANTOP study. *Psychiatry Res*. 244: 45–50.

Figure Legends

Figure 1. Top: Time course of the experiment. First, two abstract stimuli were presented. The participant had up to 2s time to make a choice. After the participant had selected one stimulus (by left or right button press), a fixation cross was presented for 4s. Finally, positive or negative feedback (monetary reward or punishment) was displayed for 1s followed by a jittered inter-trial interval (fixation cross) for 4 to 8s. **Bottom left**: **The Hierarchical Gaussian Filter (HGF).** Graphical representation of the perceptual (HGF) model used in this work. Polygons represent quantities that change with time, while circles denote time-independent, subject-specific parameters. Arrows indicate dependency of one variable on another. While hexagons represent states that satisfy the Markov property, such that the state at trial k also depends on the state at $k-1$, diamonds contain quantities that do change with time, but do not depend on their previous state. β is the inverse *decision noise*, θ the *meta-volatility* and ω the *tonic log-volatility*. x_1 is the probability of reward for each option A and B, x_2 is the tendency towards reward and x_3 is the time-dependent part of the log-volatility. y are the responses given by the participant. In our observational model y does not depend directly on the environmental volatility x_3. **Bottom right**: **The softmax choice rule.** Probability that option A is chosen according to the observational model used in this work (softmax). $v_A^{(k)} - v_B^{(k)}$ can be computed from x_1, see Supplemental Methods. A small value of *decision noise* $(1/\beta)$ implies that the most valuable option is chosen with high probability. The β values chosen correspond to the mean on the entire sample plus minus the standard deviation (see Table 2).

Figure 2. Increased learning rate after punishment in AN. The critical group×feedback interaction (significant also after Bonferroni correction across the four tested models p(corrected) = 0.04) was followed up with post-hoc comparisons which revealed that learning rate is greater in AN than in HC on punished trials (mean difference (SE) = 0.083(0.036)). Error bars reflect 95% confidence level intervals.

Figure 3. A: Correlation of BOLD activity after feedback with learning rate α. Learning rate was computed within a Hierarchical Gaussian Filter and the expected pattern of activation in the pMFC (31, 41) across all participants (whole-brain one-sample t-test) was reproduced. **B: Increased BOLD activity in AN following punishment.** Increased BOLD activity in AN relative to HC following punishment as revealed by a whole-brain independent samples t-test is depicted on the same slice. A list with the peaks of activation is reported in Table S4. We display regions where the signal is significant at a FWE<.05 level determined with random field theory. The color scale shows one sample t-test values.

Tables

Table 1. Group characteristics. Comparisons of demographic and clinical variables were examined using independent two-sample t-tests, differences in task relevant variables were examind using one-way ANCOVAs controlling for IQ. Means and standard deviations (SD) are given.

	AN		**HC**		**test statistics**	
	Mean	SD	Mean	SD		
Demographic variables					**T**	**p**
Age	16.0	2.6	16.3	2.6	-0.5	0.662
BMI	14.7	1.3	20.4	2.5	-12.0	<0.001
BMI-SDS	-2.1	0.6	0.0	0.8	-11.7	<0.001
IQ	111.9	11.1	110.9	10.0	0.4	0.673
Handedness	0.5	2.0	1.7	3.7	-1.8	0.081
Clinical variables					**T**	**p**
EDI-2 total score	197.4	50.7	139.6	28.0	5.9	<0.001
EDI-2 perfectionism	19.6	6.0	15.7	4.2	3.3	0.002
BDI-II total score	19.5	11.6	5.5	5.7	6.5	<0.001
BIS	22.0	3.7	20.8	3.3	1.12	0.269
BAS	39.8	6.3	40.5	4.2	-0.44	0.665
JTCI harm avoidance	37.3	11.5	34.1	8.0	1.36	0.178
SCL-90-R	74.9	59.8	28.6	26.8	17.4	<0.001

AN=anorexia nervosa patients; HC=healthy controls; BMI-SDS=body mass index standard deviation score; IQ=intelligence quotient; EDI-2=Eating disorder inventory; BDI-II=Beck Depression Inventory; SCL-90-R = revised Symptom Checklist 90, BIS-BAS= behavioral avoidance/inhibition (BIS/BAS) scales, computed on a sample of 19 AN and 21 HC, JTCI=Junior Temperament und Character Inventory values, computed on a sample of 34 AN and

35 HC. 32 patients were of restrictive subtype and 3 of binge-purge. *P*-values below 0.05 indicates a significant group difference.

Table 2. ANOVA on trial independent parameters. The individual parameters from the HGF perceptual model and softmax observational model were subjected to an ANOVA with group as independent factor. Group means and standard deviations (SD) are given. For the *tonic log-volatility* (ω), a Mann-Whitney test found no group differences (U=612.5, p(2-tailed)=0.089).

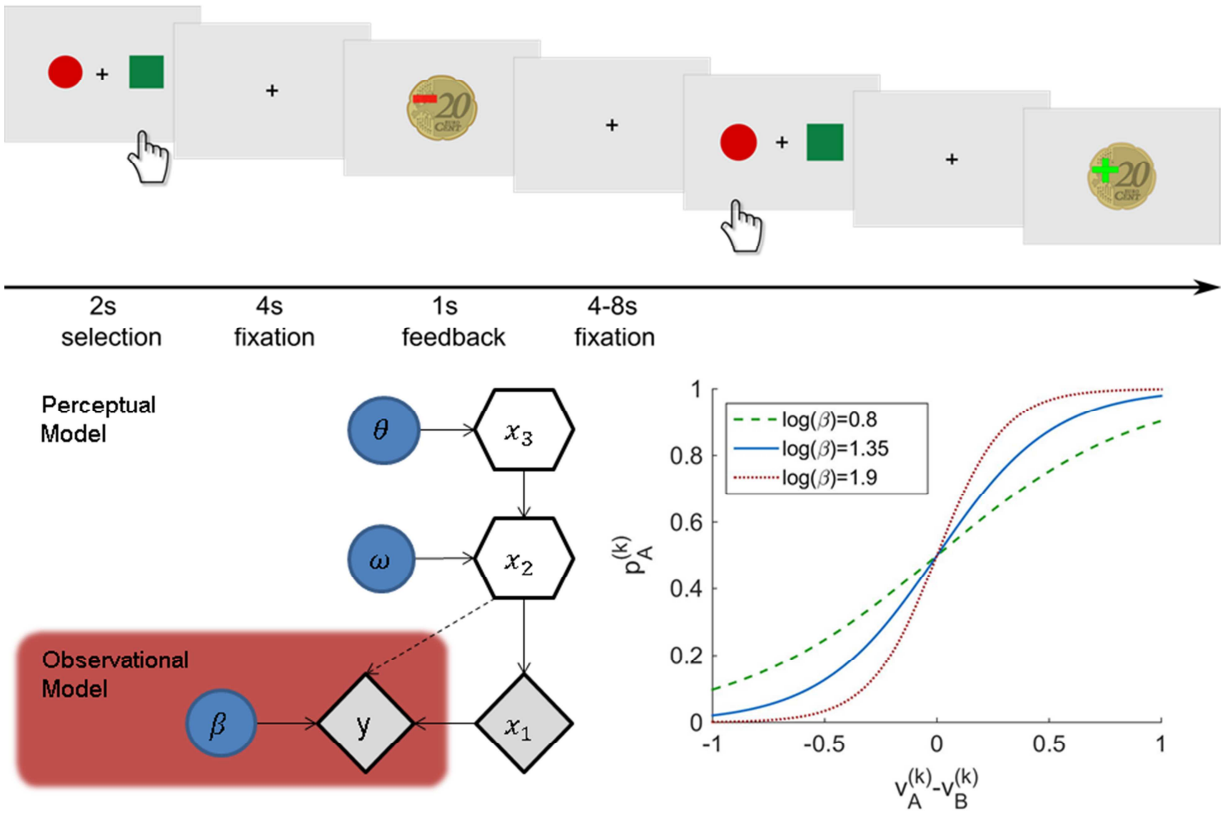
	AN		**HC**		**test statistics**	
	Mean	SD	Mean	SD	Group	
Behavioral measures					**F**	**p**
Correct answers	81.3	6.1	82.1	8.0	0.18	.675
Contingency reversal	9.2	1.4	8.7	1.9	1.27	.264
Perceptual model parameters					**F**	**p**
tonic log-volatility $[\omega]$	-1.15	.59	-1.62	1.54	2.86	.095
Log *meta-volatility* $[\log(\theta)]$	-5.87	1.38	-6.01	.64	.313	.578
Observational model parameter					**F**	**p**
Log *decision-noise* $[-log(\beta)]$	-1.33	.53	-1.39	.59	.197	.659
Quality of Fit					**F**	**p**
Log Model Evidence	-52.2	14.2	-52.9	15.5	.036	.850

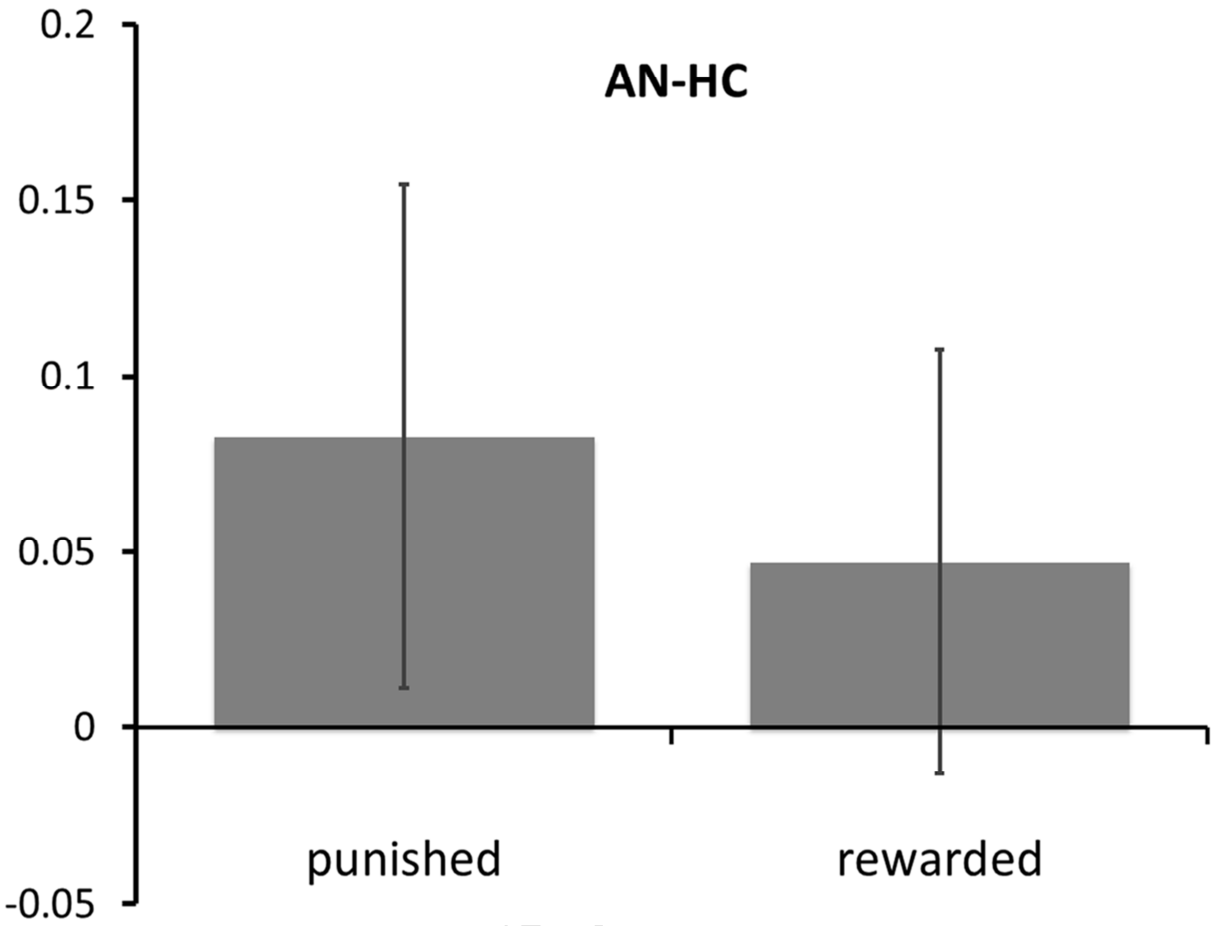
AN=anorexia nervosa patients; HC=healthy controls; *P*-values below 0.05 indicate a significant group difference. See Figure S1 for more details on performance parameters.

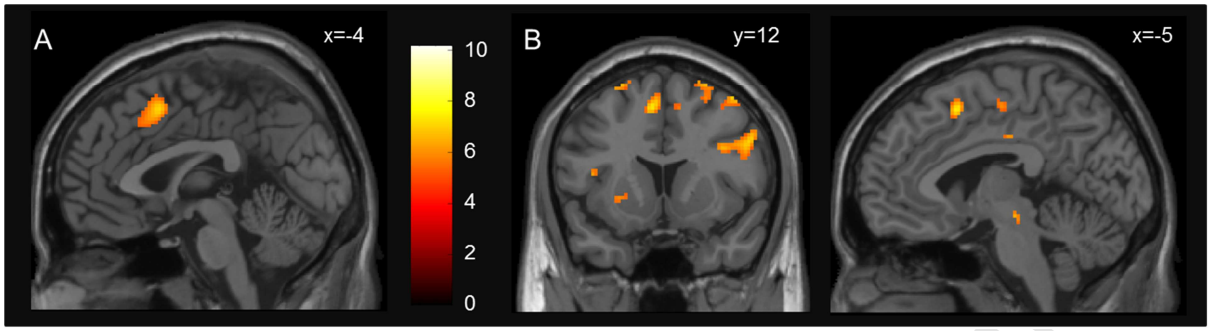
Table 3. Mixed factor ANOVA on trial dependent parameters. The individual trial dependent parameters from the HGF perceptual model and the reaction times were subjected to a 2×2×2 ANOVA after a logit and log transformation respectively (see Supplemental Methods) with group, response and feedback as factors. We provide F and p values for the main effects and interactions. Reaction times did not differ between the groups, but there was a main effect of response. The post hoc test revealed that reaction time was longer on those trials where a wrong answer was given.

Effect	**learning rate**			**prediction error**		
	df	F	p	df	F	p
response	1,8264	24.4	<.001	1,8275	823	<.001
feedback	1,8263	692.5	<.001	1,8260	13419	<.001
group	1,69.3	3.8	.055	1,83.7	.827	.366
response×feedback	1,8263	265.1	<.001	1,8260	21.4	<.001
feedback×group	1,8263	6.6	.010	1,8260	1.64	.200
response×group	1,8264	.02	.891	1,8275	.002	.964
response×feedback×group	1,8263	.46	.498	1,8260	1.925	.165

Effect	**expected value**			**reaction times**		
	df	F	p	df	F	p
response	1,8282	927	<.001	1,8274	9.99	.002
feedback	1,8272	10.7	.001	1,8270	1.06	.303
group	1,77.6	.926	.339	1,71.6	.425	.517
response×feedback	1,8273	.002	.962	1,8270	.052	.819
feedback×group	1,8272	.051	.822	1,8270	.139	.709
response×group	1,8282	.841	.359	1,8274	.577	.448
response×feedback×group	1,8273	1.35	.246	1,8270	.821	.365


2s
selection
4s
fixation
1s
feedback
4-8s
fixation
Perceptual
Model
θ
ω
x3
x2
x1
Observational
Model
β
y
log(β)=0.8
log(β)=1.35
log(β)=1.9
p(k)A
V(k)A-V(k)B





Altered Medial Frontal Feedback Learning Signals in Anorexia Nervosa

Supplemental Information

Supplemental Methods

Participants

AN participants were recruited from specialized eating disorder programs of a university child and adolescent psychiatry and psychosomatic medicine department. Since nutritional rehabilitation has to be carried out with caution at the beginning of the treatment due to possible complications such as refeeding syndrome (1, 2), our patients had at that stage of treatment a calorie intake of about 1000 to 1500 kcal per day.

Body mass index (BMI) cut-offs for AN patients were defined according to Kromeyer-Hauschild et al. (3) and Hebebrand et al. (4). AN participants had to have a BMI below the 10th age percentile (if younger than 15.5 years) or a BMI below 17.5 kg/m^2 (if older than 15.5 years) and no recent weight gain.

HCs were recruited through advertisement among middle school, high school, and university students. For case-control age-matching an implementation of the Munkres algorithm was used (5).

HC participants were excluded (before scanning) if they had any history of psychiatric illness, a lifetime BMI below the 10th age percentile (if younger than 18 years)/BMI below 18.5kg/m^2 (if older than 18 years), or were currently obese (BMI over 97th age percentile if younger than 18 years; BMI over 30kg/m² if older than 18 years). HC participants had to be eumenorrhoeic.

Participants of all study groups were excluded if they had a lifetime history of any of the following clinical diagnoses: organic brain syndrome, schizophrenia, substance dependence, psychosis NOS, bipolar disorder, bulimia nervosa, or binge-eating disorder (or "regular" binge eating - defined as bingeing at least once weekly for 3 or more consecutive months). Further exclusion criteria for all participants were IQ lower than 85; psychotropic medications within 4 weeks prior to the study, current inflammatory, neurologic or metabolic illness; chronic medical or neurological illness that could affect appetite, eating behavior, or body weight (e.g. diabetes); clinical relevant anemia; pregnancy; breast feeding**.**

The magnitude of our samples was based on a similar study (6) that found a moderately large effect size (d=.6) comparing means of model parameters between groups using a t-test. A power

analysis using Gpower (7) indicated that a sample of 36 people in each group would be needed to detect such effects with 80% statistical power at a confidence level of 1-α= .95.

Psychiatric and Psychological Assessments

Exclusion criteria and possible confounding variables, e.g. the use of psychotropic medications and medical comorbidities, were obtained using the expert version of the Structured interview for anorexia and bulimia nervosa for DSM-IV [SIAB-EX (8)] and our own semi-structured interview. SIAB-EX is a well-validated 87-item semi-standardized interview that assesses the prevalence and severity of specific eating-related psychopathology over the past three months. The interview provides diagnoses according to the ICD-10 and DSM-IV and is suitable for adolescents as well as adults. It has been used widely in eating disorder research (9–11). A good inter-rater reliability (k=.81) for the diagnostic interview has been demonstrated (12). Interviews were conducted by clinically experienced and trained research assistants under the supervision of the attending child and adolescent psychiatrist. The SIAB-EX interview was also used to determine the AN subtype, resulting in 32 patients being used in the final analysis to be of restrictive type (AN-r) and 3 of binge-purge type.

To complement the information obtained with the clinical interviews, eating disorder-specific psychopathology was assessed with the German version of the Eating Disorders Inventory [EDI-2 (13)]. Depressive symptoms were explored using the German version of the Beck Depression Inventory [BDI-II (14)]. For personality and character traits such as punishment sensitivity we used the German version of the BIS-BAS scales (15) and an adaptation of the original Temperament and Character inventory questionnaire for adolescents [JTCI (16)]. All other symptoms were gauged using the revised Symptom Checklist 90 [SCL-90-R (17)].

For AN patients, comorbid psychiatric diagnoses other than eating disorders were derived from medical records and confirmed by an expert clinician with over 10 years of experience after careful chart review (including consideration of medical and psychiatric history, physical examination, routine blood tests, urine analysis, and a range of psychiatric screening instruments). All diagnostic information was ascertained at the time of treatment and the principal investigator of this study is also the chief consultant of the eating disorder treatment center.

Psychiatric conditions in potential HC were ascertained using the same instruments as in patients. If there were any indications of psychiatric symptoms each case was discussed with a board-certified expert clinician and assessments were extended if necessary.

The BMI standard deviation score (BMI-SDS), that provides an index of weight to height ratio that is corrected for age and gender, was computed according to Hemmelmann et al. (18) and Kromeyer-Hauschild et al. (3).

Intelligence quotient (IQ) was measured with a short version of the German adaptation of the Wechsler Adult Intelligence Scale [WIE (19)] for participants aged ≥16 years or a short version of the German adaptation of the Wechsler Intelligence Scale for Children [HAWIK (20)] for participants aged <16 years. This version of the HAWIK (21) included the following subtests: vocabulary, letter number sequencing, matrix reasoning, and symbol search. The short version of the WIE (19, 22) included the subtests: picture completion, digit symbol coding, similarities and arithmetics.

Handedness was assessed using a short version of the Annett Scale of Hand Preference (23) as previously implemented in (24). This questionnaire asks for handedness in typical daily life situations as writing or brushing teeth. Response categories range from 0 'right hand', 1 'both hands' to 2 'left hand'. A mean score for handedness was calculated.

Study data were collected and managed using secure, web-based electronic data capture tools REDCap [Research Electronic Data Capture (25)].

Experimental Paradigm & Data Quality

In each trial of the task, a coloured circle and a coloured square were presented on the left and right side of a screen (spatial position randomized). Subjects were asked to choose one of the two symbols by pressing the left or right button as soon as they had decided. Trials were considered valid if the answer was given within 2 seconds after stimulus presentation. Implicitly, one symbol was designated as the 'correct' and the other one as 'wrong'. The choice of the 'correct' symbol led to a monetary reward (+20 cents) with a probability of 80% and to a punishment (-20 cents) in 20% of the cases (probabilistic errors). The choice of the 'wrong' symbol led to punishment and reward with inverted probability (-20 cents with 80% probability and +20 cents with 20% probability). Consequently, choosing the correct symbol led to accumulating monetary gain, whereas choosing the wrong symbol led to a cumulative monetary loss. With a probability of 25% the contingency reversed (change of the 'correct' symbol to the previously 'wrong' symbol) after at least four consecutive correct decisions since the last contingency switch, requiring a behavioural adaptation in the following trials. The total win was paid at the end of the session. The task performed in the scanner consisted of 120 trials (total duration of ca. 26 minutes).

Before entering the scanner, participants performed a training session of a simplified version of the task described above in which reward contingencies did not reverse. The practice session was completed after 8 consecutive choices of the 'correct' symbol or a maximum of 30 trials.

To prevent confusion, instructions were displayed again (after training) right before the scanner task started and participants were informed that the paradigm had been changed in a subtle way, and in particular that what was designated as the 'correct' option might change during the course of the experiment. Participants were instructed to maximize their gain and were informed that they will receive their total monetary reward immediately after completing the experiments, in addition to a fixed monetary compensation.

In the scanner, all stimuli were presented via a head-coil-mounted display system based on LCD technology (NordicNeuroLab AS, Bergen, Norway). Participants responded with two LUMItouch keypads (Lightwave Medical Industries, Vancouver, Canada). Stimuli were presented using Presentation® software (v11.1, Neurobehavioral Systems, Berkeley, CA ,USA).

To ensure data quality, we verified that no participant had a hit ratio (number of decisions for the 'right' figure divided by total number of trials) below or equal to 50% (performance below chance). One participant from the AN group (and her age-matched HC counterpart) was excluded for excessively poor performance (< 5 successful switches, Figure S1).

Computational Model

In this experiment participants had to make optimal decisions in an uncertain and changing environment, based on a series of inputs (feedback associated to option A and B). In the approach proposed by Daunizeau (26) the decision process at each trial k is modeled in two phases.

A so-called 'perceptual model', describes how the participant infers about the hidden "*state of affairs*" from the feedback associated to each option, specifically how she guesses the expected values $v_A^{(k)}$ and $v_B^{(k)}$ from choosing option A and B, or just the inferred probability of reward, given that the amount of money won or lost remains constant over trials. To describe this phase, three models were compared, specified below.

Subsequently, taking into account the inferred hidden states, the participant had to take action. This process is described by a so called 'observational-model', chosen to be a softmax function s, according to which option A at trial k is chosen with a probability:

$$p_A^{(k)} = \mathrm{s}\left(\beta(v_A^{(k)} - v_B^{(k)})\right), \qquad \mathrm{s}(x) = \frac{1}{1+\exp(-\mathrm{x})}$$

(and $p_B^{(k)} = 1 - p_A^{(k)}$) where β, the slope of the sigmoid curve at $v_A^{(k)} = v_B^{(k)}$, is the inverse of the subject-specific *decision noise*.

Hierarchical Gaussian Filter

While we refer to (27) for a detailed description of the HGF in a general form and its inversion (i.e. the derivation of the update equations describing the evolution of the trialwise expected values $v_A^{(k)}$, $v_B^{(k)}$, prediction errors $\delta^{(k)}$ and learning rate $\alpha^{(k)}$), we aim here to describe the specific HGF used in this study and to introduce and interpret the subject-specific parameters that couple the different levels of the hierarchy and determine the individual learning process.

The generative model

We used the HGF model introduced in the original work (27) and characterized by the presence of three levels and binary input (Figure 1). The hidden states (levels 2 and 3) are assumed to evolve as a Gaussian random walk, such that its variance depends on the state at the next higher level.

The environmental state at trial k is represented by the variable $\mathrm{x}_1^{(k)} \in \{0,1\}$, while the levels 2 and 3 are represented by $x_2^{(k)}, x_3^{(k)}$ respectively. Given our assumption of no perceptual uncertainty, knowledge of $\mathrm{x}_1^{(k)}$ would allow for an accurate prediction of the feedback received at trial k, $u^{(k)}$. Specifically it holds $u^{(k)} = x_1^{(k)}$. We make $u^{(k)} = 1$ and $u^{(k)} = 0$ correspond to reward and punishment, respectively. The probability distribution for x_1 is determined by the next higher level, $x_2 \in \{-\infty, +\infty\}$. Namely, we make $x_2 = 0$ correspond to the equiprobability of being rewarded or punished. When $x_2 \to \infty$ the probability of a reward approaches 1, conversely the probability of a punishment approaches 1 when $x_2 \to -\infty$. In the model used, this was achieved by means of the Bernoulli distribution:

$$p(x_1|x_2) = \mathrm{s}(x_2)^{x_1}\big(1 - \mathrm{s}(x_2)\big)^{1-x_1} = \text{Bernoulli}(x_1; \mathrm{s}(x_2)).$$

The value of x_2 at trial k is normally distributed around the value at trial $k-1$, with a variance depending on the highest level (for this model), the state x_3:

$$p\left(x_2^{(k)} \middle| x_2^{(k-1)}, x_3^{(k)}\right) = N\left(x_2^{(k)}; x_2^{(k-1)}, \exp(\kappa x_3^{(k)} + \omega)\right).$$

The log-volatility $\kappa x_3^{(k)} + \omega$ depends on two subject-specific parameters κ and ω, that allow for individualized Bayesian learning. Since the 'observational model' chosen does not depend on x_3, κ can be set to 1 with no loss of generality. ω is a *tonic log-volatility*, that, if low, typically induces a low learning rate. The *meta-volatility* (volatility of x_3) was set to a constant θ, which again was allowed to vary between agents.

At this point, given the priors on the initial state, which were set to be delta-functions:

$$p\left(x_2^{(0)}, x_3^{(0)}\right) = \delta\left(x_2^{(0)}\right)\delta\left(x_3^{(0)} - 1\right),$$

the generative model was defined for all times k by recursion to $k = 1$. Indeed, for any prior on the parameters $p(\kappa, \omega, \theta)$, we can compute the distribution probabilities for x_1, x_2 and x_3.

Model inversion and emerging reinforcement learning structure

Inverting the model means to compute the probability:

$$p\left(x^{(k)}, \chi \middle| u^{(1\ldots k)}\right)$$

where $x^{(k)} = \{x_1^{(k)}, x_2^{(k)}, x_3^{(k)}\}$, $\chi = \{\kappa, \omega, \theta\}$ and $u^{(1\ldots k)}$ indicates the inputs received from the beginning till trial k. It has been shown that (27), applying a chain of plausible approximations (free energy, mean-field, fixed simplified form for the marginals $q(x_i^{(k)})$, which are required to be fully characterized by mean and variance only, and a quadratic non-Laplacian approximation to the variational energies), at any level of the hierarchy i the update of the belief on trial k (i.e., posterior mean $\mu_i^{(k)}$ of the state $x_i^{(k)}$) is proportional to a precision-weighted prediction error:

$$prediction(k) = prediction(k-1) + learning\ rate(k) \times prediction\ error(k).$$

This equation has the same structure of update equations from Reinforcement Learning models, however with a time-varying learning rate such that the impact of the prediction errors is modulated by the environmental volatility and the certainty of beliefs, resulting in a bigger impact of prediction errors in more uncertain trials. In this work we focused on the implied learning rate $\alpha^{(k)}$ and prediction error $\delta^{(k)}$ at 1st level:

$$\delta^{(k)} = u^{(k)} - \hat{\mu}_1^{(k)}$$

$$\alpha^{(k)} = \frac{\hat{\mu}_1^{(k)} - \hat{\mu}_1^{(k-1)}}{\delta^{(k)}},$$

where hatted quantities represent predicted values, i.e. before associated feedback was shown. The predicted expectation values at 1st level $\hat{\mu}_1^{(k)}$ are used as input for the observational model. In particular, to predict the choice of the participant at trial k, the difference in the expected values $v_A^{(k)} - v_B^{(k)}$ is required (see Figure 1 in the manuscript). If option A is chosen, $v_A^{(k)} = \hat{\mu}_1^{(k)}$ and, since $v_B^{(k)} = 1 - v_A^{(k)}$:

$$v_A^{(k)} - v_B^{(k)} = 2\hat{\mu}_1^{(k)} - 1.$$

Similarly, $v_A^{(k)} - v_B^{(k)} = 1 - 2\hat{\mu}_1^{(k)}$ if option B is chosen at trial k.

To conclude, the free parameters of the models were set to have priors:

$$p(\omega) = N(\omega; -3,4) \qquad p(\log(\theta)) = N(\log(\theta); -6,4)$$

$$p(\log(\beta)) = N(\log(\beta); 0, \sqrt{5}).$$

For the implementation and inversion of the above model we used the HGF toolbox (v4.10; http://www.translationalneuromodeling.org/tapas/), selecting the Broyden, Fletcher, Goldfarb and Shanno (BFGS) quasi-Newton optimization algorithm with the prior means as initialization parameters.

Rescorla Wagner Models

Further 'perceptual models' considered in this study are a traditional RW model (28) with a constant learning rate and a reinforcement learning model with an adaptive learning rate (29), that grows in correspondence with switches in reward contingencies. Also these perceptual models were combined with the softmax 'observational model' described above.

The evolution of the learning rate $\alpha^{(k)}$ in the model with an adaptive learning rate is described by the following equations:

$$\delta_{abs}^{(k)} = \delta_{abs}^{(k-1)} * \left(1 - \alpha^{(k)}\right) + \left|\delta^{(k)}\right| * \alpha^{(1)}$$

$$m^{(k)} = \frac{2 * (\delta_{abs}^{(k)} - \delta_{abs}^{(k-1)})}{\delta_{abs}^{(k)} + \delta_{abs}^{(k-1)}}$$

$$f\left(m^{(k)}\right) = \text{sign}\left(m^{(k)}\right) * \left(1 - \exp\left(-\frac{m^{(k)}}{\gamma^2}\right)\right)$$

$$\alpha^{(k+1)} = \begin{cases} \alpha^{(k)} + f\left(m^{(k)}\right) * \left(1 - \alpha^{(k)}\right) & if\ m^{(k)} > 0 \\ \alpha^{(k)} + f\left(m^{(k)}\right) * \left(\alpha^{(k)}\right) & if\ m^{(k)} < 0 \end{cases}$$

where the parameter γ modulates the oscillations of the learning rate, playing a role similar to θ, ω in the HGF model. For further details, we refer to (29). We set the priors for the free parameters of the model as follows:

$$p\left(\text{logit}(\alpha^{(1)}, 1)\right) = N\left(\text{logit}(\alpha^{(1)}, 1); \text{logit}(0.5,1), \sqrt{10}\right)$$

$$p(\text{logit}(\gamma, 10)) = N(\gamma; \text{logit}(1,10), \sqrt{10})$$

$$p(\log(\beta)) = N\left(\log(\beta)\,; 0, \sqrt{5}\right) \qquad \text{logit}(x, a) = \log(\frac{x}{a - x})$$

where β is the inverse of the decision noise like above.

We choose again the BFGS as optimization algorithm, after having determined the initial parameters through a grid search for the maximum likelihood on:

$$0.1 < \alpha^{(1)} < 0.9; 0.3 < \gamma < 4;\ -1 < \log(\beta) < 2.$$

Model Comparison

The TAPAS software searches for the best fitting parameters by maximizing an approximation to the log model evidence (LME), namely the negative variational free energy under the Laplace assumption. The resulting LME values were subjected to a random-effects Bayesian model selection procedure [spm_BMS function included in SPM12 (30)] to determine Expected Posterior Probabilities (PP) and Protected Exceedance Probabilities (PXP) for each model (31). PXPs represent our belief that a particular model is more likely than any other model in the comparison set, without relying on the assumption that the frequencies of each model differ. After running BMS initially across all participants, this was then done separately for controls and patients, to account for the possibility that the groups differ with regard to which model fits their behavior best.

Behavioral Analysis

The normality distribution of all trial independent parameters was assessed through a Kolmogorov-Smirnov test. All the parameters were subjected to a standard ANOVA. There was one parameter (ω) that revealed significant deviations from normality (at p<.05) and was therefore also subjected to a non-parametric Mann-Whitney U-test.

After applying a logit transformation, the trial dependent parameters were subjected to a full-factorial 2×2×2 linear mixed model with subjects as random effects and an autocorrelation structure of order 1 (AR1), group as between subject factor and feedback and response as within subject factor. A logit transformation maps a parameter p belonging to an interval $p \in (a, b)$ to the real axis, therefore reducing violations to the normality of distribution due to the presence of the interval bounds. Specifically, the logit transformation is the inverse of the softmax function:

$$\tilde{p} = \text{logit}(p) = \log\left(\frac{p - a}{b - p}\right).$$

The restricted maximum-likelihood approach (REML) was used to estimate the model free parameters. The residual plots did not present obvious violations of normality or homoscedasticity.

When an interaction was found to be significant, post hoc t-tests were performed, using a Bonferroni correction to take into account the multiplicity of the comparisons.

Subsequently, given the findings of (6) that adults behave differently from adolescents in a task similar to ours, we supplemented our analysis by also considering extensions of the models above that included age as a covariate and all the interactions of age with the between and within subject factors.

MRI Data Acquisition

Structural and functional images were acquired between 8 and 9 am in the morning after an overnight fast using standard sequences with a 3 T whole-body MRI scanner (TRIO; Siemens, Erlangen, Germany) equipped with a standard head coil.

The T1-weighted structural brain scans were acquired with rapid acquisition gradient echo (MP-RAGE) sequence with the following parameters: number of slices=176; repetition time=1900ms; echo time=2.26ms; flip angle (FA) of 9°; slice thickness of 1 mm; voxel size of 1 x 1 x 1mm^3; field-of-view (FoV) of 256 x 224mm^2; bandwidth of 200 Hz/pixel.

The functional images were acquired by using a gradient-echo T2*-weighted echo planar imaging (EPI) with the following parameters: tilted 30° towards AC–PC line (to reduce signal dropout in orbitofrontal regions); number of volumes=656; number of slices=42; repetition time=2410ms; echo time=25ms; FA of 80°; 3mm in-plane resolution; slice thickness of 2 mm (1mm gap resulting in a voxel size of 3 x 3 x 2mm^3); FoV of 192 x 192 mm^2; bandwidth of 2112 Hz/pixel.

MRI Data Preprocessing & Quality Control

For fMRI preprocessing and analysis, we used SPM8 (http://www.fil.ion.ucl.ac.uk/spm/) unless otherwise specified. The functional images were corrected for temporal slice-timing and motion simultaneously using realign4D (32). The six realignment parameters, describing the rigid-body movement (x, y, z, pitch, roll, yaw), were saved and later used as nuisance covariates to account for the variance due to motion. The EPI volumes were coregistered to the subject's structural brain image. Then, to reduce noise due to physiological fluctuations and other sources, including motion not accounted for by realign4D, we extracted the six principal components of noise as computed from the CSF and white matter mask using the CompCor method (33) with anatomically defined ROI. This required the accurate specification of noise ROIs, in which the time series data are unlikely to be modulated by neural activity. To this end we segmented the structural images into partial volume maps of cerebral spinal fluid (CSF), white matter (WM), and gray matter (GM). Erosion (kernel of

1x1x1mm) was applied to the binarized CSF and WM maps, to minimize partial overlapping with gray matter. Finally the merging of the eroded CSF and WM masks defined the noise ROIs that were used to detect noise fluctuations. The six noise principal (34) components were then subjected to our first level GLM as nuisance regressors. This was followed by the normalization to MNI space using the individual structural images and by means of the DARTEL template (35). The resulting data were smoothed with an isotropic 8mm FWHM Gaussian kernel.

We evaluated the quality of the fMRI data by manual inspection and using artifact detection tools (ART, (36)). Volumes that exceeded a brightness intensity threshold of three standard deviations or a threshold of 1 mm normalized movement in any direction were classified as outliers [motion-outlier: acAN: median=2 HC: median=2; intensity-outlier: acAN: mean(std)=9.1(4.9) HC: 9.4 ±5.0]. The two groups did not differ regarding numbers of motion- and intensity-outliers [motion-outlier, Mann-Whitney U-test: 521; p=0.612; intensity-outlier, t-test: t(68)=-0.286; p=0.78].

MRI Data Analysis

Parametric modulators in first level analysis

In our first level analysis we implemented one GLM for each trial-dependent parameter of interest, namely the learning rate $\alpha^{(k)}$, the prediction error $|\delta^{(k)}|$ separately on punished and rewarded trials and the expected value of the chosen option $v^{(k)}$. The trialwise prediction error on rewarded and punished trials was always positive and negative, respectively. Each parametric modulator was associated with an event lasting for 1 second: starting at feedback for $\alpha^{(k)}$ and $\delta^{(k)}$ and 1 second before button press for $v^{(k)}$ and mean subtracted to achieve zero mean. Specifically, starting from the mean subtracted parameter x, we built the parametric modulator to be used in the imaging analysis p as:

$$p = \frac{x}{\max(|x|)}$$

such that $p \in [-1,1]$.

ROI analysis

As part of a more exploratory approach, we also used small volume corrections as implemented in SPM in three a priori defined ROIs (see above and Figure S1) corresponding to the ventromedial prefrontal cortex (vmPFC, $v_{A,B}^{(k)}$), ventral striatum (VS, $\delta^{(k)}$), and posterior Medial Frontal Cortex (pMFC, $\alpha^{(k)}$). In this respect, the mask of the VS was specified by binarizing a probabilistic map (37) with a threshold value of 0.4. The vmPFC mask was created by merging the left

and right frontal medial orbital cortex and rectus label from the Automated Anatomical Labelling (AAL) atlas provided in the Wake Forest University (WFU) PickAtlas for SPM (38, 39). To create the mask of the pMFC, first the Neurosynth forward inference map for the term 'learning' [www.neurosynth.org; (40)] and corresponding to a threshold of estimated FDR<.01 is eroded and subsequently dilated with a sphere kernel (radius 5mm) using fslmaths (41), to obtain a smooth volume. Subsequently, this image is intersected with the union of the left and right hemisphere regions of the AAL Atlas corresponding to the labels: anterior- and mid-cingulum cortex, supplementary motor area and frontal superior medial cortex. The resulting volume located along the medial wall corresponds well with the meta-analytic data on cognitive control from Ridderinkhof et al. (2004) and Shackman et al. (2011).

Supplemental Results

Sample Characteristics

The sample used in this study showed no significant differences with respect to the control group for the BIS and the harm avoidance scales respectively in the BIS-BAS and the JTCI questionnaires, in contrast to previous studies, documenting increased harm avoidance and BIS in AN (42, 43). To investigate whether this discrepance was due to insufficient statistical power, we considered a larger sample of AN patients and healthy controls satisfying the same inclusion and exclusion criteria described in section SM 1.1 and HA, and for whom the results of the JTCI [n(AN) = 172, n(HC) = 207) and BIS-BAS (n(AN) = 136, n(HC) = 178] questionnaires but no imaging data were available. This analysis is summarized in Table S2 and indicate that indeed both the harm avoidance and the BIS scales were increased in AN.

Model Comparison

Bayesian Model Selection across controls and AN patients, as well as for both groups separately revealed that the HGF fitted behavior best in both groups (see Table S1). Therefore the resulting trialwise learning rate $\alpha^{(k)}$, prediction error $\delta^{(k)}$ and expected value $v^{(k)}$ were used as parametric modulators in the fMRI analysis.

Behavioral Analyses

A Kolmogorov-Smirnov test revealed significant deviations from normality for the *tonic log-volatility* (ω) on the HC sample (p(HC)=0.0105, p(AN)=0.0529), but no significant deviation for the other trial independent parameters. The results of the ANOVAs and of the Mann-Whitney test were reported in Table 2 of the manuscript and revealed no significant group difference for any of the parameters. We summarize in Table S2 the results of the ANOVA including age as a covariate. Even in this case no group differences are found. However, similarly to Javadi et al. (6), the rate of correct answers, reversals, the LME and $\log(\beta)$ increased with age.

The results of the all factorial mixed models for the time dependent HGF parameters are reported in Table 3 of the main article. For the learning rate main effects of response and feedback were significant [response F(1,8264.1) = 24.4, p<.001; feedback F(1,8262.6) = 692.5, p<.001], and post hoc t-tests indicated that the learning rate was higher on trials when a wrong response was given [wrong mean(SE)=0.353(0.017), right=0.330(0.017)], and on punished trials [punished=0.408(0.018), rewarded=0.281(0.015)]. Interaction of response and feedback was found to

be significant [F(1,8262.7) = 265.1, p<.001] and post-hoc t-tests revealed that the learning rate was the highest in punished trials after a correct response (0.438(0.019)), followed by punished after wrong (0.379(0.018)) rewarded after wrong (0.329(0.018)) and rewarded after right (0.237(0.014)). Interaction of group and feedback was also significant [F(1,8262.6) = 6.6, p = .010] and in this case post hoc t-tests revealed (see also Figure 2 in main article) that AN had a higher learning rate in punished trials [AN=0.450(0.026), HC=0.367(0.025)].

For the prediction error main effects of feedback and response were significant [response F(1,8275.2) = 823, p<.001; feedback F(1,8259.9) = 13419, p<.001], showing on one hand higher and positive prediction error on rewarded trials (0.775(0.023)) and negative prediction error on punished trials (-1.507(0.011)), and on the other hand higher prediction error on wrong trials [wrong=-0.225(0.028), right=-0.859(0.020)]. Interaction of response and feedback was significant [F(1,8260.4) = 21.4, p<.001] and post-hoc t-tests revealed that the prediction error was the lowest on punished trials after a correct response (-1.656(0.010)), followed by punished after wrong (-1.307(0.016)) rewarded after right (0.513(0.022)) and rewarded after wrong (1.010(0.029)).

For the expected value we found significant main effects of response [F(1,8282.4) = 927, p<.001; right= 0.622(0.015), wrong=0.177(0.026)) and feedback (F(1,8272.4) = 10.7, p=.001; rewarded= 0.400(0.022), punished=0.449(0.020)].

There were no group differences for the reaction times, but the main effect of response was significant [F(1,8272.2) = 9.99, p=.002]. Post hoc tests revealed shorter reaction times on correct answers [right=0.604(0.013) seconds, wrong = 0.620(0.014) seconds].

The results of the mixed model with age added as a nuisance regressor are summarized in Table S3, showing that all main effects and interactions that were significant in the main model, remained significant when age was added as a covariate. Post hoc t-tests revealed that the directions of the effects also remain unaltered. For the learning rate the 3-way interaction response×feedback×age was significant [F(1,8256.4) = 12.8, p<.001]. For the prediction error the interactions response×age and feedback×age were significant [response×age: F(1,8259.0) = 29.5, p<.001; feedback×age: F(1,8251.4) = 14.7, p<.001]. For the expected value the interaction response×age was significant [F(1,8269.7) = 33.3, p<.001].

To clarify the impact of the AN subtype on our main results, of increased learning rate in AN on punished trials, we repeated our analysis only on the restrictive AN subsample, because the size of the binge-purging subtype (N=3) is too small for statistical treatment. By reproducing the results of our original analysis (for details please see Table S5 and Figure S5 below) we argue that our results are driven by participants of the restrictive subtype and it is unclear whether they also apply to patients of the binge/purge subtype.

To explore the effect of symptoms on the learning rate, particularly on punished trials, we considered mixed models restricted to the AN group where in addition to the factors response and reward and their interaction, in turn BMI-SDS, BDI-II and EDI-2 are added as covariates together with their interaction with reward. Results for these analyses are reported in Tables S6 below.

All analyses detect the expected effects of response, reward and their interaction. Moreover, a correlation at the trend level between EDI-2 and the learning rate on punished trials was found (namely, as expected, worse symptoms are associated with higher learning rate on punished trials). This same effect was observed (and tested as nominally significant) when EDI-2 is substituted by BDI-II. Malnutrition, as assessed through BMI-SDS, seems also to be associated with learning rate on punished trials. While our findings about BDI-II and EDI-2 would not withstand a Bonferroni correction for multiple comparison, they support our main finding that learning rate on punished trials is related to typical AN symptoms. However, the detected anticorrelation between BMI-SDS and the learning rate on punished trials may suggest that learning rate on punished trials is not increased in AN as a consequence of undernutrition or driven by patients with a particularly low BMI, and motivates further analysis in recovered AN patients, for whom undernutrition is not a confounding factor. No correlation between mean punishment learning rate and HA both in the whole sample and separately in the respective participant groups was detected (Pearson's r(AN+HC) = -.053, p=.668; r(AN)=-.186, p=.292; r(HC)= -.003, p=.988).

Imaging Analyses

The three GLMs with trialwise parameters from the HGF as parametric modulators revealed expected activation patterns in line with previous studies (44–46), see Chase et al. (47) for a review. We report in Table S4 a full list of activations found in this study. Additionally, the results for the learning rate were illustrated in Figure 3 (but see also Figure S4), while the results for the prediction error, separating rewarded and punished trials, and the expected value are shown in Figure S3. No group differences in the activation patterns were significant. Furthermore to test whether the subtype of AN patients had an impact on our finding of increased pMFC activation, in a region related to the learning rate, we repeated this analysis by restricting the AN sample to patients of the restrictive subtype. Results are reported in Figure S5.

While our finding of increased pMFC BOLD activation in AN after punishment seems to be related to the finding of an increased learning rate after punishment and to the fact that activity in that area correlates with the learning rate, it does not establish any causality relation. To test whether increased activation in the pMFC as observed on punished trials in AN patients leads to an increased learning rate or vice versa, we conducted mediation analysis using the PROCESS toolbox

(48). We used the MarsBaR tool (http://marsbar.sourceforge.net/ 48) to extract the mean activation on punished trials in two spheres (radius = 10mm) centered at the two peak coordinates in the pMFC, one located in the left and the other in the right hemisphere (see Table S6) where the difference in activation between the AN and HC group is at its maximum. The mean activation was correlated with the median learning rate on punished trials computed for each participant. Next, we employed the SPSS PROCESS toolbox with 5000 bootstrapping samples to compute 95% confidence intervals for inference about indirect effects. Since no indirect effect of group on the pMFC activation as mediated by the learning rate on punished trials or on the learning rate as mediated by the pMFC activation on punished trials was found to be significant (see Table S8 below), no causal relationship can be inferred.

Supplemental Tables

Table S1. Model comparison across groups and for each group separately. Results of Bayesian Model Selection of both patients and healthy controls together as well as for both groups separately, for all the models preliminarily considered in this work: a traditional RW, a RW model with adaptive learning rate and a HGF with three levels and binary input as described above.

	AN		**HC**		**AN&HC**	
	PP	**PXP**	**PP**	**PXP**	**PP**	**PXP**
RW (constant learning rate)	0.082	<.001	0.066	.001	0.059	<.001
RW (adaptive learning rate)	0.098	<.001	0.243	.002	0.152	<.001
HGF	0.820	>.999	0.691	.997	0.799	>.999

PXP=Protected Exceedance Probability, PP=posterior probability.

Table S2. Harm avoidance (JTCI) and inhibition (BIS-BAS) on a larger sample. Independent sample t-test results as computed on a larger sample than the one included in this study.

	AN			**HC**			**Test statistics**		
	N	**Mean**	**SD**	**N**	**Mean**	**SD**	**Dof**	**T**	**p**
BIS	136	22.4	3.5	178	19.6	3.3	312	7.265	<.001
JTCI harm avoidance	172	38.2	9.5	207	33.4	8.8	377	5.075	<.001

Table S3. ANOVA on trial independent parameters. The individual parameters from the HGF perceptual model and softmax observational model were subjected to a one way ANOVA with age as a covariate. Group means and standard deviations (SD) are given.

	AN		**HC**		**test statistics**	
	Mean	**SD**	**Mean**	**SD**	**group**	
Behavioral measures					**F**	**p**
Correct answers	81.7	5.9	82.7	7.9	0.39	.536
Contingency reversal	9.2	1.4	8.8	1.9	1.33	.253
Perceptual model parameters					**F**	**p**
tonic log-volatility $[\omega]$	-1.15	.59	-1.62	1.54	2.86	.095
Log *meta-volatility* $[\log(\theta)]$	-5.87	1.38	-6.01	.64	.257	.614
Observational model parameter					**F**	**p**
Log *decision-noise* $[-\log(\beta)]$	-1.33	.53	-1.39	.59	.10	.75
Quality of Fit					**F**	**p**
Log Model Evidence	-52.2	14.2	-52.9	15.6	.148	.702

AN=anorexia nervosa patients; HC=healthy controls; *P*-values below 0.05 indicates a significant group difference.

Table S4. Mixed factor ANOVA on trial dependent parameters with age as a covariate. The individual trial dependent parameters from the HGF perceptual model and the reaction times were subjected to a 2×2×2 full factorial mixed model with group, response and feedback as factors and age at date of scan as a covariate, after being respectively logit and log transformed (see above). We provide F and p values for the main effects and interactions of group, response and feedback only.

Effect	learning rate			prediction error		
	df	F	p	df	F	p
response	1,8258	23.2	<.001	1,8267	968	<.001
feedback	1,8257	702	<.001	1,8253	15832	<.001
group	1,67.2	3.63	.023	1,80.8	1.10	.298
response×feedback	1,8257	264	<.001	1,8254	26.4	<.001
feedback×group	1,8257	6.48	.011	1,8253	1.63	.202
response×group	1,8258	.016	.900	1,8267	.013	.909
response×feedback×group	1,8257	.245	.621	1,8254	1.41	.235

Effect	expected value			reaction times		
	df	F	p	df	F	p
response	1,8275	915	<.001	1,8267	10.4	.001
feedback	1,8266	10.2	.001	1,8264	1.15	.283
group	1,75.1	1.01	.319	1,69.5	.489	.487
response×feedback	1,8266	.044	.833	1,8264	.052	.820
feedback×group	1,8266	.072	.788	1,8264	.130	.718
response×group	1,8275	.415	.519	1,8267	.675	.411
response×feedback×group	1,8266	1.22	.270	1,8264	.713	.398

Table S5. Mixed factor ANOVA on trial dependent learning rate for the AN restrictive subgroup (AN-r). The individual trial dependent learning rate from the HGF perceptual model was subjected to a 2×2×2 ANOVA after a logit transformation (see above) with group, response and feedback as factors. In this supplementary analysis, AN patients not belonging to the AN-r subgroup (n=3) were excluded. We provide F and p values for the main effects and interactions. A post hoc t-test revealed that on punished trials the learning rate is greater in AN-r than in HC (mean difference(SE) = 0.088(0.036)).

Effect	**learning rate**		
	df	F	p
response	1,7910	25.4	<.001
feedback	1,7909	677	<.001
group	1,66.3	4.8	.033
response×feedback	1,7909	266	<.001
feedback×group	1,7909	4.3	.037
response×group	1,7910	.04	.843
response×feedback×group	1,7909	.36	.550

Table S6. Relationships between the trial dependent learning rate in the AN group and symptom variables. After a logit transformation (see SM above), the individual trial dependent learning rate from the HGF perceptual model was subjected to four 2×2 ANOVA with response and feedback as factors and in turn EDI-2, BDI-II and BMI-SDS as covariates, together with their interaction with the feedback factor. We provide F and p values for the main effects and interactions.

Effect	**learning rate**		
	df	F	p
response	1,3905	10.0	.002
feedback	1,3904	17.3	<.001
response×feedback	1,3904	117.3	<.001
EDI-2	1,31.11	0.10	.753
EDI-2×feedback	1,3904	3.86	.050

Effect	**learning rate**		
	df	F	p
response	1,4143	11.5	.001
feedback	1,4141	110	<.001
response×feedback	1,4141	131	<.001
BDI-II	1,33.11	0.08	.774
BDI-II×feedback	1,4141	4.07	.044

Effect	**learning rate**		
	df	F	p
response	1,4144	11.0	.001
feedback	1,4142	286	<.001
response×feedback	1,4141	134	<.001
BMI-SDS	1,33.2	14.9	<.001
BMI-SDS×feedback	1,4141	63.8	<.001

Table S7. Correlation of BOLD activity and HGF trial dependent parameters**.** Peaks voxels of clusters (k>100) where correlation between bold signal and parametric modulator is significant at FWE level. Peaks in the cerebellum were excluded. For each cluster only the global maximum is reported.

Brain region	**Peak Voxel MNI Coordinates**			**T**	**Cluster Extension**
	x	y	z		k
		α			
Globus Pallidus	10	0	0	7.81	231
Thalamus	8	-16	4	5.89	
Insula	34	18	6	7.18	231
pMFC	-4	14	50	7.04	474
Precuneus	8	-64	46	6.73	123
Insula	-30	16	-6	6.29	132
		$\lvert\delta\rvert$, rewarded trials			
superior frontal	0	34	46	9.47	1085
Insula	32	20	-6	8.25	199
middle frontal	46	14	44	7.16	414
Thalamus	10	-6	-2	6.96	254
Thalamus	-8	-6	0	6.09	
Insula	-30	18	-4	6.78	149
Parietal Inf	44	-50	46	6.74	118
		$\lvert\delta\rvert$, punished trials			
Precuneus	-10	-38	2	6.64	110
	-2	-44	40	6.55	437
		v			
Parahippocampal	-30	-36	-14	9.45	282
Middle Temporal	-62	-6	-18	8.42	729
vmPFC	-2	46	-10	8.00	1373
Precuneus	-10	-52	10	7.90	1545
Precuneus	8	-52	10	7.28	
...	-46	-68	28	7.07	413
	28	-26	-22	6.41	146

pMFC = posterior medial frontal cortex, vmPFC = ventromedial prefrontal cortex, VS = ventral striatum.

Table S8. Increased BOLD activity in AN after feedback on punished trials. Peaks voxels of clusters (k>50) where BOLD activation was stronger in AN than in HC after punishing stimuli was presented (at FWE level). For each of these peaks we also report the T-values for each group's activation, as computed within the 2×2 model. Clusters in the Cerebellum were excluded.

Brain region	**Peak Voxel MNI Coordinates**			**T**			**Cluster Extension**
				HC	AN	AN>HC	AN>HC
	x	y	z				
		α					
pMFC	-6	14	52	4.94	8.04	7.20	71
pMFC	6	2	64	2.30	6.78	7.39	270
Precentral	-42	-4	60	1.05	4.63	10.09	340
Occipital Mid	-32	-89	-4	5.39	8.37	12.15	1704
Occipital Inf	36	-84	-2	4.16	7.03	12.72	1313
Insula	-26	14	-2	2.40	5.82	6.91	111
Temporal Mid	40	-58	10	0.04	3.31	8.90	112
Cuneus	14	-96	14	-0.27	2.70	9.45	102
Precentral	-44	4	24	3.01	5.82	9.41	386
Frontal Mid	34	4	50	2.18	5.34	9.15	1686
Occipital	-22	-72	30	4.07	6.70	9.40	260
Angular	26	-58	44	4.21	7.25	11.27	512
Parietal Sup	-24	-58	46	4.90	7.62	8.50	171

Table S9a. Mediation analysis between brain activation in the pMFC on punished trials and diagnostic group with learning rate as a mediator. LLCI and ULCI indicate the lower and upper level of 95% confidence intervals respectively. Both the direct effect and indirect effect refer to confidence intervals for the group effect on pMFC activation with the learning rate on punished trials included as a mediator in the model. No indirect effect was significant.

	pMFC (left hemisphere)				pMFC (right hemisphere)			
	Direct Effect		Indirect Effect		Direct Effect		Indirect Effect	
	LLCI	ULCI	LLCI	ULCI	LLCI	ULCI	LLCI	ULCI
Learning Rate	-.6171	-.0069	-.1424	.0350	-.9391	-.3195	-.0976	.0960

Table S9b. Mediation analysis between learning rate on punished trials and diagnostic group with activation in the pMFC as a mediator. LLCI and ULCI indicate the lower and upper level of 95% confidence intervals respectively. Both the direct effect and indirect effect refer to confidence intervals for the group effect on learning rate on punished trials with the pMFC activation on punished trials in the right and left hemisphere spheres included as a mediator in the model. No indirect effect was significant.

	pMFC (left hemisphere)				pMFC (right hemisphere)			
	Direct Effect		Indirect Effect		Direct Effect		Indirect Effect	
	LLCI	ULCI	LLCI	ULCI	LLCI	ULCI	LLCI	ULCI
Learning Rate	-.0533	.0057	-.0139	.0051	-.0573	.0067	-.0192	.0174

Supplemental Figures

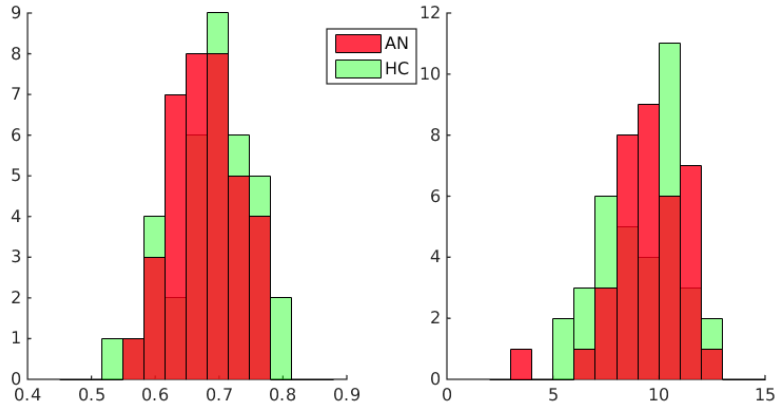


Figure S1. Performance indicators. Histogram for rate of correct answers (left) and number of achieved contingency switches (right) for our entire sample, before behavioral exclusion criteria were applied. The number of contingency switches was normalized to keep into account the (slightly varying) number of valid trials for each participant. The AN patient with sensibly lower achieved number of contingency switches was excluded from our final sample, together with her age-matched HC partner.

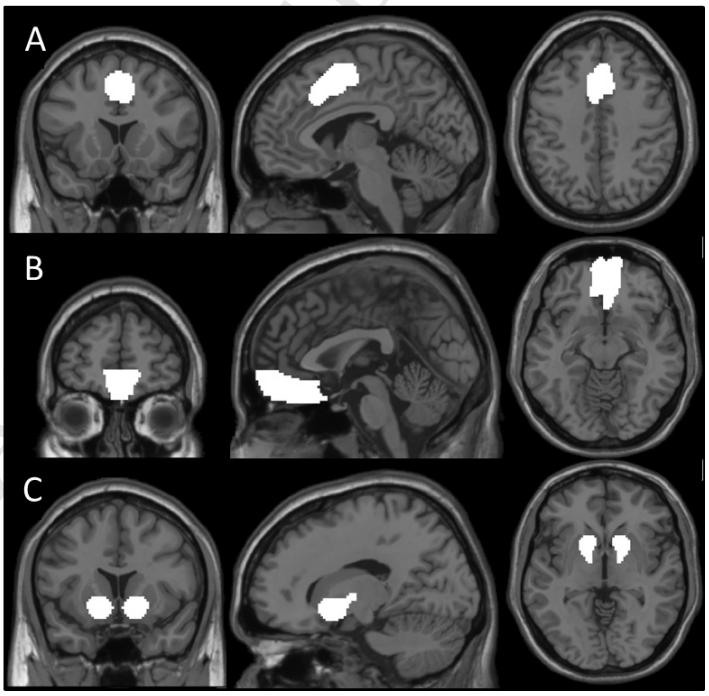


Figure S2. Regions of interest. (A) posterior medial frontal cortex – pMFC, (B) ventromedial prefrontal cortex – vmPFC, (C) ventral striatum – VS.

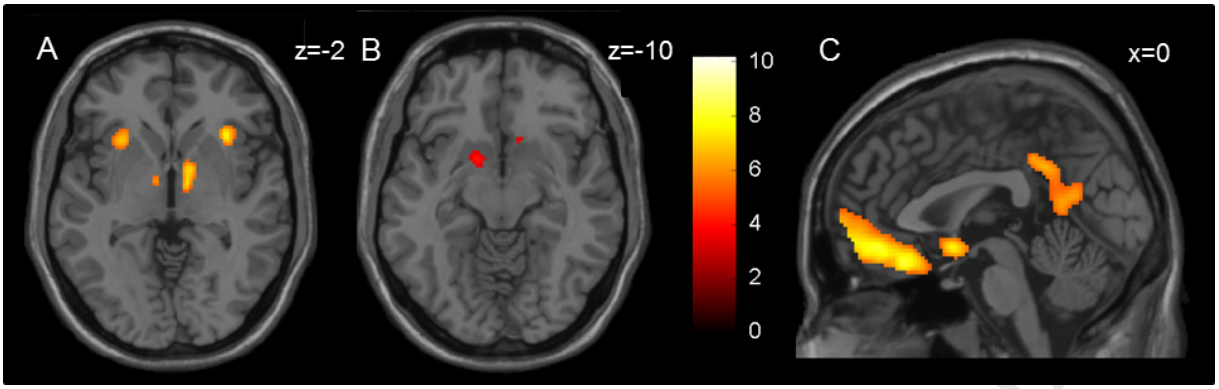


Figure S3. Correlation of BOLD activity with the absolute value of the prediction error $|\delta^{(k)}|$ on punished and rewarded trials (A and B, respectively) and expected value $v^{(k)}$ (C) as computed within a Hierarchical Gaussian Filter model. The color scale shows T-values. This correlation was significant at whole-brain FWE level for the expected value and the prediction error on rewarded trials, but only at a more lenient threshold (FWE-SVC level, namely all voxels were individually significant at the small volume corrected threshold) for the prediction error on punished trials. Therefore in (A) and (C) we display voxels where the signal was significant at FWE level (whole-brain), while in (B) we display voxels in the VS (see SM 1.9 for a definition of the VS mask) where the signal was significant at FWE-SVC level.

Figure S4. Overlap between the cluster in the pMFC showing a significant correlation with the learning rate in the whole-brain one-sample analysis (blue scale, FWE<.05) and the cluster where the AN group presented increased activation after punishment (red scale, whole-brain FWE<.05). A similar overlap of learning rate related activation and increased activation in AN also occurs in the left insula (see Tables S4, S5).

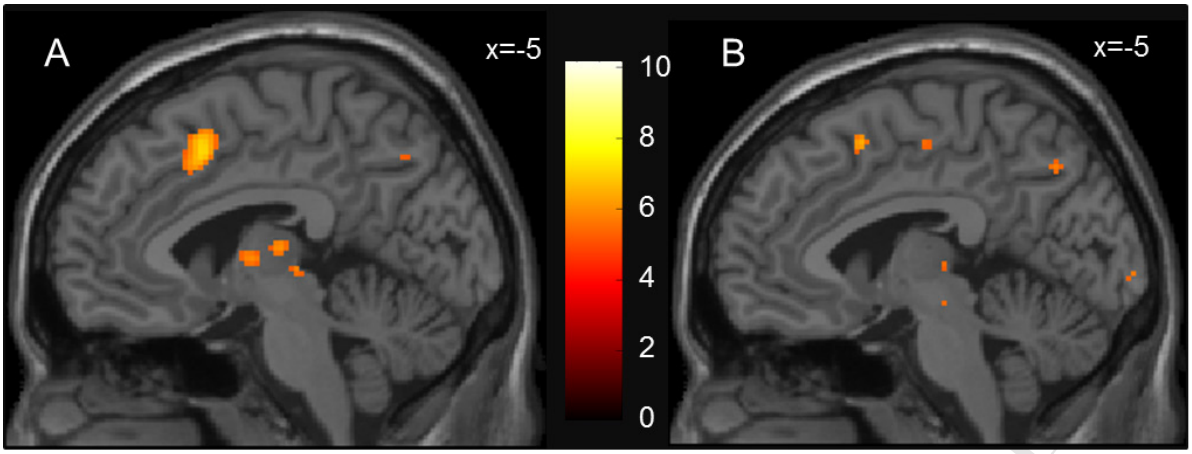


Figure S5. A: Correlation of BOLD activity after feedback with the learning rate α. The learning rate was computed within a Hierarchical Gaussian Filter and the expected pattern of activation in the pMFC across all healthy and AN-r participants (whole-brain one-sample t-test) was reproduced. **B: Increased BOLD activity in AN-r following punishment.** Increased BOLD activity in AN-r patients relative to HC following punishment is depicted on the same slice. We display regions where the signal is significant at a FWE<.05 level determined with random field theory (whole-brain). The color scale shows one sample t-test values.

Supplemental References

1. Ehrlich S, Querfeld U, Pfeiffer E (2006): Refeeding oedema : an important complication in the treatment of anorexia nervosa. *Eur Child Adolesc Psychiatry*. 15: 241–243.

2. Tey HL, Lim SC, Snodgrass AM (2005): Refeeding oedema in anorexia nervosa. *Singapore Med J*. 46: 308–310.

3. Kromeyer-Hauschild K, Wabitsch M, Kunze D, Geller F, Geiß HC, Hesse V, *et al.* (2001): Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. *Monatsschr Kinderheilkd*. 149: 807–818.

4. Hebebrand J, Casper R, Treasure J, Schweiger U (2004): The need to revise the diagnostic criteria for anorexia nervosa. *J Neural Transm (Vienna)*. 111: 827–840.

5. Munkres J (1957): *Algorithms for the Assignment and Transportation Problems*. .

6. Javadi AH, Schmidt DHK, Smolka MN (2014): Adolescents Adapt More Slowly than Adults to Varying Reward Contingencies. *J Cogn Neurosci*. 26: 2670–2681.

7. Faul F, Erdfelder E, Lang A-G, Buchner A (2007): G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 39: 175–191.

8. Fichter MM, Quadflieg N (1999): *Strukturiertes Inventar für anorektische und bulimische Eßstörungen: (SIAB) ; Fragebogen (SIAB-S) und Interview (SIAB-EX) nach DSM-IV und ICD-10 ; Handanweisung*. Hogrefe, Verlag für Psychologie.

9. Giel KE, Friederich H-C, Teufel M, Hautzinger M, Enck P, Zipfel S (2011): Attentional Processing of Food Pictures in Individuals with Anorexia Nervosa—An Eye-Tracking Study. *Biological Psychiatry*, Reduced Behavioral Flexibility in Addiction. 69: 661–667.

10. Kaye WH, Bulik CM, Thornton L, Barbarich N, Masters K (2004): Comorbidity of Anxiety Disorders With Anorexia and Bulimia Nervosa. *AJP*. 161: 2215–2221.

11. Zipfel S, Wild B, Groß G, Friederich H-C, Teufel M, Schellberg D, *et al.* (2014): Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study): randomised controlled trial. *Lancet*. 383: 127–137.

12. Fichter M, Quadflieg N (2001): The structured interview for anorexic and bulimic disorders for DSM-IV and ICD-10 (SIAB-EX): reliability and validity. *Eur Psychiatry*. 16: 38–48.

13. Paul T, Thiel A (2005): *Eating Disorder Inventory-2 (EDI-2): deutsche Version*. Hogrefe.

14. Hautzinger M, Keller F, Beck AT, Kühner C (2009): *Beck Depressions-Inventar : BDI II ; Manual*. Pearson Assessment.

15. Strobel A, Beauducel A, Debener S, Brocke B (2001): Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White. *Zeitschrift für Differentielle und Diagnostische Psychologie*. 22: 216–227.

16. Goth K, Schmeck K (2005): *JTCI - Junior Temperament und Charakter Inventar (JTCI-12-18 R)*. Hogrefe: Göttingen. Retrieved November 10, 2016, from http://www.equals.ch/software/testverfahren/jtci-junior-temperament-und-charakter-inventar-jtci-12-18-r.

17. Franke GH (2002): *Symptom-Checkliste von Derogatis (SCL-90-R)*. Beltz Test.

18. Hemmelmann C, Brose S, Vens M, Hebebrand J, Ziegler A (2010): Perzentilen des Body-Mass-Index auch für 18- bis 80-Jährige? Daten der Nationalen Verzehrsstudie II. *DMW - Deutsche Medizinische Wochenschrift*. 135: 848–852.

19. von Aster M, Neubauer AC, Horn R (2006): *WIE. Wechsler Intelligenztest für Erwachsene. Deutschsprachige Bearbeitung und Adaptation des WAIS–III von David Wechsler*. Frankfurt/Main: Harcourt Test Services.

20. Petermann F, Petermann U (2008): *Hamburg Wechsler Intelligenztest fuer Kinder IV (HAWIK-IV)*. Bern: Huber.

21. Waldmann H-C (2008): Kurzformen des HAWIK-IV: Statistische Bewertung in verschiedenen Anwendungsszenarien. *Diagnostica*. 54: 202–210.

22. Donnell AJ, Pliskin N, Holdnack J, Axelrod B, Randolph C (2007): Rapidly-administered short forms of the Wechsler Adult Intelligence Scale-3rd edition. *Arch Clin Neuropsychol*. 22: 917–924.

23. Annett M (1970): A Classification of Hand Preference by Association Analysis. *British Journal of Psychology*. 61: 303–321.

24. Gollub RL, Shoemaker JM, King MD, White T, Ehrlich S, Sponheim SR, *et al.* (2013): The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia. *Neuroinformatics*. 11: 367–388.

25. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG (2009): Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 42(2): 377–81.

26. Daunizeau J, Ouden HEM den, Pessiglione M, Kiebel SJ, Stephan KE, Friston KJ (2010): Observing the Observer (I): Meta-Bayesian Models of Learning and Decision-Making. *PLOS ONE*. 5: e15554.

27. Mathys C, Daunizeau J, Friston KJ, Stephan KE (2011): A bayesian foundation for individual learning under uncertainty. *Front Hum Neurosci*. 5: 39.

28. Rescorla R, Wagner A (1972): A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black A, Prokasy W, editors. *Classical Conditioning II: Current Research and Theory*. Appleton-Century-Crofts, pp 64–99.

29. Krugel LK, Biele G, Mohr PNC, Li S-C, Heekeren HR (2009): Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *PNAS*. 106: 17951–17956.

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

31. Rigoux L, Stephan KE, Friston KJ, Daunizeau J (2014): Bayesian model selection for group studies — Revisited. *NeuroImage*. 84: 971–985.

32. Roche A (2011): A four-dimensional registration algorithm with application to joint correction of motion and slice timing in fMRI. *IEEE Trans Med Imaging*. 30: 1546–1554.

33. Behzadi Y, Restom K, Liau J, Liu TT (2007): A Component Based Noise Correction Method (CompCor) for BOLD and Perfusion Based fMRI. *Neuroimage*. 37: 90–101.

34. Behzadi Y, Restom K, Liau J, Liu TT (2007): A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*. 37: 90–101.

35. Ashburner J (2007): A fast diffeomorphic image registration algorithm. *Neuroimage*. 38: 95–113.

36. Whitfield-Gabrieli S (2009): *Artifact Detection Tool*. MIT. Retrieved from http://web.mit.edu/swg/art/art.pdf.

37. Nielsen F ˚arup, Hansen LK (2002): Automatic anatomical labeling of Talairach coordinates and generation of volumes of interest via the BrainMap database. *Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, 2002*. .

38. Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003): An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage*. 19: 1233–1239.

39. 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.

40. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011): Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods*. 8: 665–670.

41. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM (2012): FSL. *Neuroimage*. 62: 782–790.

42. Bischoff-Grethe A, McCurdy D, Grenesko-Stevens E, Irvine LEZ, Wagner A, Yau W-YW, *et al.* (2013): Altered brain response to reward and punishment in adolescents with Anorexia nervosa. *Psychiatry Res*. 214: 331–340.

43. Jappe LM, Frank GKW, Shott ME, Rollin MDH, Pryor T, Hagman JO, *et al.* (2011): Heightened Sensitivity to Reward and Punishment in Anorexia Nervosa. *Int J Eat Disord*. 44: 317–324.

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

45. Gläscher 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.

46. 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.

47. Chase HW, Kumar P, Eickhoff SB, Dombrovski AY (2015): Reinforcement learning models and their neural correlates: An activation likelihood estimation meta-analysis. *Cogn Affect Behav Neurosci*. 15: 435–459.

48. Hayes AF (2013): *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach*. Methodology in the social sciences. New York, NY: Guilford Press.

49. Brett M, Anton J-L, Valabregue R, Poline J-B (2002): Region of interest analysis using the MarsBar toolbox for SPM 99. *Neuroimage*. 16: S497.