

Diminished Adaptation, Satisfaction, and Neural Responses to Advantageous Social Signals in Anorexia Nervosa and Bulimia Nervosa

Yi Luo, Dustin Pluta, Brooks B. Brodrick, Jayme M. Palka, Jordan McCoy, Terry Lohrenz, Xiaosi Gu, Marina Vannucci, P. Read Montague, and Carrie J. McAdams

ABSTRACT

BACKGROUND: Development and recurrence of 2 eating disorders (EDs), anorexia nervosa and bulimia nervosa, are frequently associated with environmental stressors. Neurobehavioral responses to social learning signals were evaluated in both EDs.

METHODS: Women with anorexia nervosa ($n = 25$), women with bulimia nervosa ($n = 30$), or healthy comparison women ($n = 38$) played a neuroeconomic game in which the norm shifted, generating social learning signals (norm prediction errors [NPEs]) during a functional magnetic resonance imaging scan. A Bayesian logistic regression model examined how the probability of offer acceptance depended on cohort, block, and NPEs. Rejection rates, emotion ratings, and neural responses to NPEs were compared across groups.

RESULTS: Relative to the comparison group, both ED cohorts showed less adaptation ($p = .028$, $\eta_p^2 = 0.060$), and advantageous signals (positive NPEs) led to higher rejection rates ($p = .014$, $\eta_p^2 = 0.077$) and less positive emotion ratings ($p = .004$, $\eta_p^2 = 0.111$). Advantageous signals increased neural activations in the orbitofrontal cortex for the comparison group but not for women with anorexia nervosa ($p = .018$, $d = 0.655$) or bulimia nervosa ($p = .043$, $d = 0.527$). More severe ED symptoms were associated with decreased activation of dorsomedial prefrontal cortex for advantageous signals.

CONCLUSIONS: Diminished neural processing of advantageous social signals and impaired norm adaptation were observed in both anorexia nervosa and bulimia nervosa, while no differences were found for disadvantageous social signals. Development of neurocognitive interventions to increase responsivity to advantageous social signals could augment current treatments, potentially leading to improved clinical outcomes for EDs.

<https://doi.org/10.1016/j.bpsc.2023.10.010>

Eating disorders (EDs), such as anorexia nervosa (AN) and bulimia nervosa (BN), are complex psychiatric disorders with strong evidence for both psychosocial and biological risks (1). Several psychological constructs related to social comparisons are shared risk factors for the development of AN and BN, such as body dissatisfaction (2), low self-esteem (3), and social anxiety (4). Changes and stressors in the social environment have also been associated with the development of EDs (5,6). Deficits in social learning have been hypothesized as contributing to many types of psychopathology, including EDs (7–9). Neural impairments in social learning might provide a mechanism that explains how social environment contributes to EDs.

The norm-adaptation ultimatum game (UG) evaluates social learning and adaptation from norm violations. Adaptation refers to changing a behavior in response to a change in the environment. The UG asks responders to accept or reject offers to split money, akin to real-world scenarios such as dividing a sum among coworkers (e.g., cook, waiter) (10). Participants can learn to expect offers of a typical amount, and unpredictable shifts change rejection rates as participants

adapt (11,12). The norm prediction error (NPE), the difference between the expected offer (i.e., the social norm) and the actual offer, is a key social learning signal parametrically related to neural activations in the orbitofrontal cortex (OFC) (13). The valence of the NPE signals whether the outcome was better (positive) or worse (negative) than expected.

Internalizing beliefs about social norms requires the integration of stored beliefs about one's world with new information acquired through discrete interactions. Prior studies have found that women with AN and BN have altered neurobehavioral responses both when accessing beliefs about themselves and others (14,15) and when engaged in reciprocating interactions with a specific other (16,17). The purpose of this study was to determine how participants with EDs respond to changes in their social environment. We hypothesized that participants with EDs would show less adaptation and reduced neural activations to NPE in the OFC. Differences in processing NPEs were explored using whole-brain group comparisons and regressions examining acute psychiatric symptoms.

METHODS AND MATERIALS

Participants

A total of 93 women (biological sex at birth; 25 with AN, 30 with BN, and 38 healthy comparison [HC] participants) between 18 and 46 years were recruited and provided written informed consent as per the University of Texas Southwestern Institutional Review Board. The Eating Disorder Assessment for DSM-5 (18) assessed ED diagnosis; the Mini-International Neuropsychiatric Interview for DSM-5 (19) determined comorbidities. Participants in the AN and BN cohorts met DSM-5 criteria for AN or BN, respectively, during the last year. Participants in the HC cohort did not have any psychiatric disorders. See the Supplement for details about screening, stabilization, weight, acute symptoms, comorbidities, medications, ethnicity, and race.

Scales

As the study examined learning in response to social norm shifts, and learning can depend on cognitive ability, emotional state, and motivations for positive and negative feedback, participants completed assessments about these factors. For cognitive ability, the Wechsler Abbreviated Scale of Intelligence (20) estimated IQ, while the Trail Making Test (21) provided an estimate of set-shifting that is independent of social expectations. The 26-item Eating Disorder Examination Questionnaire (22), the 16-item Quick Inventory of Depressive Symptomatology, clinician rating (23), and the 14-item Structured Interview Guide for the Hamilton Anxiety Rating Scale (24) evaluated disordered eating cognitions and behaviors, depression, and anxiety, respectively. Finally, the 48-item Sensitivity to Punishment and Sensitivity to Reward Questionnaire (25) assessed reactivity to negative and positive motivations.

UG Paradigm

There are 2 players in the norm-adaptation UG, a proposer and a responder (Figure 1). Here, participants played 60 rounds as the responder. The proposer decides how to split a \$20 endowment, and the responder accepts or rejects it. Participants were informed that each round begins with a new proposer, and one round would be randomly picked for payment. If accepted, both players get that split, and if rejected, both get nothing. The participant was told that each offer came from a prior participant, and that payment on each round matters to both players. To simulate adaptation to different social norms, the mean value of the distribution from which the proposer's offers were drawn changed during the game. This approach has been validated to simulate real-world social behaviors (11,12,26). During the first 20 rounds (preconditioning), participants received offers from the medium distribution (mean [SD] = \$8.00 [\$1.50]). During the next 20 rounds (conditioning), participants received offers from either the high (high conditioning; \$12.00 [\$1.50]) or the low (low conditioning; \$4.00 [\$1.50]) distributions. For the last 20 rounds (postconditioning), participants received offers from the medium distribution. In 3 of 5 rounds, participants rated their feelings on a 1-to-9 scale of unhappy to happy emoticons.

Behavioral Data Analysis

Group Difference in Adaptation During the Game. To test adaptation, 2-factor analysis of variance (ANOVA) evaluated how group (AN, BN, HC), conditioning type (low, high), and block (pre [round 1–20], during [round 21–40], post [round 41–60]) impacted rejection rates and emotion ratings using IBM SPSS Statistics Version 21.0 (IBM Corp.), with significance set at $p < .05$ (2-tailed). For significant ANOVA findings, η_p^2 effect size values were calculated, with post hoc

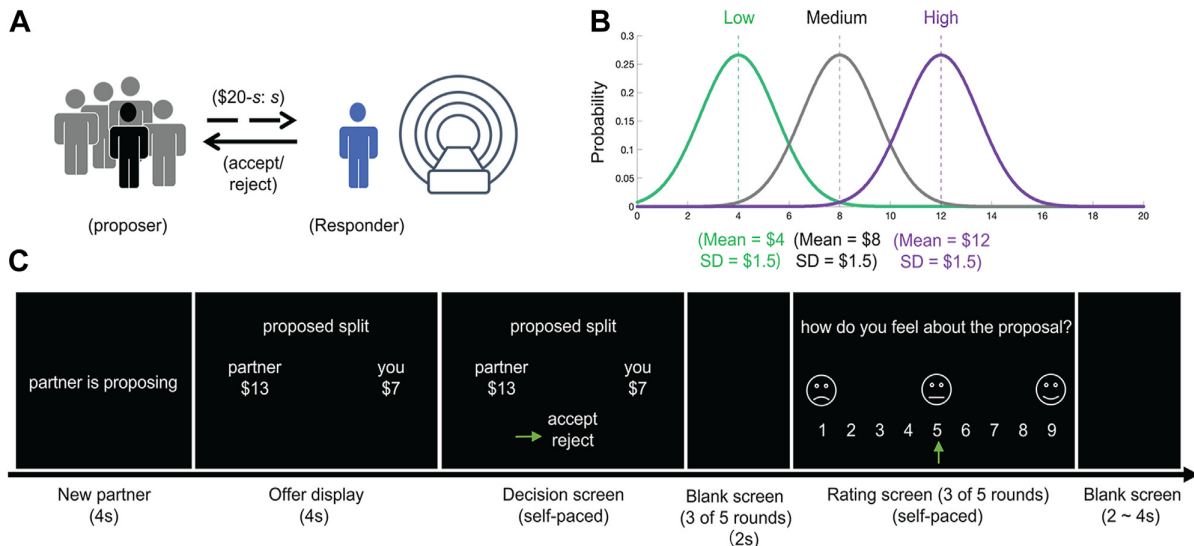


Figure 1. The ultimatum game. (A) All participants played as the responder in the game for 60 rounds in the scanner. The game involved accepting and rejecting offers that are splits of \$20. (B) The monetary split in each round was drawn from 1 of 3 Gaussian distributions (i.e., low, medium, or high). (C) Schematic illustrating the screens shown during 1 round of the game.

comparisons evaluated using 2-tailed pairwise tests with Sidak correction.

Modeling Norm Adaptation. As per prior studies (9,10), participants were assumed to have internal norms updated as offers occurred, and responses were fit to 2 types of norm-adaptation models, Bayesian observer and Rescorla-Wagner models (Supplement). Briefly, the sensitivity to negative NPE (α), sensitivity to positive NPE (β), randomness of choice (γ), learning rate in Rescorla-Wagner models (ϵ), and the starting norm ($\hat{\mu}_0$ or V_0) for variable models were estimated by maximizing the log likelihood of choices across 60 trials for each participant. Models were compared by calculating the Bayesian information criterion score for each model for each participant, with the lowest mean Bayesian information criterion score model winning. NPEs were computed for the winning model as current offer subtracted from the preceding norm ($V_{i-1} - s_i$), with valence indicating whether an offer was better (positive NPE) or worse (negative NPE) than expected. Estimates of sensitivity to better-than-expected (i.e., guilt, β) and worse-than-expected (i.e., envy, α) offers were compared across groups using ANOVA.

Bayesian Logistic Regression. Differential effects of positive and negative NPEs on round response concerning diagnosis and block type were evaluated using Bayesian logistic regression models. Response values x_i were treated as independent binary observations drawn from a Bernoulli distribution with probability π_i , where

$$\text{logit}(\pi_i) = \alpha + \beta^T z_i,$$

for covariate vector z_i associated with round $i = 1, \dots, n$. The covariates of interest were the Rescorla-Wagner NPE (continuous scalar), diagnosis (3 levels: AN, BN, HC), and block (5 levels: first medium block, second block in the low condition, third block in the low condition, second block in the high condition, third block in the high condition).

The full model contained main effects and 2-way interactions of NPE, diagnosis (dx), and block.

$$\begin{aligned} \text{logit}(\pi) &\sim \text{NPE}^- \times dx + \text{NPE}^+ \times dx + \text{NPE}^- \times \text{block} \\ &+ \text{NPE}^+ \times \text{block} + \text{block} \times dx \times z \\ &\sim \text{Bern}(\pi) \end{aligned}$$

We also considered reduced models: 1) diagnosis main effects; 2) main effect and interaction models of diagnosis and block; 3) NPE main effects; 4) NPE \times diagnosis interaction models, with and without block effects. Models were fit via Markov chain Monte Carlo (27) sampling with 4 chains of 2000 iterations each, computed in the R package brms (28). The intercept term and all coefficients were given the prior $N(\mu, \sigma^2)$, with $\mu = 0$, $\sigma^2 = 4$. All parameters sampled obtained convergence as determined by $\hat{R} = 1.0$ (29).

Inference for all quantities of interest was conducted via the Markov chain Monte Carlo samples of the posterior. Estimated coefficient effects were computed as the mean of the posterior samples; corresponding 95% credible intervals were calculated as the posterior quantiles. The posterior

probability of accepting was calculated from the Markov chain Monte Carlo draws of the model coefficients. Covariate effects were significant if the 95% confidence interval of the coefficient did not contain zero; similarly, a difference in the posterior probability of accepting for 2 sets of covariates was significant if the associated posterior 95% confidence interval of the difference did not contain zero.

Rejection Rates and Emotional Responses. Rejection rates and emotion ratings were examined by grouping offers by the sign of NPE and completing a group \times sign of NPE ANOVA. To characterize impact of NPE magnitude, offers were also binned by NPE for each participant and evaluated with a group \times NPE bin ANOVA.

Neuroimaging Data Acquisition and Preprocessing

Images were acquired with an Achieva 3T magnetic resonance imaging scanner (Philips Healthcare), using a single-shot gradient T2*-weighted echo-planar image sequence with a repetition time of 2 seconds. The echo time was 25 ms, and the flip angle was 90°. Volumes were composed of 38 axial slices, each acquired with a matrix size of 64 \times 64 and a voxel size of 3.4 \times 3.4 \times 4 mm³. High-resolution magnetization-prepared rapid acquisition gradient-echo 3-dimensional T1-weighted images were acquired with the following parameters: repetition time = 8100 ms, echo time = 3.7 ms, 12° flip angle, and 1 mm³ voxels. Preprocessing was implemented in SPM12 (<http://www.fil.ion.ucl.ac.uk/spm/software/spm12/>) following standard procedures (Supplement). Participants with excessive (more than 6 mm during task) head movements were excluded (1 AN, 2 BN, and 2 HC).

Neural Region-of-Interest Analyses

We compared brain responses in regions of interest (ROIs) to offers with positive or negative NPEs between groups. The study hypotheses focused on the OFC because prior research identified a parametric relationship between its activations and NPE (12). Comparison ROIs were selected from other UG studies (12,30) (details in Supplement). Five ROIs were examined: OFC (center [4 40 -16]), right anterior insula (32 24 -4), nucleus accumbens (-2 8 -4), left dorsal caudate (-6 18 4), and right dorsal caudate (4 20 6). For each ROI, a spherical mask was created with a 10-mm radius centered at the reported peak activation using MarsBaR (<https://marsbar-toolbox.github.io/index.html>), and the average β values across all voxels within the mask for offers with positive NPEs and negative NPEs were extracted and averaged, respectively. Group differences among AN, BN, and HC cohorts on the averaged β values in each ROI for both positive and negative NPEs were compared using a series of 1-way Kruskal-Wallis tests of k means, a nonparametric approach as activations were leptokurtic, with Bonferroni corrections for multiple comparisons (threshold: corrected $p < .05$). Residuals from simple regressions estimated the impact of acute psychopathology symptoms and medications on any identified differences related to cohorts in the ROI analyses.

Exploratory Whole-Brain Analyses

General linear models were specified for each subject. Visual stimuli and motor responses were modeled in a design matrix constructed by convolving each event onset with a canonical hemodynamic response function in SPM12. Residual effects of head motion were corrected by including 6 estimated motion parameters as covariates. β maps at offer display were estimated for offers with positive NPEs and those with negative NPEs, respectively, at the subject level. Because neuro-behavioral results were similar for AN and BN cohorts, a pooled ED cohort was constructed. ED versus HC differences in responses to positive NPE and negative NPE offers were examined. Associations between clinical symptom severity (Eating Disorder Examination Questionnaire, Structured Interview Guide for the Hamilton Anxiety Rating Scale, Quick Inventory of Depressive Symptomatology, clinician rating) and neural responses were examined by including each assessment as a covariate in the group-level model for positive NPEs and negative NPEs, respectively. The threshold for statistical significance for whole-brain analyses was set at a familywise error cluster-corrected $p < .05$ with an individual voxel-height minimum of $p < .005$. Statistics maps were overlaid on a standard brain in Montreal Neurological Institute space using Mango (<http://ric.uthscsa.edu/mango/mango.html>).

RESULTS

Assessments

Cohorts differed in several measures, but did not differ in age, intelligence (Wechsler Abbreviated Scale of Intelligence), or set-shifting (Trail Making Test) (Table 1). The AN and BN cohorts had more disordered eating, depression, and anxiety symptoms than the HC cohort. Body mass index was lower for the AN cohort relative to both the BN and the HC cohorts, and

the BN cohort had higher levels of weight concern than the AN cohort. Both the AN and the BN cohorts had higher sensitivity to punishment than the HC cohort. The BN cohort had higher sensitivity to reward than both the AN and the HC cohorts. At scanning, acute ED symptoms varied; most participants remained symptomatic (20 AN; 22 BN), with some in partial (5 AN; 7 BN) or full (2 BN) remission (Supplement).

Behavioral Results

Group Difference in Norm-Adaptation Effect.

Adaptation was evaluated by determining how much participants in each group adjusted rejection rates and emotion ratings following conditioning to a new norm. The interaction of group (AN, BN, HC) \times conditioning type (low, high) \times block (pre, during, post) on rejection rates was significant ($F_{4,174} = 2.796$, $p = .028$, $\eta_p^2 = 0.060$) (Figure S1A). HC participants adapted for high offers had a higher rejection rate in the postconditioning block ($53.8 \pm 8.8\%$, third block in the high condition) than the HC participants adapted for low offers ($27.1 \pm 5.4\%$, $p = .011$, third block in the low condition), whereas the postconditioning rejection rate was not different for low and high blocks for either AN ($p = .896$) or BN ($p = .139$) participants. The group \times conditioning type \times block interaction was not significant for emotion ratings ($F_{4,174} = 2.164$, $p = .075$, $\eta_p^2 = 0.047$) (Figure S1A).

Norm-Adaptation Model Estimates. The Rescorla-Wagner model with a fixed starting norm had the best fit in all groups (Table S1) and provided the estimates for sensitivity to positive (i.e., guilt, denoted as β) and negative (i.e., envy, denoted as α) NPEs. Consistent with previous findings (10), HC participants had a significantly higher level of envy, α (mean \pm SE = 3.76 ± 0.56) than guilt, β (0.94 ± 0.39) ($t_{37} = 4.688$, $p < .001$), whereas this difference was not observed for AN

Table 1. Demographic Information and Scales

	AN, Mean (SD)	BN, Mean (SD)	HC, Mean (SD)	AN vs. BN vs. HC		Post Hoc Comparisons
				F (df)	p Value	
Age, Years	27.3 (7.2)	28.6 (6.3)	27.6 (5.0)	0.4 (2,90)	.699	–
BMI	18.8 (1.6)	26.0 (6.1)	25.3 (4.5)	20.6 (2,90)	<.001	AN < HC and BN
WASI	116.6 (10.4)	117.9 (11.4)	116.7 (13.1)	0.1 (2,86)	.905	–
EDE-Q						
Global	2.5 (1.6)	3.0 (1.3)	0.5 (0.5)	43.7 (2,88)	<.001	AN and BN > HC
Restraint	2.3 (1.7)	2.2 (1.6)	0.4 (0.8)	19.3 (2,88)	<.001	AN and BN > HC
Eating concern	2.0 (1.6)	2.0 (1.3)	0.1 (0.2)	33.4 (2,88)	<.001	AN and BN > HC
Shape concern	3.2 (1.7)	3.9 (1.5)	0.9 (0.9)	43.3 (2,88)	<.001	AN and BN > HC
Weight concern	2.7 (1.8)	3.7 (1.6)	0.6 (0.6)	44.1 (2,88)	<.001	BN > AN > HC
Depression (QIDS-CR)	6.0 (5.6)	6.6 (3.6)	1.9 (2)	16.1 (2,89)	<.001	AN and BN > HC
Anxiety (SIGH-A)	9.0 (8.5)	9.7 (5.8)	2.6 (2.9)	15.8 (2,89)	<.001	AN and BN > HC
Reward Sensitivity	9.4 (3.4)	12.6 (3.4)	8.8 (3.8)	8.9 (2,85)	<.001	BN > AN and HC
Punishment Sensitivity	15.9 (5.7)	15.3 (5.1)	9.8 (5.8)	11.6 (2,85)	<.001	AN and BN > HC
TMT-A	21.0 (7.9)	20.5 (8.1)	21.8 (8.7)	0.2 (2,86)	.810	–
TMT-B	53.2 (29.4)	43.1 (16.7)	51.2 (31.8)	1.1 (2,86)	.338	–
TMT-B:A	2.6 (1.1)	2.2 (0.8)	2.4 (1.1)	1.0 (2,86)	.377	–

AN, anorexia nervosa; BMI, body mass index; BN, bulimia nervosa; EDE-Q, Eating Disorder Examination Questionnaire; HC, healthy comparison; QIDS-CR, Quick Inventory of Depressive Symptomatology, clinician rating; SIGH-A, Structured Interview Guide for the Hamilton Anxiety Rating Scale; TMT, Trail Making Test; WASI, Wechsler Abbreviated Scale of Intelligence.

($t_{24} = 0.881, p = .387$) or BN ($t_{29} = 1.550, p = .132$) participants (Table S2).

Bayesian Logistic Regression Results. The main effects models considered how diagnosis, block, and NPE impacted offer acceptances. In the diagnosis main effect model, participants with AN and BN accepted more offers overall than HC participants, with no significant difference between AN and BN participants (Figure 2A; Table S3). The block main effects showed the expected increases and decreases in acceptance probabilities for the high and low blocks, respectively, relative to the first medium block baseline (Figure 2B; Table S3). The third block in the high condition offer block returned to a level slightly lower than baseline, whereas the third block in the low condition block showed an increase. The NPE model showed a strongly significant effect for negative NPE, but not for positive NPE (Figure 2C; Table S3). Although model comparison by Bayes factors and leave-one-out information criterion (31) suggests the best overall fit for the full model (Table S4), the associated plots (Figure S1A) are similar to the NPE \times diagnosis interaction model (Figure S1B), and the block \times NPE interaction terms in the full model changed behavior for positive NPEs, widening associated confidence intervals. As such, the simpler NPE \times diagnosis interaction model provides a more parsimonious understanding of how clinical diagnosis alters processing of social stimuli.

As shown in Figure 2D, the model including all main effects and 2-way interactions of diagnosis with NPEs found significant main effects for positive NPE, negative NPE, and AN

(Table S3). For AN, the difference relative to the HC cohort was most significant for positive NPE, while the effect of BN was distributed with significant interactions for both positive and negative NPE. Both AN and BN cohorts had stronger guilt terms than the HC cohort, with the AN positive NPE coefficient (-0.80 ± 0.13) higher than the BN positive NPE coefficient (-0.29 ± -0.15). For the negative NPE interactions, both AN and BN cohorts exhibited significant envy terms, but with opposite effects: The AN interaction coefficient (-0.15 ± 0.05) reduced the acceptance probability for negative NPEs relative to the HC cohort, whereas the BN interaction coefficient (0.12 ± 0.04) increased the acceptance probability for negative NPEs relative to the HC cohort. These results suggest that, on average, AN and BN cohorts had different responses to both positive (guilt) and negative (envy) NPEs.

Rejection Rates and Emotional Responses to Positive and Negative NPEs. The group \times sign of NPE interaction was significant for both rejection rates ($F_{2,89} = 3.373, p = .039, \eta_p^2 = 0.070$) and emotion ratings ($F_{2,88} = 3.737, p = .028, \eta_p^2 = 0.078$), with the Sidak-corrected post hoc analyses showing participants with AN differed from HC participants for positive NPEs for both rejection rates (AN: $12.7 \pm 4.6\%$; HC: $3.0 \pm 1.2\%$; $p = .051$) and emotion ratings (AN: 6.4 ± 0.3 ; HC: 7.3 ± 0.2 ; $p = .018$). For negative NPEs, no group differences were found for rejection rates ($F_{2,89} = 0.930, p = .399$) or emotional ratings ($F_{2,88} = 0.637, p = .531$). The group \times NPE bin analysis was also significant for both rejection rates ($F_{14,343} = 2.044, p = .014, \eta_p^2 = 0.077$) (Figure 3A) and emotion

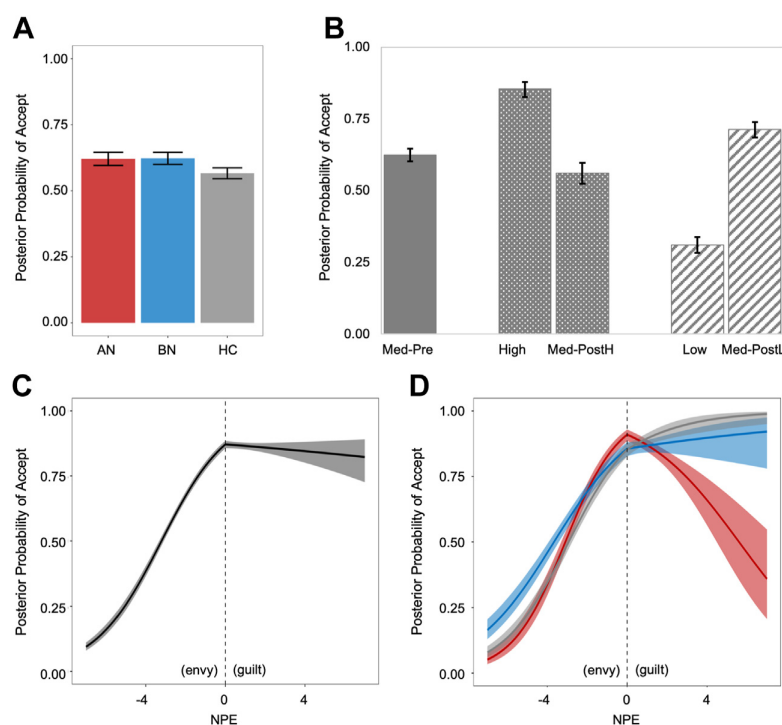


Figure 2. Bayesian logistic regression results. (A) Fitted acceptance probabilities in the diagnosis main-effect model of the 3 cohorts, anorexia nervosa (AN), bulimia nervosa (BN), and healthy comparison (HC). (B) Fitted acceptance probabilities in the block main-effect model. The block types included before-adaptation offers from medium distribution (Med-Pre) (solid bar), high conditioning blocks including high offers (High) followed by medium offers (Med-PostH) (stippled bars), and low conditioning blocks including low offers (Low) followed by medium offers (Med-PostL) (slashed bars). (C) Fitted acceptance probabilities in the norm prediction error (NPE) main-effect model. (D) Fitted acceptance probabilities in the NPE \times diagnosis interaction model.

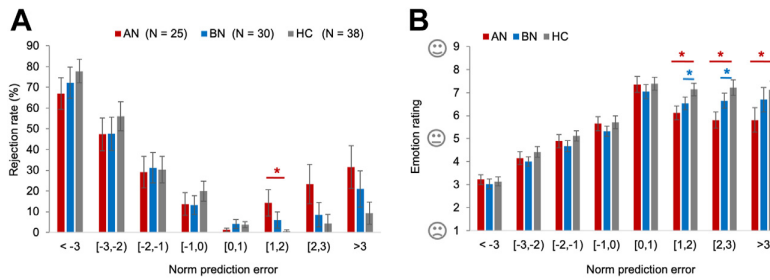


Figure 3. Rejection rate and emotion rating results. **(A)** Rejection rates grouped by norm prediction error bins illustrate increased willingness to accept better-than-expected offers by the healthy comparison (HC) than the anorexia nervosa (AN) and bulimia nervosa (BN) cohorts ($F_{14,343} = 2.044, p = .014, \eta_p^2 = 0.077$). **(B)** Emotion ratings grouped by norm prediction error bins also show that the HC cohort reported more positive emotion ratings after receiving better-than-expected offers than the AN and BN cohorts ($F_{14,266} = 2.365, p = .004, \eta_p^2 = 0.111$). *Sidak-corrected $p < .05$. A \sqsubseteq NPE < -2 .

bracket indicates that the end point is included in the interval, while a parenthesis means it is not. For example, $[-3, -2)$ indicates $-3 \leq \text{NPE} < -2$.

ratings ($F_{14,266} = 2.365, p = .004, \eta_p^2 = 0.111$) (Figure 3B), with both AN and BN cohorts differing from HC cohorts (Table S5).

Brain Results

ROI Analysis. OFC responses to positive NPE offers were significantly different across the 3 groups (Figure 4A) ($H_2 = 8.516, p = .014$), with Bonferroni post hoc pairwise comparisons showing differences between AN and HC (AN: -0.62 ± 0.25 ; HC: 0.38 ± 0.32 ; $p = .018, d = 0.655$) cohorts and BN and HC (BN: -0.43 ± 0.23 ; $p = .043, d = 0.527$) cohorts. No group differences were observed for OFC responses to negative NPE offers ($H_2 = 1.656, p = .437$). No group differences in activations were observed in other ROIs (anterior insula, nucleus accumbens, left and right dorsal caudate) for either positive NPE or negative NPE offers (Table S6). OFC differences were not related to psychiatric symptoms or psychoactive medications (Supplement).

Whole-Brain Analysis. The pooled ED group (AN and BN) had decreased activations in the ventromedial prefrontal cortex (vmPFC)/anterior cingulate cortex and the dorsomedial PFC (dmPFC) relative to the HC cohort for positive NPE offers (Figure 4B; Figure S4; Table S7). No whole-brain clusters were observed for negative NPE offers.

Correlation Between Brain Responses to NPEs and Clinical Symptoms.

Only the Eating Disorder Examination Questionnaire regression for positive NPE identified a significant cluster in the right dmPFC (peak = $[18, 50, 18], t = 4.09, k = 722$) (Figure 4C). As detailed in Figure 4D, participants with more severe ED symptoms showed less positive activation in the dmPFC for positive NPE offers, with similar relationships observed for both AN ($r = -0.49, p = .015$) and BN ($r = -0.57, p = .002$) cohorts. No significant clusters were identified in the depression or anxiety regressions.

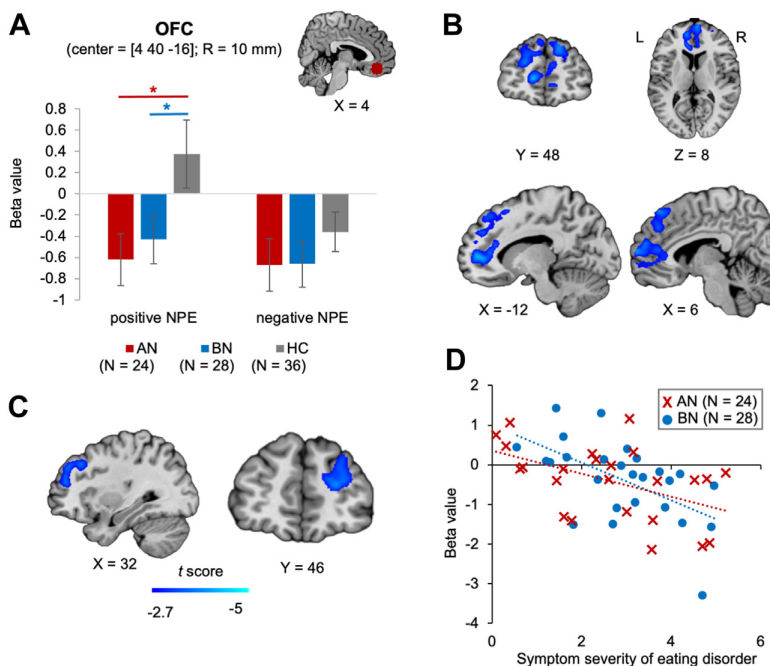


Figure 4. Brain results. **(A)** The orbitofrontal cortex (OFC) was differentially activated for positive norm prediction error (NPE) for the healthy comparison (HC) cohort, but not for the anorexia nervosa (AN) or bulimia nervosa (BN) cohorts; no differences were seen for the 3 groups for negative NPE. **(B)** Whole-brain comparison showed attenuated activations in medial prefrontal cortex in response to offers with positive NPE in AN and BN relative to HC cohorts. **(C)** Whole-brain regression revealed that the AN and BN cohort responses to offers with positive NPE in the dorsomedial prefrontal cortex were negatively correlated with current symptom severity of the eating disorder. **(D)** Activation in the dorsomedial prefrontal cortex cluster shown in panel (C) was significantly correlated with symptom severity for both AN ($r = -0.49, p = .015$) and BN ($r = -0.57, p = .002$) cohorts. For panels (B, C), individual voxel height $p < .005$ followed by familywise error cluster-corrected $p < .05$. *Sidak-corrected $p < .05$.

DISCUSSION

Responding appropriately to unpredictable stimuli in one's social environment requires identification of whether such a stimulus is advantageous or disadvantageous to the individual (32). Advantageous offers led to diminished neural, behavioral, and emotional responses in both AN and BN cohorts, while no differences were found for disadvantageous offers. Participants with AN and BN did not adapt when the norms in the environment changed. Both participants with AN and participants with BN showed less activation for advantageous offers in the OFC, vmPFC/anterior cingulate cortex, and dmPFC relative to HC. ED severity was associated with reduced neural activations to advantageous offers in the right dmPFC.

NPEs are socially ambiguous and more complex than reward prediction errors because social rewards can include a personal obligation, while rewards from a game of chance do not (33). In a previous study with the UG, women with AN playing the proposer role offered higher splits than comparison women, interpreted as higher guilt in women with AN (34). Here, guilt also reduced acceptances for participants with AN and participants with BN in the responder role, while the HC cohort changed acceptance rates more in response to envy than guilt, consistent with prior studies (12,26). Other studies have shown that ambiguous social signals are interpreted more negatively in participants with EDs than comparison participants (35,36).

Social context appears to be important to identify neurobehavioral differences in EDs. No differences were observed for the Wechsler Abbreviated Scale of Intelligence or Trail Making Test, measures of intelligence and set-shifting that are independent of social context. Second, neural differences were observed only in frontal brain regions, not subcortical regions as in depression (30). Decreased OFC activation in EDs might indicate alterations in reward processing and learning (37) and in creating models connecting oneself to one's environment (38). The OFC overlaps with the vmPFC, an area involved in emotion processing, social cognition, and decision making that is dysregulated in many types of psychopathology (39). Evaluation of social context is a key role attributed to the dmPFC (40); we observed that the severity of the ED reduced activations in the dmPFC to positive NPE. Changes in these frontal brain regions may contribute to differences in social cognition previously reported for individuals with acute AN relative to individuals with recovered AN and BN (9,41,42).

Social stressors can lead to the manifestation of psychiatric disorders when biological liabilities are present. EDs often manifest after experiences of social exclusion (6,43,44). The reduced frontal brain activations observed for positive NPEs provide a candidate neural mechanism that may contribute to EDs in multiple ways. The lower emotional ratings for positive NPE offers suggest that advantageous interactions are less satisfying. People with EDs tend to have smaller social networks (45) and spend less time in social activities (46). We hypothesize that differences in response to positive NPEs may impair development of supportive relationships. Building relationships is complex, but a first step involves awareness of advantageous social signals. In the real world, these data suggest that people with EDs may not notice when someone values them. Future studies should evaluate how social brain

responses are related to real-world social networks. EDs are also closely associated with interpersonal trauma (47). Understanding how exposure to trauma alters social brain responses and contributes to disordered eating requires more research.

Increased rejection of advantageous offers in the UG was also reported for adults who had attended a rule-based intervention in preschool (26). Advantageous but inequitable interactions may be less satisfying for people who prize social rules. Overvaluation of shape and weight is a core feature of both AN and BN (48). The identity-value model of social decision making proposes that self-regulation depends on weighting information about concepts internalized as high-value components of one's identity (e.g., appearance, rules) with immediate choices (e.g., eat, share), and proposed integration occurs in the vmPFC (49,50). We hypothesize that one neurodevelopmental risk for EDs may involve prioritizing social rules, such as equity, over self-interest by the vmPFC. When such individuals experience inequity, restrictive or compensatory eating behaviors may begin in an attempt to conform to societal rules about appearance such as the thin ideal.

Responses to fair and unfair splits in the UG, but not NPEs, have also been evaluated in anxiety and depression. Relative to comparison, individuals with anxiety disorders reject fewer unfair splits (51), but no differences were reported for depression (30), for a combined cohort of depression and/or social anxiety (52), or in our sample (Supplement). The depression cohort also reported more sadness following unfair offers and showed lower neural activations for fair offers in both the nucleus accumbens and the dorsal caudate, but no differences in the vmPFC (30). In sum, the UG differences observed in EDs appear to be distinct from those reported for depression and anxiety.

Limitations

We studied stabilized outpatients with AN and BN. This choice may reduce detection of problems associated with severe ED pathology but minimizes the impact of medical instability and low body mass index on the brain (53). Both symptomatic and remitted participants were included, but sample size was insufficient to separate these groups in analyses. The study included only women between 18 and 46 years, to minimize confounds, as both sex and age alter social brain responses (54,55). Studies of social processing in men with EDs are needed. Participants had psychiatric comorbidities and took medications, as is common in individuals with EDs (56). Causality about differences observed cannot be determined from a single assessment. The stability of the brain-symptom associations should be validated in larger samples.

Conclusions

AN and BN showed impaired processing of positive social learning signals in frontal brain regions. From a treatment perspective, psychotherapy requires learning, and while psychotherapy is the first-line treatment for AN and BN, it is effective for less than half of individuals with AN (57) and BN (58). A number of interventions being developed for EDs may also affect social learning, including interventions that increase cognitive flexibility (59,60), change social behaviors (61–64), or

target frontal brain regions (65–67). Evaluation of social processing should be included in such research in EDs so that we can better understand whether changes occur in the social domain and impact clinical course.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the National Institute of Mental Health (Grant Nos. R01 MH112927 and R21 MH124016 [to CJM], R01 MH124115 [to MV], and R01 MH122948 [to DP]), Ware Family Clinical Research Award (to CJM), Research Project of Shanghai Science and Technology Commission (Grant No. 20dz2260300 [to YL]), Fundamental Research Funds for the Central Universities (to YL), National Natural Science Foundation of China (Grant No. 32200886 [to YL]), Young Elite Scientists Sponsorship Program by China Association for Science and Technology (Grant No. 2022QNRC001 [to YL]), and Wellcome Trust Principal Research Fellowship (to PRM). Data collection with RedCap was supported by the National Institutes of Health Clinical and Translational Science Award (Grant No. UL1 RR024982). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

We thank all the research participants for choosing to participate in this project.

Data are available in the National Institute of Mental Health Data Archive (dataset identifiers: C2752; 10.15154/nkvz-zm72).

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Shanghai Key Laboratory of Mental Health and Psychological Crisis Intervention, School of Psychology and Cognitive Science, East China Normal University, Shanghai, China (YL); Fraikin Biomedical Research Institute, Virginia Tech, Roanoke, Virginia (YL, TL, PRM); Department of Statistics, Rice University, Houston, Texas (DP, MV); Department of Psychiatry, University of Texas at Southwestern Medical School, Dallas, Texas (BBB, JMP, JM, CJM); Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York (XG); Nash Family Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, New York (XG); Center for Computational Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York (XG); Department of Physics, Virginia Tech, Blacksburg, Virginia (PRM); and Virginia Tech-Wake Forest School of Biomedical Engineering and Mechanics, Blacksburg, Virginia (PRM).

Address correspondence to Carrie J. McAdams, M.D., Ph.D., at Carrie.McAdams@UTSouthwestern.edu.

Received Jul 5, 2023; revised Sep 22, 2023; accepted Oct 23, 2023.

Supplementary material cited in this article is available online at <https://doi.org/10.1016/j.bpsc.2023.10.010>.

REFERENCES

- Schaumberg K, Welch E, Breithaupt L, Hubel C, Baker JH, Munn-Chernoff MA, *et al.* (2017): The science behind the Academy for Eating Disorders' Nine Truths About Eating Disorders. *Eur Eat Disord Rev* 25:432–450.
- Suarez-Albor CL, Galletta M, Gomez-Bustamante EM (2022): Factors associated with eating disorders in adolescents: A systematic review. *Acta Biomed* 93:e2022253.
- Kastner D, Lowe B, Gumz A (2019): The role of self-esteem in the treatment of patients with anorexia nervosa—a systematic review and meta-analysis. *Int J Eat Disord* 52:101–116.
- Kerr-Gaffney J, Harrison A, Tchanturia K (2018): Social anxiety in the eating disorders: A systematic review and meta-analysis. *Psychol Med* 48:2477–2491.
- Gerbase ME, Richards LK, Thomas JJ, Agnew-Blais JC, Thompson-Brenner H, Gilman SE, *et al.* (2014): Globalization and eating disorder risk: Peer influence, perceived social norms, and adolescent disordered eating in Fiji. *Int J Eat Disord* 47:727–737.
- Rojo L, Conesa L, Bermudez O, Livianos L (2006): Influence of stress in the onset of eating disorders: Data from a two-stage epidemiologic controlled study. *Psychosom Med* 68:628–635.
- Fonagy P, Campbell C, Constantinou M, Higgitt A, Allison E, Luyten P (2021): Culture and psychopathology: An attempt at reconsidering the role of social learning. *Dev Psychopathol* 34:1205–1220.
- Konstantakopoulos G, Ioannidi N, Patrikelis P, Gonidakis F (2020): Empathy, mentalizing, and cognitive functioning in anorexia nervosa and bulimia nervosa. *Soc Neurosci* 15:477–488.
- Mason TB, Lesser EL, Dolgon-Krutolow AR, Wonderlich SA, Smith KE (2021): An updated transdiagnostic review of social cognition and eating disorder psychopathology. *J Psychiatr Res* 143:602–627.
- van Dijk E, De Dreu CK (2021): Experimental games and social decision making. *Annu Rev Psychol* 72:415–438.
- Gu X, Wang X, Hula A, Wang S, Xu S, Lohrenz TM, *et al.* (2015): Necessary, yet dissociable contributions of the insular and ventromedial prefrontal cortices to norm adaptation: Computational and lesion evidence in humans. *J Neurosci* 35:467–473.
- Xiang T, Lohrenz T, Montague PR (2013): Computational substrates of norms and their violations during social exchange. *J Neurosci* 33:1099–1108.
- Apps MAJ, Sallet J (2017): Social learning in the medial prefrontal cortex. *Trends Cogn Sci* 21:151–152.
- McAdams CJ, Krawczyk DC (2013): Neural responses during social and self-knowledge tasks in bulimia nervosa. *Front Psychiatry* 4:103.
- Mendoza CJ, Palka JM, Pelfrey SE, Hunt BJ, Krawczyk DC, McAdams CJ (2022): Neural processes related to negative self-concept in adult and adolescent anorexia nervosa. *Eur Eat Disord Rev* 30:23–35.
- McAdams CJ, Lohrenz T, Montague PR (2015): Neural responses to kindness and malevolence differ in illness and recovery in women with anorexia nervosa. *Hum Brain Mapp* 36:5207–5219.
- Luo Y, Mendoza C, Pelfrey S, Lohrenz T, Gu X, Montague PR, *et al.* (2022): Elevated neurobehavioral responses to negative social interactions in women with bulimia nervosa. *Biol Psychiatry Cogn Neurosci Neuroimaging* 7:696–705.
- Attia E, Marcus MD, Walsh BT, Guarda AS (2017): The need for consistent outcome measures in eating disorder treatment programs: A proposal for the field. *J Eat Disord* 50:231–234.
- Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, *et al.* (1998): The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 59(Suppl 20):22–33;quiz 34–57.
- Wechsler D (1999): Wechsler Abbreviated Scale of Intelligence (WASI). San Antonio, TX: The Psychological Corporation.
- Tombaugh TN (2004): Trail Making Test A and B: Normative data stratified by age and education. *Arch Clin Neuropsychol* 19:203–214.
- Mond JM, Hay PJ, Rodgers B, Owen C, Beumont PJ (2004): Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behav Res Ther* 42:551–567.
- Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, *et al.* (2003): The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biol Psychiatry* 54:573–583.
- Shear MK, Vander Bilt J, Rucci P, Endicott J, Lydiard B, Otto MW, *et al.* (2001): Reliability and validity of a structured interview guide for the Hamilton Anxiety Rating Scale (SIGH-A). *Depress Anxiety* 13:166–178.
- Torrubia R, Ávila C, Moltó J, Caseras X (2001): The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Pers Individ Dif* 31:837–862.
- Luo Y, Hetu S, Lohrenz T, Hula A, Dayan P, Ramey SL, *et al.* (2018): Early childhood investment impacts social decision-making four decades later. *Nat Commun* 9:4705.

Impaired Social Learning in Eating Disorders

27. Gilks WR, Richardson S, Spiegelhalter D (1995): Markov Chain Monte Carlo in Practice. Boca Raton, FL: CRC Press.
28. Bürkner P (2017): An R package for bayesian multilevel models using Stan. *J Stat Softw* 80:1–28.
29. Brooks SP, Gelman A (1998): General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 7:434–455.
30. Gradin VB, Perez A, MacFarlane JA, Cavin I, Waiter G, Engelmann J, et al. (2015): Abnormal brain responses to social fairness in depression: An fMRI study using the Ultimatum Game. *Psychol Med* 45:1241–1251.
31. Vehtari A, Gelman A, Gabry J (2017): Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput* 27:1413–1432.
32. Joiner J, Piva M, Turrin C, Chang SWC (2017): Social learning through prediction error in the brain. *NPJ Sci Learn* 2:8.
33. Ruff CC, Fehr E (2014): The neurobiology of rewards and values in social decision making. *Nat Rev Neurosci* 15:549–562.
34. Isobe M, Kawabata M, Murao E, Noda T, Matsukawa N, Kawada R, et al. (2018): Exaggerated envy and guilt measured by economic games in Japanese women with anorexia nervosa. *Biopsychosoc Med* 12:19.
35. Cardi V, Turtur R, Schifano S, Leppanen J, Hirsch CR, Treasure J (2017): Biased interpretation of ambiguous social scenarios in anorexia nervosa. *Eur Eat Disord Rev* 25:60–64.
36. Bronstein MV, Everaert J, Kummerfeld E, Haynos AF, Vinogradov S (2022): Biased and inflexible interpretations of ambiguous social situations: Associations with eating disorder symptoms and socioemotional functioning. *Int J Eat Disord* 55:518–529.
37. O’Doherty JP (2004): Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Curr Opin Neurobiol* 14:769–776.
38. Buckner RL, DiNicola LM (2019): The brain’s default network: Updated anatomy, physiology and evolving insights. *Nat Rev Neurosci* 20:593–608.
39. Hiser J, Koenigs M (2018): The multifaceted role of the ventromedial prefrontal cortex in emotion, decision making, social cognition, and psychopathology. *Biol Psychiatry* 83:638–647.
40. Lieberman MD, Straccia MA, Meyer ML, Du M, Tan KM (2019): Social, self, (situational), and affective processes in medial prefrontal cortex (MPFC): Causal, multivariate, and reverse inference evidence. *Neurosci Biobehav Rev* 99:311–328.
41. Bora E, Kose S (2016): Meta-analysis of theory of mind in anorexia nervosa and bulimia nervosa: A specific impairment of cognitive perspective taking in anorexia nervosa? *Int J Eat Disord* 49:739–740.
42. Preti A, Siddi S, Marzola E, Abbate Daga G (2022): Affective cognition in eating disorders: A systematic review and meta-analysis of the performance on the “Reading the Mind in the Eyes” test. *Eat Weight Disord* 27:2291–2307.
43. Grilo CM, Pagano ME, Stout RL, Markowitz JC, Ansell EB, Pinto A, et al. (2012): Stressful life events predict eating disorder relapse following remission: Six-year prospective outcomes. *Int J Eat Disord* 45:185–192.
44. Lie SO, Ro O, Bang L (2019): Is bullying and teasing associated with eating disorders? A systematic review and meta-analysis. *Int J Eat Disord* 52:497–514.
45. Tiller JM, Sloane G, Schmidt U, Troop N, Power M, Treasure JL (1997): Social support in patients with anorexia nervosa and bulimia nervosa. *Int J Eat Disord* 21:31–38.
46. Krug I, Penelo E, Fernandez-Aranda F, Anderlueh M, Bellodi L, Cellini E, et al. (2013): Low social interactions in eating disorder patients in childhood and adulthood: A multi-centre European case control study. *J Health Psychol* 18:26–37.
47. Mitchell KS, Sciofi ER, Galovski T, Belfer PL, Cooper Z (2021): Post-traumatic stress disorder and eating disorders: Maintaining mechanisms and treatment targets. *Eat Disord* 29:292–306.
48. Forrest LN, Jones PJ, Ortiz SN, Smith AR (2018): Core psychopathology in anorexia nervosa and bulimia nervosa: A network analysis. *Int J Eat Disord* 51:668–679.
49. Berkman ET, Livingston JL, Kahn LE (2017): The identity-value model of self-regulation: Integration, extension, and open questions. *Psychol Inq* 28:157–164.
50. Berkman ET, Livingston JL, Kahn LE (2017): Finding the “self” in self-regulation: The identity-value model. *Psychol Inq* 28:77–98.
51. Grecucci A, Giorgetta C, Brambilla P, Zuanon S, Perini L, Balestrieri M, et al. (2013): Anxious ultimatums: How anxiety disorders affect socioeconomic behaviour. *Cogn Emot* 27:230–244.
52. Nicolaisen-Sobesky E, Paz V, Cervantes-Constantino F, Fernandez-Theoduloz G, Perez A, Martinez-Montes E, et al. (2023): Event-related potentials during the ultimatum game in people with symptoms of depression and/or social anxiety. *Psychophysiology* 60:e14319.
53. Walton E, Bernardoni F, Batury V-L, Bahnsen K, Larivière S, Abbate-Daga G, et al. (2022): Brain structure in acutely underweight and partially weight-restored individuals with anorexia nervosa—a coordinated analysis by the ENIGMA Eating Disorders Working Group. *Biol Psychiatry* 92:730–738.
54. Pfeifer JH, Berkman ET (2018): The development of self and identity in adolescence: Neural evidence and implications for a value-based choice perspective on motivated behavior. *Child Dev Perspect* 12:158–164.
55. Proverbio AM (2023): Sex differences in the social brain and in social cognition. *J Neurosci Res* 101:730–738.
56. Van Alsten SC, Duncan AE (2020): Lifetime patterns of comorbidity in eating disorders: An approach using sequence analysis. *Eur Eat Disord Rev* 28:709–723.
57. Solmi M, Wade TD, Byrne S, Del Giovane C, Fairburn CG, Ostinelli EG, et al. (2021): Comparative efficacy and acceptability of psychological interventions for the treatment of adult outpatients with anorexia nervosa: A systematic review and network meta-analysis. *Lancet Psychiatry* 8:215–224.
58. Svaldi J, Schmitz F, Baur J, Hartmann AS, Legenbauer T, Thaler C, et al. (2019): Efficacy of psychotherapies and pharmacotherapies for bulimia nervosa. *Psychol Med* 49:898–910.
59. Timko CA, Bhattacharya A, Fitzpatrick KK, Howe H, Rodriguez D, Mears C, et al. (2021): The shifting perspectives study protocol: Cognitive remediation therapy as an adjunctive treatment to family based treatment for adolescents with anorexia nervosa. *Contemp Clin Trials* 103:106313.
60. Meneguzzo P, Tenconi E, Todisco P, Favaro A (2021): Cognitive remediation therapy for anorexia nervosa as a rolling group intervention: Data from a longitudinal study in an eating disorders specialized inpatient unit. *Eur Eat Disord Rev* 29:770–782.
61. Cardi V, Tchanturia K, Treasure J (2018): Premorbid and illness-related social difficulties in eating disorders: An overview of the literature and treatment developments. *Curr Neuropharmacol* 16:1122–1130.
62. Harrison A, Watterson SV, Bennett SD (2019): An experimental investigation into the use of eye-contact in social interactions in women in the acute and recovered stages of anorexia nervosa. *Int J Eat Disord* 52:61–70.
63. Hunt BJ, Smith Hagan W, Pelfrey S, Mericle S, Harper JA, Palka JM, et al. (2021): Pilot data from the Self-Blame and Perspective-Taking Intervention for eating disorders. *J Behav Cogn Ther* 31:57–66.
64. Hagan WS, Mericle S, Hunt BJ, Harper JA, Palka JM, Pelfrey S, et al. (2021): Qualitative patient experiences from the Self-Blame and Perspective-Taking Intervention for eating disorders. *J Eat Disord* 9:127.
65. McClelland J, Kekic M, Bozhilova N, Nestler S, Dew T, Van den Eynde F, et al. (2016): A randomised controlled trial of neuronavigated repetitive transcranial magnetic stimulation (rTMS) in anorexia nervosa. *PLoS One* 11:e0148606.
66. Gay A, Jaussent I, Sigaud T, Billard S, Attal J, Seneque M, et al. (2016): A lack of clinical effect of high-frequency rTMS to dorsolateral prefrontal cortex on bulimic symptoms: A randomised, double-blind trial. *Eur Eat Disord Rev* 24:474–481.
67. Dalton B, Bartholdy S, McClelland J, Kekic M, Rennalls SJ, Werthmann J, et al. (2018): Randomised controlled feasibility trial of real versus sham repetitive transcranial magnetic stimulation treatment in adults with severe and enduring anorexia nervosa: The TIARA study. *BMJ Open* 8:e021531.