



Published in final edited form as:

Int J Eat Disord. 2017 April ; 50(4): 415–423. doi:10.1002/eat.22709.

Decreased Feedback Learning in Anorexia Nervosa Persists After Weight Restoration

Karin Foerde^a and Joanna E. Steinglass^{a,b}

^aNew York State Psychiatric Institute, New York, NY

^bDepartment of Psychiatry, Columbia University Medical Center, New York, NY

Abstract

Objective—Anorexia Nervosa (AN) is a serious disorder, with a mortality rate the highest of any psychiatric illness. It is notoriously challenging to treat and mechanisms of illness are not well understood. Reward system abnormalities have been proposed across theoretical models of the persistence of AN. Feedback learning is an important component of how reward systems shape behavior and we hypothesized that individuals with AN would show poorer learning from feedback.

Method—We administered the Acquired Equivalence Task to measure both learning from incremental feedback and generalization of that learning to novel stimuli. Participants were individuals with AN ($n=36$) before and after intensive weight restoration treatment and healthy controls (HC, $n=26$) tested twice. Performance was assessed as accuracy during the Learning and Test phases, for both trained and novel stimuli. The relationship between task performance and eating disorder severity at baseline was also assessed.

Results—Both before and after treatment, individuals with AN showed reduced learning from feedback in the Learning phase ($F_{3,180}=2.75$, $p=0.048$) and lower accuracy during the Test phase ($F_{1,60}=4.29$, $p=0.043$), as compared with HC. Individuals with AN did not differ from HC in accuracy for novel stimuli ($F_{1,60}=1.04$, $p=0.312$), indicating no deficit in generalization. Decreased acquisition of feedback learning was associated with longer illness duration and with greater eating disorder symptom severity at baseline.

Discussion—Individuals with AN show reduced learning from feedback or reinforcement, which may contribute to difficulties in changing maladaptive behaviors.

Keywords

feedback learning; acquired equivalence; reward; cognitive neuroscience; neuropsychology; anorexia nervosa; eating disorders; longitudinal

Anorexia Nervosa (AN) is a serious disorder, with a mortality rate among the highest of any psychiatric illness (1, 2). AN is characterized by persistent restrictive intake that is inadequate to maintain a healthy body weight, combined with a disturbance in body image (3); it is notoriously challenging to treat, and for older adolescents and adults, no

psychological or pharmacologic intervention has emerged as a treatment of choice (4). As many as 50% of hospitalized patients with AN require rehospitalization within a year of discharge (5), and the mortality rate increases with duration of illness (6, 7). Treatment that helps an individual succeed with weight restoration leads to improvement in many psychiatric symptoms—including mood and anxiety—yet the core behavioral disturbance in eating behavior (i.e., restricted caloric intake) persists even after intensive treatment (8, 9). Neuropsychological tasks can be used as a probe of brain systems, and can help clarify mechanisms that contribute to the persistent and entrenched nature of illness.

Abnormalities in reward systems are a common element in theoretical models attempting to elucidate the persistence and maintenance of this complex illness, even though the models differ in emphasis, (10-12). There is mounting evidence that neural processing of reward is anomalous in AN (13-15) and neuroimaging studies have found differences compared with healthy controls (HC) in response to rewarding stimuli (both hyper- (14, 16) and hypo-activation (17) in fMRI). Additionally, it has been proposed that individuals with AN may experience “reward contamination,” such that stimuli generally thought to be aversive are experienced as rewarding and vice versa (16, 18). These studies suggest that reward processes differ in AN (as compared with healthy controls), yet reward processing is a multifarious construct and the aspects of response to reward and learning from reward that may be abnormal are not well characterized.

Behavior is shaped through reward processing via learning from feedback. Through direct experience with the consequences of behavior, individuals learn from rewards and feedback to adapt choices in the future. Moreover, once learned, memories can be used to generalize to similar and overlapping contexts and guide responses to stimuli that have not been directly experienced (19). This raises questions about how maladaptive behavior forms and persists. It has been proposed that reward-based learning resulting in habit formation is central to the persistence of maladaptive eating behavior in AN (10, 11): that is, restrictive food intake becomes a habit. Once established, the core behaviors of AN are remarkably resistant to change, even among those seeking treatment. Individuals with AN have great difficulty learning new behaviors or changing maladaptive behaviors. In this way, individuals, once ill, seem impaired at learning from feedback. Additionally, stereotyped behaviors are thought to be responses specific to stimuli and are not easily applied in novel contexts. In contrast, generalization of learning from one context to another is important for normal, flexible behavior. Many maladaptive behaviors appear stereotyped and generalization of learning in AN merits further exploration. Furthermore, acutely ill patients with AN are also, by definition, significantly underweight and starvation itself may contribute to problems with reward-based learning (20-22). Reward based choices and neural reward systems have been shown to change with weight restoration in AN (13), raising the question of whether learning from feedback changes with acute treatment and weight restoration.

Multiple brain systems are involved in learning (23), and the systems contributing to feedback learning and generalization have been dissociated both behaviorally and neurally (19, 24) using the Acquired Equivalence task—a learning task designed to measure both learning from feedback and the ability to generalize learning to novel contexts. In the

Acquired Equivalence task, participants are first trained to learn associations between different images (stimuli) through feedback; participants are then tested on whether they have learned these associations *and* how well they are able to generalize what they have learned to new sets of associations (see Figure 1). Previous studies comparing patients with striatal or medial temporal lobe damage have elucidated components of learning and memory processes, showing that feedback learning depends critically on the striatum, whereas generalization relies on the hippocampus (19, 24). The task has been used in populations with schizophrenia (25), substance abuse (26), and Alzheimer's Disease (27) to demonstrate learning abnormalities associated with psychopathology and neuropathology, and is therefore well-suited to test these learning mechanisms in individuals with AN.

In this study, we tested the main hypotheses that: 1) individuals with AN, as compared with healthy peers, show reduced feedback learning and 2) this persist after weight restoration. In addition, we explored the hypotheses that 3) individuals with AN have reduced generalization of learning and 4) reduced feedback learning is related to severity of illness in AN.

METHODS

Participants

Participants were individuals with AN and HC who presented to the Columbia Center for Eating Disorders (Table 1). Eligible patients were between 16 and 45 years old, met DSM-5 (3) criteria for AN—restricting (AN-R) or binge-purge (AN-BP) subtype, and were receiving inpatient treatment at the New York State Psychiatric Institute (NYSPI). HC were recruited through the community and were compensated for participation. Individuals were excluded if they had an estimated IQ less than 80 (measured by Wechsler Test of Adult Reading; WTAR, Wechsler (28)), a history of a neurological, bipolar, or psychotic disorder, or substance abuse in the last 6 months. Anxiety and depressive disorders were not an exclusion, as these commonly co-occur with AN (29). HC were group-matched for age, sex, and ethnicity and were included if they had no psychiatric or significant medical illness, no psychotropic medications, and had a body mass index (BMI) in the normal range (18-25 kg/m²). This study was approved by the NYSPI Institutional Review Board, and after complete description of the study to the participants, written informed consent was obtained.

Procedures

Diagnoses were established via Semi Structured Interview for DSM-IV (SCID; First, Spitzer (30)), with the exception that amenorrhea was not required, consistent with DSM-5 criteria) and Eating Disorder Examination (EDE; Fairburn and Cooper (31)). Testing occurred twice. Individuals with AN were tested within one week of hospital admission (Session 1) and again after weight restoration treatment to 90% ideal body weight (32) (Session 2). Time between sessions was group matched for HC ($M_{HC}=61.4\pm 37.6$ days, $M_{AN}=53.7\pm 15.1$ days, $t(30.87)=0.997$, $p=0.33$). Height and weight were obtained on a beam balance scale (Detecto, Webb, MO) on the day of testing. Severity of eating disorder psychopathology was measured by the Eating Disorder Examination Questionnaire (EDE-Q; Fairburn (33)) and Eating Disorder Inventory (EDI; Garner and Olmsted (34)). Three subscales of the EDI were

examined—Drive for Thinness, Bulimia, and Body Dissatisfaction—as these three subscales are most relevant for AN psychopathology (35).

Neurocognitive battery

To assess global cognitive functioning, participants were administered a cognitive battery including the WTAR (estimated IQ) (28), the Stroop Task (attention and cognitive control) (36), Trail Making Test (TMT, attention) (37), and Letter Number Sequencing (LNS; a working memory subtest of the Wechsler Adult Intelligence Scale-III (WAIS-III)) (38) in counterbalanced order. Details of task administration are provided in the Supplement. Higher scores indicate better performance.

Task

The Acquired Equivalence Task ((19); Fig. 1) was presented using Matlab (Natick, Massachusetts) and the Psychophysics toolbox (39). In the Learning phase, participants were presented with a face and two landscapes and learned which landscape the face “preferred” by trial and error (associative learning). Participants indicated by button press whether the face “preferred” the left or right landscape and received feedback (the word “Correct” or “Incorrect,” Fig. 1A). The response window was 2.5 seconds. Once a choice was made, the selected option was highlighted and remained on the screen until the end of the response window. Next, feedback was displayed for 1.5 seconds followed by a 0.5 second intertrial interval.

Eight face and eight landscape images were used. The associations between faces and scenes were structured such that two faces (F_1 , F_2) were paired with the same two scenes (S_1 , S_2). Four associations stem from this relationship: F_1 - S_1 , F_2 - S_1 , F_1 - S_2 , and F_2 - S_2 . Two additional faces were presented with the same scene stimuli, but with opposite preferences, so that no single scene was always a correct response. Only three of the four associations were trained in the Learning phase (F_1 - S_1 , F_2 - S_1 , F_1 - S_2 ; Fig. 1B), resulting in 12 individual associations across the eight faces and eight landscapes; associations were presented in an interleaved fashion throughout training (19). Participants completed 96 learning trials across four blocks (eight repetitions of each association).

In the Test phase, trials included the trained associations (Trained trials) as well as the novel untrained associations (F_2 - S_2 ; Generalized trials; Fig. 1C). The Test phase comprised 96 trials: 72 Trained trials and 24 Generalized trials (six repetitions of each association). No feedback was provided in the Test phase.

Two versions of the task were used with different combinations of stimuli, such that different associations had to be learned at Session 1 versus Session 2. Performance was assessed by accuracy and response times (see details below).

Data processing and analysis

Clinical characteristics and neuropsychological tasks were compared using Student t tests for independent samples. Performance accuracy on the Acquired Equivalence Task Learning phase was measured as the proportion of correct trials on each of four blocks of 24 trials. In

the Test phase, proportion correct was averaged separately for Trained and Generalized trials. Because generalization of learning is limited by how well an individual has learned the trained associations, we also assessed generalization relative to the performance reached by each individual participant by dividing their performance score on trials with novel associations by their performance score on directly learned associations (Generalized/Trained). Response times were calculated as the median response times on correct trials for each block in the Learning phase, and for Trained and Generalized trials separately in the Test phase (see Supplemental materials for results of response time analyses). Data were analyzed using mixed ANOVAs with Huynh–Feldt correction for nonsphericity. Effect size was measured using both partial eta squared (η_p^2) and generalized eta (η_G^2) squared as recommended for mixed ANOVAs (40-42). For the partial eta-squared small, medium and large effect sizes are 0.01, 0.06, and 0.14 respectively, whereas 0.02, 0.13, and 0.26 are suggested for generalized eta squared (40).

To assess correlations between learning and clinical characteristics, accuracy was averaged across Learning phases at Session 1 and Session 2 to generate an overall learning score per individual. Relationships between overall accuracy and measures of eating disorder severity (EDE-Q, EDI-Drive for Thinness, EDI-Bulimia, EDI-Body Dissatisfaction, BMI, and duration of illness) in the AN group at baseline were assessed using Pearson correlations or Spearman rank-order correlations for non-normal data. For the exploratory analyses of correlations between learning and eating disorder severity, Bonferroni correction for multiple comparisons results in an alpha level of 0.008 (0.05/6). Data were analyzed using the IBM SPSS Statistics 22 analysis package.

One hundred one individuals met inclusion criteria for the study (61 AN, 40 HC). Data were lost from one AN and two HC due to computer errors, 13 AN and eight HC did not participate in Session 2 (12 AN left treatment prior to weight restoration, one AN did not want to participate at Session 2), leaving 77 participants (47 AN, 30 HC) with data at both Session 1 and 2. The proportion of individuals participating only once did not differ significantly between groups ($\chi^2(1, N=98)=0.005, p=1$; further information about the group that did not participate at Session 2 is provided in the Supplement).

Response rates and accuracy from 77 participants (47 AN, 30 HC) who completed both Sessions were examined. We excluded participants if they responded on fewer than 90% of trials (4 AN, 2 HC). Following standard procedure (43), we excluded participants who showed less than 50% accuracy during the Test phase on Trained trials (7 AN and 2 HC), i.e., those who failed to demonstrate any learning. Data from these non-learners were excluded from analyses to aid interpretation of performance on the novel trials in the Test phase, as “generalization” on those trials relies on learning of the trained associations in the Learning phase (see Fig. 1). The proportion of participants excluded based on poor performance did not differ significantly between groups ($\chi^2(1, N=77)=1.18, p=0.38$) and the pattern of results reported below did not differ when including all 77 participants (see Supplemental analyses). Both AN and HC individuals excluded due to poor performance showed poorer cognitive performance on neuropsychological tests (see Supplement).

Sixty-two participants, 36 AN and 26 HC, were included in the analyses reported here (AN, restricting subtype $n=19$, AN, binge-purge subtype $n=17$; subtype comparisons are provided in the Supplement). The AN and HC groups did not differ in age, education, gender, ethnicity, and measures of general cognitive functioning (see Table 1). As expected, groups differed in BMI and psychological measures at Session 1. In the AN group, 11 participants had a comorbid diagnosis. Seven participants had an anxiety disorder (2 specific phobia, 3 social phobia, 4 PTSD, 1 OCD, 1 panic disorder), three had a depressive disorder (2 MDD, 1 dysthymia); some participants had multiple diagnoses. Four participants were taking antidepressant medication (SSRI or SNRI) and one was taking a benzodiazepine.

RESULTS

Learning phase

Performance across the learning phases was examined in a 2 (Session 1/Session 2) \times 2 (HC/AN) \times 4 (Block) repeated measures ANOVA. There was no main effect of Diagnosis ($F(1,60)=1.72$, $p=0.195$). There was a main effect of Block, indicating that, as expected, participants improved performance with practice ($F(3,180)=93.31$, $p < 0.001$, $\eta_p^2=0.609$, $\eta_G^2=0.24$). There was a significant interaction between Block and Diagnosis ($F(3,180)=2.75$, $p=0.048$, $\eta_p^2=0.044$, $\eta_G^2=0.009$), suggesting that individuals with AN exhibited reduced feedback learning relative to HC. As seen in Figure 2A, patients with AN and HC performed similarly at the beginning of learning, but the AN group improved less than HC. There were no effects of Session or interactions with Session ($ps > 0.14$), suggesting that the decreased feedback learning in individuals with AN was present both before and after treatment.

Test phase

In the Test phase, performance was compared between trials in which the associations had been experienced directly in the Learning phase (Trained, Fig. 1) and novel trials in which generalization to new pairs was required (Generalized) using a 2 (Session 1/Session 2) \times 2 (HC/AN) \times 2 (Trial type: Trained/Generalized) mixed ANOVA (Fig. 2B). There were main effects of Diagnosis ($F(1,60)=4.29$, $p=0.043$, $\eta_p^2=0.067$, $\eta_G^2=0.038$) and Trial type ($F(1,60)=97.23$, $p < 0.001$, $\eta_p^2=0.618$, $\eta_G^2=0.24$), but no interaction between Diagnosis and Trial type ($F(1,60)=0.74$, $p=0.40$). This pattern of results indicated that, as in the Learning phase, AN participants' feedback learning was poorer than that of healthy controls, but generalization of learning on trials with novel associations was not specifically worse. As in the Learning phase, there was no main effect of Session ($F(1,60)=0.66$, $p=0.42$), or any interactions with Session ($ps > 0.2$).

Relative generalization accuracy (Generalized trials divided by Trained trials) did not differ between groups ($F(1,60)=1.04$, $p=0.312$; Fig. 2C), suggesting that AN generalized learning at a level similar to controls when taking into account how well they had learned the associations.

Changes in clinical severity with treatment

As expected, weight restoration treatment led to a significant increase in BMI from Session 1 to Session 2 in the AN group (16.8 ± 1.3 vs. 20.3 ± 0.7 , $t(35) = -17.5$, $p < 0.001$). Several additional measures of clinical severity improved from Session 1 to Session 2: EDE-Q Global score (4.0 ± 1.5 vs. 2.6 ± 1.4 , $t(34) = 5.7$, $p < 0.001$); EDI-Drive for Thinness (12.8 ± 6.9 vs. 10.2 ± 7.0 , $t(34) = 2.3$, $p = 0.029$); EDI-Bulimia (2.9 ± 4.4 vs. 0.9 ± 1.7 , $t(34) = 3.1$, $p = 0.003$). EDI-Body Dissatisfaction did not change significantly from Session 1 to Session 2 (15.9 ± 7.8 vs. 17.0 ± 8.1 , $t(33) = -0.9$, $p = 0.34$).

Associations between task performance and clinical severity

Within the AN group, feedback learning (across Session 1 and 2) showed a significant negative correlation with eating disorder severity, as measured by the EDE-Q Global score at Session 1 ($\rho(34) = -0.50$, $p = 0.002$, Fig. 3A). The correlations between overall feedback learning and the EDI subscales were all negative (Drive for Thinness ($\rho(34) = -0.35$, $p = 0.039$), Body Dissatisfaction ($\rho(33) = -0.39$, $p = 0.021$), and Bulimia ($\rho(34) = -0.34$, $p = 0.044$)), but did not reach significance when correcting for multiple comparisons. BMI at baseline (Session 1) was not correlated with feedback learning ($\rho(34) = -0.07$, $p = 0.69$).

Longer illness duration was associated with poorer performance on overall feedback learning ($\rho(34) = -0.35$, $p = 0.037$; Fig. 3B), but this correlation did not reach the threshold of significance when correcting for multiple comparisons. Duration of illness was strongly associated with age, as expected ($\rho(34) = 0.74$, $p < 0.001$). However, feedback learning was not related to age among HC ($\rho(34) = 0.09$, $p = 0.66$) or among AN ($\rho(34) = -0.32$, $p = 0.058$).

DISCUSSION

Individuals with AN showed decreased acquisition of learning from feedback (compared with HC) both before and after weight restoration treatment. Moreover, the reduction in feedback learning was associated with eating disorder psychopathology, as suggested by associations with some eating disorder symptom severity scales and with illness duration (though not with BMI). In contrast, relative to their feedback learning performance, patients with AN performed similarly to healthy controls when generalizing from what they had learned. Thus, individuals with AN had poorer feedback learning, which was not remediated by weight restoration and intensive treatment, yet were able to generalize learning and apply what they have learned to novel stimuli. Psychological symptoms did improve with weight restoration treatment, yet feedback learning did not. In this population, AN and HC were well matched on measures of global cognition (IQ, working memory, attention), suggesting a circumscribed cognitive abnormality.

Our results come from a behavioral task, and do not directly assess brain function. Nonetheless, the findings are consistent with other studies that have investigated forms of striatal-based learning in AN (44-46). Additionally, previous studies using the Acquired Equivalence task in patients with Parkinson's disease, who are known to have striatal dysfunction, have shown a pattern of learning impairment similar to that seen in the present study (24), suggesting that striatal function may be impaired in AN.

Functional neuroimaging studies in AN also suggest striatal abnormalities. In one study, individuals with AN showed equivalent striatal response for receipt of monetary gains and losses during a card guessing task (14). In a study of temporal discounting, individuals with AN showed decreased striatal activity when choosing delayed rewards before treatment, and increased striatal activity after treatment (13). These studies and others like them measure neural systems directly, yet do not assess learning and therefore do not assess how feedback and rewards may be used to shape behavior. Our results suggest that the fronto-striatal abnormalities found in prior studies may have consequences for learning, and thereby for behavior, and that these do not improve with short-term weight gain.

In contrast to the compromised feedback learning, and counter to our hypothesis, we found that generalization processes thought to rely on the medial temporal lobe (hippocampus) were intact both before and after weight restoration. This finding is consistent with a study by Lawrence et al. (46), who used a learning task sensitive to striatal dysfunction and a memory test sensitive to medial temporal lobe damage in a small sample of individuals with AN, tested while underweight, and found a dissociation between learning thought to depend on the striatum vs. the medial temporal lobes. Here we find that this pattern remains after weight restoration. While hippocampal volume abnormalities have been demonstrated in AN (47), such abnormalities may not be associated with changes in cognitive function (48). Thus a host of neural changes may accompany AN, but it is critical to link neural changes to relevant cognition and behavior.

If feedback learning is indeed compromised in AN, a central question is, how is this related to maladaptive behavior in AN? We have proposed that persistent maladaptive restrictive intake is related to overlearned behavior and related to habit mechanisms (11, 49). The maladaptive, yet very effective, learning that may occur when establishing AN appears paradoxical in light of our finding of decreased feedback learning. One possibility is that maladaptive behaviors may be established during a developmental window of learning sensitivity. For example, adolescence, a time when maladaptive eating behavior is most commonly established (50), also represents a time of accelerated learning. Hypersensitivity to reward in adolescence has been established across many studies (51), and a recent study demonstrated enhanced learning from feedback during adolescence (52). If so, it may be that adolescence is a window during which it is very easy to acquire new behavioral patterns. However, the learning mechanisms supporting this learning may change as a result of normal development (aging). Or, perhaps having AN affects neurocognitive processes such that feedback learning worsens the longer an individual is affected by illness, making it increasingly difficult to overcome maladaptive behavior. Our data lend some support to the idea that illness may affect learning by showing a potential relationship between reduced feedback learning and duration of illness. Although duration of illness and age are highly related, the absence of an association between age and feedback learning in HC compels further consideration of this relationship in AN. In particular it will be important to assess learning across a broader range of illness duration paired with a broader range of illness onset to disentangle the contributions of illness and age.

Decreased learning from feedback may have implications for treatment response. If the presence of a feedback learning problem at baseline (regardless of how and when the

behavioral pattern is acquired) is associated with failure to respond to treatment, it could be a potential biomarker for a more chronic course of illness. Treatment interventions may be improved by considering the possibility of reinforcement learning difficulties. Individuals with poor learning from feedback may take longer, or require more intensive feedback, to integrate the behavioral learning that is central to treatment of eating disorders. Distinguishing between feedback learning difficulties at baseline versus acquired as a consequence of disease is challenging, and points to the need for more longitudinal studies of individuals with AN.

Limitations and future directions

This study found that individuals with AN had reduced learning from feedback relative to HC, in the absence of impairment in ability to generalize from what had been learned. This is consistent with reward-based models of AN (10-12). To better evaluate these models, this finding needs to be replicated in other samples of individuals with AN, and the link between feedback learning and eating behavior needs to be examined. Additionally, tasks that include losses would allow for better evaluation of learning impairment related to negative feedback that cannot be distinguished in the Acquired Equivalence task.

Data from the current study need to be considered cautiously in light of the following limitations. Menstrual cycle influences on learning (53) were not controlled for. It would be beneficial to assess feedback learning after more long-term recovery and restoration of menstrual cycle has occurred and to study participants during the early follicular phase. Additionally, the finding of decreased performance in the learning phase represented a small effect, and the association with duration of illness did not survive correction for multiple comparisons. These findings, therefore, warrant further study to replicate the findings and test the hypothesis that there is a relationship with chronicity of illness.

In this initial study of feedback learning before and after treatment, we aimed to examine an inclusive and representative clinical sample of individuals diagnosed with AN. As a result the sample is heterogeneous. Studies of more homogeneous samples, or with larger samples would be needed to test hypotheses about subtypes of AN.

Further consideration should also be given to the role of comorbid depression and anxiety in learning impairments. Previous work examining similar learning in patients with Parkinson's disease found particular learning impairments in those with depressive illness (54). Individuals with AN in the present study were significantly more depressed and anxious, as measured with the BDI and STAI, than HC (Table 1). BDI and STAI scores improved with treatment, whereas learning performance did not. In AN, anxiety and depression are likely tightly linked with overall illness severity, and in this sample BDI and STAI were both positively correlated with EDE-Q Global (BDI: $\rho=0.65$, $p < 0.001$; STAI: $\rho= 0.55$, $p=0.001$) and with each other ($\rho=0.71$, $p<0.001$). It is therefore difficult to discern whether there are important distinct associations. Future work should further assess the contributions of depression and anxiety to learning.

The data from this study suggest a relationship between reduced feedback learning and clinical features during the intensive treatment phase of illness. However, the current study

did not directly measure the involvement of brain systems and future studies are needed to directly link fronto-striatal dysfunction observed in AN to learning and to probe the role of fronto-striatal circuits in the persistence of AN.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Yaakov Stern, PhD, for his mentorship on this project; B. Timothy Walsh, MD and Daphna Shohamy, PhD, for their thoughtful contributions; and the patients and staff of the Eating Disorders Research Unit at the New York State Psychiatric Institute.

This work was supported by National Institute of Mental Health (J.E.S., grant number K23MH076195). The content of this study is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Arcelus J, Mitchell AJ, Wales J, Nielsen S. Mortality Rates in Patients With Anorexia Nervosa and Other Eating Disorders: A Meta-analysis of 36 Studies. *Archives of General Psychiatry*. 2011; 68(7):724–31. [PubMed: 21727255]
2. Rosling AM, Sparen P, Norring C, von Knorring AL. Mortality of eating disorders: a follow-up study of treatment in a specialist unit 1974-2000. *Int J Eat Disord*. 2011; 44(4):304–10. [PubMed: 21472749]
3. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders FE*. Arlington, VA: American Psychiatric Association; 2013.
4. Watson HJ, Bulik CM. Update on the treatment of anorexia nervosa: review of clinical trials, practice guidelines and emerging interventions. *Psychological medicine*. 2013; 43(12):2477–500. [PubMed: 23217606]
5. Pike KM. Long-term course of anorexia nervosa: response, relapse, remission, and recovery. *Clin Psychol Rev*. 1998; 18(4):447–75. [PubMed: 9638357]
6. Sullivan PF. Mortality in anorexia nervosa. *American Journal of Psychiatry*. 1995; 152(7):1073–5. [PubMed: 7793446]
7. Fichter MM, Quadflieg N. Mortality in eating disorders - results of a large prospective clinical longitudinal study. *Int J Eat Disord*. 2016; 49(4):391–401. [PubMed: 26767344]
8. Mayer LE, Schebendach J, Bodell LP, Shingleton RM, Walsh BT. Eating behavior in anorexia nervosa: before and after treatment. *Int J Eat Disord*. 2012; 45(2):290–3. [PubMed: 21495053]
9. Sysko R, Walsh BT, Schebendach J, Wilson GT. Eating behavior among women with anorexia nervosa. *Am J Clin Nutr*. 2005; 82(2):296–301. [PubMed: 16087971]
10. O'Hara CB, Campbell IC, Schmidt U. A reward-centred model of anorexia nervosa: A focussed narrative review of the neurological and psychophysiological literature. *Neuroscience and biobehavioral reviews*. 2015; 52:131–52. [PubMed: 25735957]
11. Walsh BT. The enigmatic persistence of anorexia nervosa. *Am J Psychiatry*. 2013; 170(5):477–84. [PubMed: 23429750]
12. Godier LR, Park RJ. Compulsivity in anorexia nervosa: a transdiagnostic concept. *Frontiers in psychology*. 2014; 5:778. [PubMed: 25101036]
13. Decker JH, Figner B, Steinglass JE. On Weight and Waiting: Delay Discounting in Anorexia Nervosa Pretreatment and Posttreatment. *Biol Psychiatry*. 2015; 78(9):606–14. [PubMed: 25641636]
14. Wagner A, Aizenstein H, Venkatraman VK, Fudge J, May JC, Mazurkewicz L, et al. Altered reward processing in women recovered from anorexia nervosa. *American Journal of Psychiatry*. 2007; 164(12):1842–9. [PubMed: 18056239]

15. Frank GK, Reynolds JR, Shott ME, Jappe L, Yang TT, Tregellas JR, O'Reilly RC. Anorexia nervosa and obesity are associated with opposite brain reward response. *Neuropsychopharmacology*. 2012; 37(9):2031–46. [PubMed: 22549118]
16. Cowdrey FA, Park RJ, Harmer CJ, McCabe C. Increased neural processing of rewarding and aversive food stimuli in recovered anorexia nervosa. *Biol Psychiatry*. 2011; 70(8):736–43. [PubMed: 21714958]
17. Holsen LM, Lawson EA, Blum J, Ko E, Makris N, Fazeli PK, et al. Food motivation circuitry hypoactivation related to hedonic and nonhedonic aspects of hunger and satiety in women with active anorexia nervosa and weight-restored women with anorexia nervosa. *J Psychiatry Neurosci*. 2012; 37(5):322–32. [PubMed: 22498079]
18. Keating C. Theoretical perspective on anorexia nervosa: the conflict of reward. *Neuroscience and biobehavioral reviews*. 2010; 34(1):73–9. [PubMed: 19619579]
19. Shohamy D, Wagner AD. Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron*. 2008; 60(2):378–89. [PubMed: 18957228]
20. Carroll ME, France CP, Meisch RA. Intravenous self-administration of etonitazene, cocaine and phencyclidine in rats during food deprivation and satiation. *The Journal of pharmacology and experimental therapeutics*. 1981; 217(2):241–7. [PubMed: 6112257]
21. Carroll ME, Meisch RA. The effects of feeding conditions on drug-reinforced behavior: maintenance at reduced body weight versus availability of food. *Psychopharmacology*. 1980; 68(2):121–4. [PubMed: 6107944]
22. Thanos PK, Michaelides M, Piyis YK, Wang GJ, Volkow ND. Food restriction markedly increases dopamine D2 receptor (D2R) in a rat model of obesity as assessed with in-vivo muPET imaging ([11C] raclopride) and in-vitro ([3H] spiperone) autoradiography. *Synapse*. 2008; 62(1):50–61. [PubMed: 17960763]
23. Squire LR. Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. *J Cogn Neurosci*. 1992; 4(3):232–43. [PubMed: 23964880]
24. Myers CE, Shohamy D, Gluck MA, Grossman S, Kluger A, Ferris S, et al. Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *J Cogn Neurosci*. 2003; 15(2):185–93. [PubMed: 12676056]
25. Shohamy D, Mihalakos P, Chin R, Thomas B, Wagner AD, Tamminga C. Learning and generalization in schizophrenia: effects of disease and antipsychotic drug treatment. *Biol Psychiatry*. 2010; 67(10):926–32. [PubMed: 20034612]
26. Mattyassy A, Keri S, Myers CE, Levy-Gigi E, Gluck MA, Kelemen O. Impaired generalization of associative learning in patients with alcohol dependence after intermediate-term abstinence. *Alcohol Alcohol*. 2012; 47(5):533–7. [PubMed: 22582184]
27. Bodi N, Csibri E, Myers CE, Gluck MA, Keri S. Associative learning, acquired equivalence, and flexible generalization of knowledge in mild Alzheimer disease. *Cogn Behav Neurol*. 2009; 22(2):89–94. [PubMed: 19506424]
28. Wechsler, D. Wechsler Test of Adult Reading. San Antonio, TX: The Psychological Corporation; 2001.
29. Hudson JI, Hiripi E, Pope HG Jr, Kessler RC. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biological Psychiatry*. 2007; 61(3):348–58. [PubMed: 16815322]
30. First, MB., Spitzer, RL., Gibbon, M., Williams, JB. Structured Clinical Interview for DSM-IV-TR Axis I Disorders-Patient Edition (SCID-I/P). New York, NY: Biometrics Research Department, New York State Psychiatric Institute; 2002.
31. Fairburn, CG., Cooper, PJ. The Eating Disorder Examination. In: Fairburn, CG., Wilson, GT., editors. *Binge Eating: Nature, Assessment, and Treatment*. New York: Guilford Press; 1993. p. 317-60.
32. MetropolitanLifeInsurance. New weight standards for men and women. *Statistical Bulletin*. 1959; 40:1–4.
33. Fairburn, CG. *Cognitive Behavior Therapy and Eating Disorders*. New York: Guilford Press; 2008.
34. Garner, DM., Olmsted, MP. *The eating disorder inventory manual*. Odessa, FL: Psychological Assessment Resources; 1984.

35. Garner DM, Olmsted MP, Polivy J. Development and validation of a multidimensional eating disorder inventory for anorexia and bulimia. *International Journal of Eating Disorders*. 1983; 2:15–34.
36. Golden, C. Stroop color and word test: a manual for clinical and experimental uses. Wood Dale, IL: Stoelting; 1978.
37. Bowie CR, Harvey PD. Administration and interpretation of the Trail Making Test. *Nature protocols*. 2006; 1(5):2277–81. [PubMed: 17406468]
38. Wechsler, D. WAIS-III administration and scoring manual. San Antonio, TX: The Psychological Corporation; 1997.
39. Brainard DH. Psychophysics software for use with MATLAB. *Spatial Vision*. 1997; 10:433–6. [PubMed: 9176952]
40. Bakeman R. Recommended effect size statistics for repeated measures designs. *Behav Res Methods*. 2005; 37(3):379–84. [PubMed: 16405133]
41. Olejnik S, Algina J. Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychol Methods*. 2003; 8(4):434–47. [PubMed: 14664681]
42. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology*. 2013; 4:863. [PubMed: 24324449]
43. Shohamy D, Myers CE, Onlaor S, Gluck MA. Role of the basal ganglia in category learning: how do patients with Parkinson's disease learn? *Behav Neurosci*. 2004; 118(4):676–86. [PubMed: 15301595]
44. Shott ME, Filoteo JV, Jappe LM, Pryor T, Maddox WT, Rollin MD, et al. Altered implicit category learning in anorexia nervosa. *Neuropsychology*. 2012; 26(2):191–201. [PubMed: 22201300]
45. Firk C, Mainz V, Schulte-Ruether M, Fink G, Herpertz-Dahlmann B, Konrad K. Implicit sequence learning in juvenile anorexia nervosa: neural mechanisms and the impact of starvation. *Journal of child psychology and psychiatry, and allied disciplines*. 2015; 56(11):1168–76.
46. Lawrence AD, Dowson J, Foxall GL, Summerfield R, Robbins TW, Sahakian BJ. Impaired visual discrimination learning in anorexia nervosa. *Appetite*. 2003; 40(1):85–9. [PubMed: 12631509]
47. Burkert NT, Koschutnig K, Ebner F, Freidl W. Structural hippocampal alterations, perceived stress, and coping deficiencies in patients with anorexia nervosa. *Int J Eat Disord*. 2015; 48(6):670–6. [PubMed: 25864963]
48. Connan F, Murphy F, Connor SE, Rich P, Murphy T, Bara-Carill N, et al. Hippocampal volume and cognitive function in anorexia nervosa. *Psychiatry research*. 2006; 146(2):117–25. [PubMed: 16510268]
49. Foerde K, Steinglass JE, Shohamy D, Walsh BT. Neural mechanisms supporting maladaptive food choices in anorexia nervosa. *Nat Neurosci*. 2015; 18(11):1571–3. [PubMed: 26457555]
50. Treasure J, Stein D, Maguire S. Has the time come for a staging model to map the course of eating disorders from high risk to severe enduring illness? An examination of the evidence. *Early Interv Psychiatry*. 2015; 9(3):173–84. [PubMed: 25263388]
51. Casey BJ, Getz S, Galvan A. The adolescent brain. *Developmental Review*. 2008; 28(1):62–77. [PubMed: 18688292]
52. Davidow JY, Foerde K, Galvan A, Shohamy D. An Upside to Reward Sensitivity: The Hippocampus Supports Enhanced Reinforcement Learning in Adolescence. *Neuron*. 2016; 92(1):93–9. [PubMed: 27710793]
53. Diekhof EK, Ratnayake M. Menstrual cycle phase modulates reward sensitivity and performance monitoring in young women: Preliminary fMRI evidence. *Neuropsychologia*. 2016; 84:70–80. [PubMed: 26471712]
54. Herzallah MM, Moustafa AA, Misk AJ, Al-Dweib LH, Abdelrazeq SA, Myers CE, Gluck MA. Depression impairs learning whereas anticholinergics impair transfer generalization in Parkinson patients tested on dopaminergic medications. *Cogn Behav Neurol*. 2010; 23(2):98–105. [PubMed: 20535058]

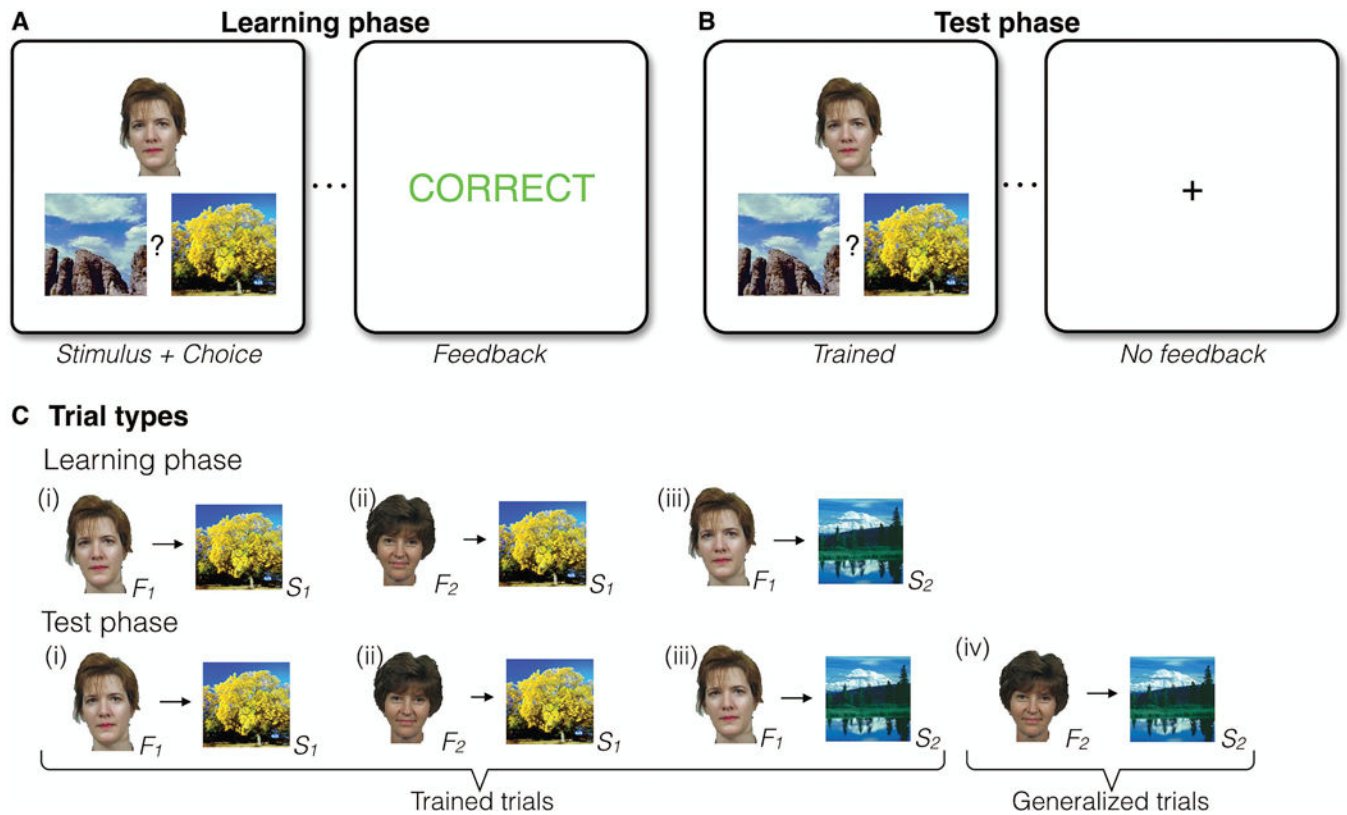


Figure 1. Task structure and trial types

(A) In the Learning phase, participants learned a series of individual face-scene associations based on feedback (12 individual associations in total). A face-scene pair was presented on each trial, followed by performance-dependent feedback. (B) The Learning phase was followed by a Test phase, where participants received no feedback and were asked to respond to both previously experienced (Trained) trials and novel untrained face-scene association (Generalized) trials. (C) The Learning phase had three event types: Individual associations shared overlapping features such that two faces always were associated with a common scene (i and ii), and one of those faces was also associated with a second scene (iii). A scene that was the incorrect choice for one face was the correct choice for another face, such that simple stimulus-response learning strategies could not support learning. In the Test phase, novel associations (iv, Generalized trials) were presented intermixed with trials that tested knowledge for previously trained associations (Trained trials).

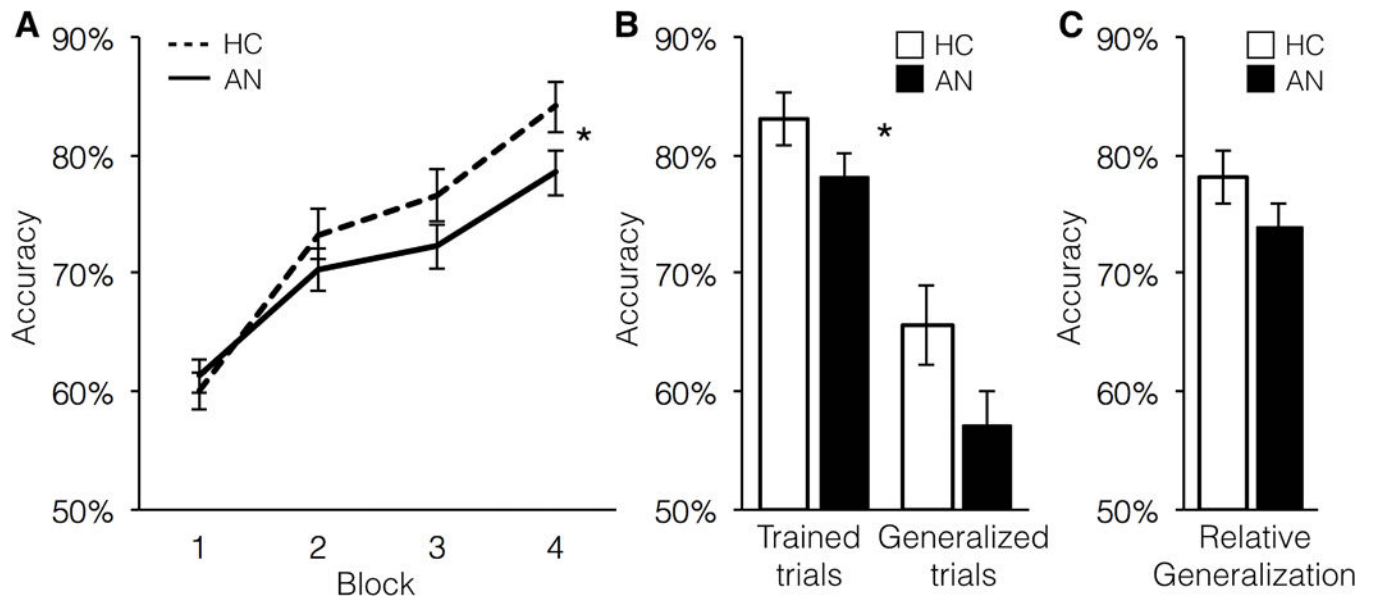


Figure 2. Acquired Equivalence Task Performance

(A) In the Learning phase, AN showed significantly decreased learning compared with HC ($F(3,180)=2.75, p=0.048$). Shown is performance collapsed across Sessions 1 and 2, as there was no effect of Session. (B) Accuracy in the Test phase was significantly lower for AN than HC ($F(1,60)=4.29, p=0.043$). As expected, participants were more accurate for the Trained trials than for the novel Generalized trials ($F(1,60)=97.23, p < 0.001$). The interaction was not significant, indicating that AN were impaired on both trial types, without a specific impairment in Generalization. (C) To account for baseline learning, groups were compared on Generalized trials divided by Trained trials; they did not differ significantly ($F(1,60)=1.04, p=0.312$). HC=healthy controls; AN=anorexia nervosa.

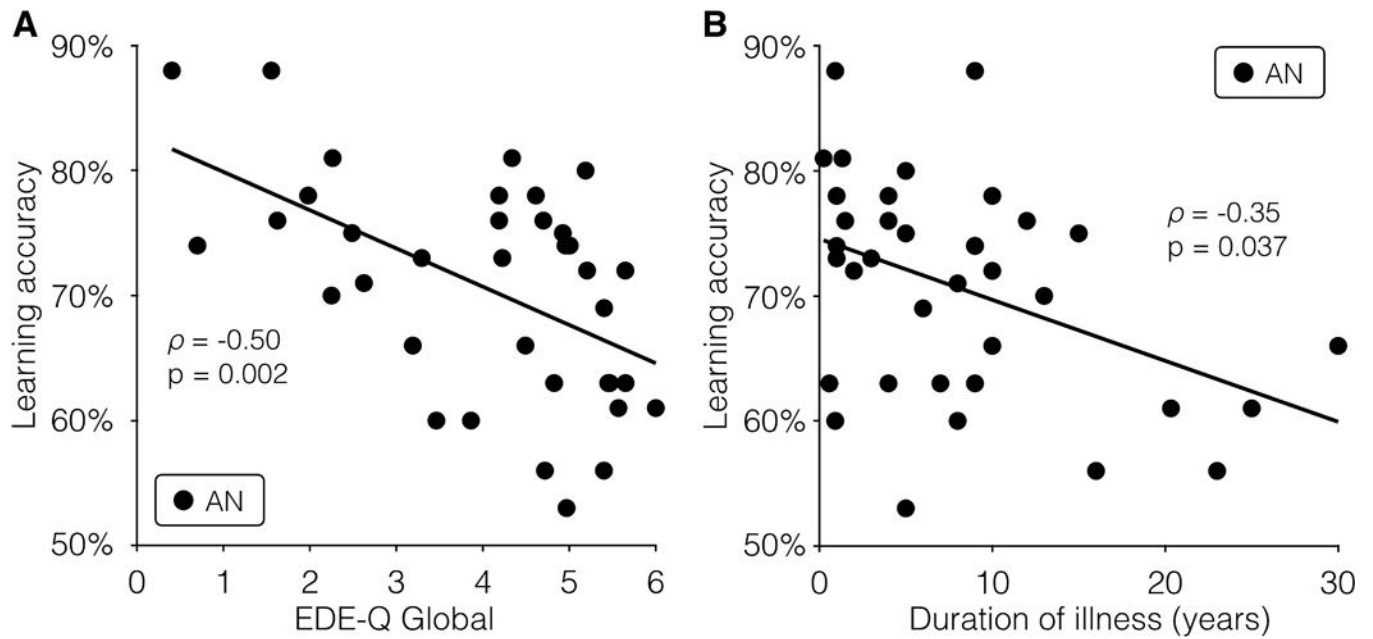


Figure 3. Task Performance and Clinical Severity

(A) Accuracy in the learning phase was significantly associated with the Eating Disorder Examination Questionnaire (EDE-Q) global score at Session 1 ($\rho(34)=-0.50$, $p=0.002$). (B) Accuracy in the learning phase was also associated with duration of illness ($\rho(34)=-0.35$, $p=0.037$) but this correlation did not reach the threshold of significance when correcting for multiple comparisons. AN=anorexia nervosa.

Table 1

Participant clinical characteristics at baseline (Session 1)

	HC (n=26)		AN (n=36)		HC v AN		<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
<i>Demographics</i>							
Age (years)	22.9	5.6	24.7	8.1			0.28
Gender (F/M)	24/2		35/1				
Ethnicity (C/AA+H/As)	22/3/1		34/1/1				
Years of education	15.1	3.0	14.0	2.2			0.11
Body Mass Index (kg/m ²)	21.6	1.7	16.8	1.3			<0.001
Duration illness (years)			8.1	7.4			
Subtype (Restricting/Binge-Purge)				19/17			
<i>Neuropsychological Tests</i>							
WTAR estimated IQ	110.5	11.4	108.8	7.9			0.51
LNS (Total number correct)*	12.2	2.2	11.6	2.2			0.36
Stroop (Interference)*	8.34	9.7	4.94	7.0			0.13
Trail Making Test*	0.46	0.2	0.56	0.3			0.15
<i>Psychological Measures</i>							
EDE-Q Global	0.1	0.2	4.0	1.5			<0.001
EDI-Drive for Thinness	0.3	0.8	12.8	6.9			<0.001
EDI-Body Dissatisfaction	1.6	3.1	16.2	7.8			<0.001
EDI-Bulimia scale	0.1	0.6	2.8	4.4			0.001
BDI	1.2	1.8	25.8	12.2			<0.001
STAI	27.5	5.7	59.6	9.8			<0.001

HC=Healthy control, AN=Anorexia Nervosa, F/M=Female/Male, C/AA+H/As=Caucasian/African American and/or Hispanic/Asian, WTAR=Wechsler Test of Adult Reading, LNS=Letter-Number Sequencing, EDE-Q=Eating Disorder Examination Questionnaire, EDI=Eating Disorder Inventory, Data were missing from 1 HC for WTAR, LNS, EDE-Q, and EDI. Data were missing from 1 AN for WTAR, Trail Making Test, and EDI-Body Dissatisfaction.

* Higher scores indicate better performance. See Supplement for scoring and metrics of neuropsychological tests.