**WILEY**

**International Journal of**
**EATING DISORDERS**

ORIGINAL ARTICLE

# Mild, moderate, meaningful? Examining the psychological and functioning correlates of DSM-5 eating disorder severity specifiers

Loren Gianini, PhD[1] (ID) | Christina A. Roberto, PhD[2] (ID) | Evelyn Attia, MD[1] | B. Timothy Walsh, MD[1] | Jennifer J. Thomas, PhD[3,4] (ID) | Kamryn T. Eddy, PhD[3,4] | Carlos M. Grilo, PhD[5,6] | Thomas Weigel, MD[7] | Robyn Sysko, PhD[8]

[1]Department of Psychiatry, Columbia University Medical Center, New York State Psychiatric Institute, 1051 Riverside Drive, Unit 98, New York, New York

[2]Department of Medical Ethics & Health Policy, Perelman School of Medicine, University of Pennsylvania, Blockley Hall, Philadelphia, Pennsylvania

[3]Eating Disorders Clinical and Research Program, Massachusetts General Hospital, 2 Longfellow Place, Suite 200, Boston, MA

[4]Department of Psychiatry, Harvard Medical School, 25 Shattuck Street, Boston, Massachusetts

[5]Department of Psychiatry, Yale University School of Medicine, 301 Cedar Street, New Haven, Connecticut

[6]Department of Psychology, Yale University, Box 208205, New Haven, Connecticut

[7]Klarman Center, McLean Hospital and Harvard Medical School, Belmont, Massachusetts

[8]Eating and Weight Disorders Program, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1230, New York, New York

**Correspondence**
Loren Gianini, Columbia Center for Eating Disorders, New York State Psychiatric Institute, 1051 Riverside Drive, Unit 98, New York, NY 10032.
Email: lg2753@cumc.columbia.edu

**Funding information**
The National Institutes of Health, Grants Numbers: K24 DK070052 (Dr. Grilo), R01 DK49587 (Dr. Grilo), T32-MH096679 (Dr. Attia and Dr. Gianini), DK088532 (Dr. Sysko), P30AG034546 (Dr. Roberto) and the Hilda and Preston Davis Foundation (Dr. Thomas and Dr. Eddy).

## Abstract

**Objective:** This study evaluated the *DSM-5* severity specifiers for treatment-seeking groups of participants with anorexia nervosa (AN), the purging form of bulimia nervosa (BN), and binge eating disorder (BED).

**Method:** Hundred and sixty-two participants with AN, 93 participants with BN, and 343 participants with BED were diagnosed using semi-structured interviews, sub-categorized using *DSM-5* severity specifiers and compared on demographic and cross-sectional clinical measures.

**Results:** In AN, the number of previous hospitalizations and the duration of illness increased with severity, but there was no difference across severity groups on measures of eating pathology, depression, or measures of self-reported physical or emotional functioning. In BN, the level of eating concerns increased across the severity groups, but the groups did not differ on measures of depression, self-esteem, and most eating pathology variables. In BN, support was also found for an alternative severity classification scheme based upon number of methods of purging. In BED, levels of several measures of eating pathology and self-reported physical and emotional functioning increased across the severity groups. For BED, however, support was also found for an alternative severity classification scheme based upon overvaluation of shape and weight. Preliminary evidence was also found for a transdiagnostic severity index based upon overvaluation of shape and weight.

**Discussion:** Overall, these data show limited support for the *DSM-5* severity specifiers for BN and modest support for the *DSM-5* severity specifiers for AN and BED.

**KEYWORDS**
anorexia nervosa, bulimia nervosa, binge-eating disorder, diagnosis, DSM-5, severity

# 1 | INTRODUCTION

The recently released fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (*DSM-5*) (American Psychiatric Association, 2013) sparked interest in evaluating new aspects of the feeding and eating disorder classification. An addition to the *DSM-5* was the inclusion of specifiers to characterize the severity of the eating disorders using a dimensional measure. For anorexia nervosa (AN), the severity specifiers are based on body mass index (BMI; kg/m$^2$); for bulimia nervosa (BN), on frequency of inappropriate compensatory behaviors (e.g., self-induced vomiting); for binge-eating disorder (BED), on frequency of binge eating episodes. The Feeding Disorders, Other Specified Feeding or Eating Disorder (OSFED) and Unspecified Feeding or Eating Disorder (UFED) do not have severity specifiers. *DSM 5* does not clearly articulate what these severity specifiers should predict, but it does note that when assessing severity one may consider the intensity of the symptoms of the illness, degree of functional impairment, and the need for medical or clinical supervision (American Psychiatric A, Force DSMT, 2013). In addition to indicating need for supervision/treatment and functional impairment, a severity indicator ideally would also provide information regarding risk of complications, prognosis, and intensity of other contemporaneous symptoms.

There is limited empirical research on whether the *DSM-5* specifiers provide clinically meaningful information. In a treatment-seeking sample of individuals with AN, Machado and colleagues (Machado, Grilo, and Crosby, 2017) found that the severity groups did not differ on cross-sectional measures of eating pathology or frequency of binge eating or purging. In a non-clinical community sample, Grilo and colleagues (Grilo, Ivezaj, & White, 2015a) found modest differences in eating disorder pathology and depression between the extreme severity group for BN and the other severity groups, but effect sizes were small, offering modest evidence that the severity specifiers detect differences on cross-sectional measures of pathology. In a treatment-seeking sample receiving cognitive behavioral therapy (CBT), the severity groups differed significantly on abstinence from compensatory behaviors at the conclusion of a course of treatment, with abstinence being less likely with increase in severity (Dakanalis et al., 2016). Although the *DSM-5* severity classification system has not been directly compared to alternative classification systems proposed for BN, research has found that individuals who report multiple methods of purging versus only one method exhibit greater eating disorder pathology, anxiety, self-injurious behavior, and suicide attempts (Eddy et al., 2009; Edler, Haedt, & Keel, 2007; Favaro & Santonastaso, 1996).

Grilo and colleagues also examined *DSM-5* severity specifiers in a non-clinical community sample of individuals with BED and found that those in the moderate severity category reported higher levels of eating disorder pathology compared to those in the mild category, with small effect sizes (Grilo, Ivezaj, & White, 2015c). These groups did not differ significantly on severity of depression. However, using an alternative classification scheme, they found that individuals reporting overvaluation of shape and weight had significantly greater levels of eating disorder pathology and depression than individuals without overvalua-

tion of shape and weight (Grilo et al., 2015c). Effect sizes for these analyses were considerably larger than those reported when comparing the *DSM-5* severity specifier groups, and the authors concluded that overvaluation of shape and weight may provide a more clinically meaningful measure of severity (Grilo et al., 2015c). In a clinical sample with BED, Grilo and colleagues found that those in the severe/extreme group reported a higher level of eating disorder pathology compared to the mild and moderate groups, and the severe/extreme and moderate groups reported higher levels of depression than those in the mild Group (Grilo et al., 2015b). The BED groups categorized based on overvaluation of shape and weight, however, exhibited more pronounced statistically significant differences reflecting moderate-to-large effect sizes on these measures (Grilo et al., 2015b). Dakanalis and colleagues (Dakanalis, Colmegna, Riva, & Clerici, 2017) found that the *DSM 5* severity groups differed significantly on abstinence from binge eating at the conclusion of a course of treatment with CBT, with abstinence being less likely with increase in severity. Though not a direct examination of the *DSM* severity specifiers, Wilson and Sysko (2009) synthesized findings comparing groups of individuals with BED based upon frequency of binge episodes. Their review found that individuals with BED who reported binge eating once per week did not differ significantly from individuals with BED who reported binge eating twice per week or more on measures of eating pathology, self-esteem, or depression. Taken together, these studies suggest that the *DSM-5* severity specifiers may have some utility in discriminating groups based upon cross-sectional measures of psychopathology and treatment outcome.

It is useful to examine meaningful markers of severity within the diagnostic groups of AN, BN, and BED. There has also been enthusiasm within the field for considering transdiagnostic approaches to classification and treatment of the eating disorders. Restrictive eating, binge eating, compensatory behaviors, and overvaluation of shape and weight are just some of the cognitive and behavioral features which may be present in varying forms across AN, BN, and BED, and crossover between these diagnoses is fairly common (Castellini et al., 2011). Therefore, examining a potential severity classification scheme that can be implemented transdiagnostically across AN, BN, and BED is also indicated. Some have suggested that the "core pathology" across many of the eating disorders is the overvaluation of shape and weight and this construct is commonly assessed during the course of diagnosis of AN, BN, and BED (Fairburn, Cooper, & Shafran, 2003). Given the commonalities between AN, BN, and BED, the utility of a transdiagnostic severity classification scheme, the importance of overvaluation of shape and weight across AN, BN, and BED and the promising results in examining this as an alternative severity classification scheme in BED (Grilo et al., 2015a, 2015b), we believe further exploration of overvaluation of shape and weight as a transdiagnostic alternative severity classification scheme is warranted.

The aim of the present study was to examine measures of eating disorder pathology and impairment, depression, self-esteem, and health status across the newly developed *DSM* severity specifiers among treatment-seeking participants with AN, the purging form of BN, or BED. As noted above, one purpose of the *DSM 5* severity specifiers is

to discriminate groups based upon intensity of eating disorder psychopathology, as well as functional impairment. We hypothesized that severity of pathology and impairment would vary in a monotonically increasing manner across the severity categories. Additional aims were to test other existing severity classifications, specifically single vs. multiple purging methods as an alternative severity specifier scheme for BN and overvaluation of shape and weight as an alternative severity specifier scheme for BED. We hypothesized that those with BN who engage in multiple purging methods versus a single purging method and those with BED who overvalue shape and weight compared to those without overvaluation of shape and weight would report greater levels of pathology and impairment. Finally, we sought to examine overvaluation of shape and weight as a potential transdiagnostic severity classification scheme across the three diagnoses, predicting that those who overvalue shape and weight would report greater levels of pathology and impairment compared to those without overvaluation of shape and weight.

## 2 | METHODS

### 2.1 | Participants

#### 2.1.1 | Anorexia nervosa

Participants were inpatients with AN admitted to a structured behavioral inpatient treatment program designed to normalize eating and weight. Upon admission, consent was obtained and participants completed a battery of assessments and their height and weight were measured. The New York State Psychiatric Institute (NYSPI) Institutional Review Board reviewed and approved this study. The final dataset included 162 participants with a mean age of $25.42 \pm 7.76$ years and a mean BMI of $16.06 \pm 1.84$ kg/m$^2$. Fifty percent of the study group had the binge/purge subtype of AN.

#### 2.1.2 | Bulimia nervosa

Participants were treatment-seeking participants with BN. Prior to treatment, consent was obtained and participants completed a battery of assessments and height and weight were measured. The final dataset included 93 participants with a mean age of $23.03 \pm 4.67$ years and a mean BMI of $22.08 \pm 3.32$ kg/m$^2$. Sixty-five participants were recruited through the New York State Psychiatric Institute (which provided both inpatient and outpatient treatment) and the remaining 28 were recruited through McLean Hospital's Klarman Eating Disorders Center (which provided residential treatment). The Partners Human Research Committee approved data collection and sharing at the McLean site, and the NYSPI Institutional Review Board approved all methods at the Columbia site.

#### 2.1.3 | Binge-eating disorder

Participants were consecutively evaluated treatment-seeking obese (BMI $\geq$ 30) adults who met DSM-5 criteria for BED. Prior to treatment, consent was obtained and participants completed a battery of assessments and height and weight were measured. All methods were approved by the Yale University School of Medicine Institutional

Review Board. The final dataset included 343 participants with a mean age of $46.71 \pm 10.35$ years and a mean BMI of $38.6 \pm 5.7$ kg/m$^2$.

### 2.2 | Procedures

Participants provided basic demographic information and height and weight were measured in-person. All diagnoses were made via the Eating Disorder Examination (described below) and/or Structured Clinical Interview for DSM-IV (First, Spitzer, Gibbon, & Williams, 2007).

The *Eating Disorder Examination (EDE)* is a semi-structured investigator-based interview assessing frequency and severity of cognitions and behaviors found in eating disorders (Cooper & Fairburn, 1987). The EDE yields four subscales: Dietary Restraint, Eating Concern, Shape Concern, and Weight Concern, and quantifies behaviors such as objective binge episodes (OBEs), and compensatory behaviors such as vomiting, use of laxative and/or diuretics, and excessive or compulsive exercise. The EDE has demonstrated test-retest reliability, and the four subscales are internally consistent, although internal consistencies are higher in samples of individuals with AN and BN compared to BED (Berg, Peterson, Frazier, & Crow, 2012). In the current analyses, the full EDE was available for participants with AN and BED and participants with BN from the McLean site. Abbreviated EDEs were available for participants with BN from the NYSPI site where the frequency of OBEs and compensatory behaviors could be calculated, but subscale scores were not available.

The *Eating Disorder Examination Questionnaire (EDE-Q)* is the self-report version of the EDE that assesses frequency of behaviors found in eating disorders, including OBEs and compensatory behaviors, and it also produces Dietary Restraint, Eating Concern, Shape Concern and Weight Concern subscale scores (Fairburn & Beglin, 1994). The EDE-Q has adequate test-retest reliability (Peterson et al., 2007; Reas, Grilo, & Masheb, 2006) and good convergence with the EDE (Grilo, Masheb, & Wilson, 2001a, Grilo, Masheb, & Wilson, 2001b; Mond, Hay, Rodgers, Owen, & Beumont, 2004; Wilfley, Schwartz, Spurrell, & Fairburn, 1997). In the current analyses, the EDE-Q was used to calculate the four subscale scores and the overvaluation classification scheme in the study group with BN. The EDE-Q was available for participants from both the NYSPI and McLean sites.

The *Clinical Impairment Scale* (CIA) is a 16-item self-report scale measuring psychosocial impairment due to an eating disorder (Bohn et al., 2008). The CIA has demonstrated good test-retest reliability, and construct and discriminant validity in samples with AN and BN (Bohn et al., 2008). In the current study, the CIA was administered to participants with AN and BN.

The *Beck Depression Inventory* (BDI) assessed depressive symptoms and was available for study groups with AN, BN, and BED (Beck, Ward, & Mendelson, 1961). The BDI has strong psychometric properties, including good internal consistency and concurrent validity (Beck, Steer, & Carbin, 1988).

The *Rosenberg Self-Esteem Scale* (RSES) is a well-established and widely-used measure of global self-esteem in which higher scores reflect higher self-esteem (Rosenberg, 1965). The RSES was available for a subset of participants with BN and BED. The RSES has

demonstrated good test-retest reliability and internal consistency (Schmitt & Allik, 2005).

The *MOS SF-36* is a 36-item self-report measure of health-related quality of life (Ware & Sherbourne, 1992). The MOS SF-36 scoring yields two component scores: a Physical Health $t$-score, and a Mental Health $t$-score. The Physical Health $t$-score provides information about physical functioning, role limitations due to physical health problems, and bodily pain. For example, a question used to calculate this score is "Does your health now limit you in climbing several flights of stairs?," which is rated on a 3-point scale from 1: Limited A Lot, to 3: Not Limited At All. The Mental Health $t$-score provides information about social and emotional functioning and includes questions such as "How much of the time during the past 4 weeks have you felt so down in the dumps that nothing could cheer you up?," rated on a 6-point scale from 1: All Of The Time, to 6: None Of The Time. The MOS-36 has demonstrated good internal consistency and concurrent validity (Jenkinson, Wright, & Coulter, 1994). The MOS SF-36 was available for participants with AN and BED.

## 2.3 | Severity groups using DSM-5 specifiers

For each disorder, we divided participants into groups based on the DSM-5 minimum severity criteria. For AN, where the severity specifiers are based on BMI, group designations were made following an in-person assessment of height and weight at time of admission to the NYSPI inpatient unit. The severity categories are: mild: $\geq 17.0$ kg/m$^2$, moderate: 16–16.99 kg/m$^2$, severe: 15–15.99 kg/m$^2$, extreme: $<15$ kg/m$^2$. For BN, the severity specifiers are based on frequency of inappropriate compensatory behaviors (e.g., self-induced vomiting), which was based on the EDE interview assessment. The phrasing of the EDE questions assessing fasting and exercise preclude our ability to determine whether these behaviors are compensatory or aimed only at decreasing weight and shape in general and therefore severity specifiers in the current analyses are based on frequency of compensatory self-induced vomiting, laxative, and diuretic use. Severity categories are: mild: 1–3 weekly episodes, moderate: 4–7 weekly episodes, severe: 8–13 weekly episodes, extreme: 14 or more weekly episodes. The EDE assesses the number of purging episodes per month, and thus we assigned participants to the severity categories by dividing this monthly total by four. For BED the severity specifiers are based on binge episode frequency, which was captured on the EDE. Severity groups are as follows: mild: 1–3 weekly episodes, moderate: 4–7 weekly episodes, severe 8–13 weekly episodes, extreme: 14 or more weekly episodes. The EDE assesses the number of binge episodes per month, and thus we assigned participants to the severity categories by dividing this monthly total by four.

## 2.4 | Severity groups for BN based on number of purging methods

The "single method" group was composed of participants who reported use of only one method of purging (e.g., self-induced vomiting, laxatives, or diuretics) in the previous 3 months, while the "multiple method" group was composed of participants who reported using more than one method of purging in the previous 3 months (Edler et al., 2007).

## 2.5 | Severity groups based on overvaluation of shape and weight: For BED and transdiagnostically

Consistent with prior research, for the AN and BED groups overvaluation of shape and weight was operationalized using two items from the EDE, and for the BN group the two corresponding items from the EDE-Q were used (Grilo et al., 2008; Ojserkis, Sysko, Goldfein, & Devlin, 2012): "Over the past 4 weeks, has your shape been important in influencing how you feel about (judge, think, evaluate) yourself as a person?" and "Over the past 4 weeks, has your weight been important in influencing how you feel about (judge, think, evaluate) yourself as a person?" Both items are rated from 0 to 6, with 0 indicating "no importance" and 6 indicating "supreme importance." The overvaluation of shape and weight group consisted of participants who rated either of these items at or above the clinical cutoff score of 4 (moderate importance).

## 2.6 | Statistical analyses

We conducted analysis of variance (ANOVA) with post hoc Tukey tests to examine pairwise differences between severity groups for demographic variables. We also conducted $\chi^2$ tests ($\chi^2$) to examine differences across severity groups for categorical demographic variables. To test our hypothesis that that severity of pathology and impairment would vary in a monotonic, increasing manner across the severity categories we conducted one-way ANOVA with linear contrast weights. To examine the alternative classification scheme using number of methods of purging in BN, two ANCOVAs were conducted with the *DSM-5* severity specifier categories and purging frequency as covariates. We examined the alternative classification scheme using overvaluation of shape and weight in BED and then again transdiagnostically. In BED, two ANCOVAs were done, with the *DSM-5* severity specifier categories and binge frequency as covariates, respectively. In examining the transdiagnostic scheme, two ANCOVAs were done, first with diagnosis and BMI as covariates and then adding *DSM 5* severity category as an additional covariate. When exploring the transdiagnostic classification scheme, we included results only for measures that were present in at least two of the three diagnostic groups. Partial $\eta^2$ was calculated as a measure of effect size, representing the proportion of variance in the dependent measure that was accounted for by the severity group classification, and effect sizes were classified as follows, small: 0.02, medium: 0.13, large: 0.26 (Cohen, 2013). All tests were two-tailed and an alpha of .05 was used as the threshold for statistical significance.

## 3 | RESULTS

Table 1 summarizes the demographic characteristics of each group by severity category. The AN and BN severity groups did not differ significantly on age, sex, or ethnicity/race, and the BED severity groups only

**TABLE 1** Demographic and physical characteristics of participants with anorexia nervosa, bulimia nervosa, and binge-eating disorder across DSM-5 severity groups

| | Anorexia nervosa | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Mild N = 39 | | Moderate N = 38 | | Severe N = 27 | | Extreme N = 32 | | Test statistic | p value | Effect size[c] |
| Age in years, mean (SD) | 23.58 | (5.71) | 23.94 | (9.59) | 25.84 | (6.85) | 28.17 | (8.03) | $F(3,118) = 2.31$ | .080 | .055 |
| Female, no (%) | 38 | (97.4%) | 34 | (89.5%) | 25 | (92.6%) | 31 | (96.9%) | $\chi^2(3, N = 136) = 2.81$ | .422 | .144 |
| White, no (%)[a] | 37 | (94.8%) | 35 | (92.1%) | 25 | (92.6%) | 30 | (93.8%) | $\chi^2(3, N = 136) = 2.75$ | .965 | .045 |
| Body mass index[b], mean, (SD) | 18.07 | (0.90) | 16.45 | (0.30) | 15.58 | (0.27) | 13.54 | (1.22) | $F(3,132) = 199.19$ | <.001 | .819 |
| | **Bulimia nervosa** | | | | | | | | | | |
| | Mild/moderate N = 30 | | Severe N = 25 | | Extreme N = 38 | | | | Test statistic | p value | Effect size[c] |
| Age in years, mean (SD) | 22.76 | | (4.25) | 24.24 | | (5.54) | 22.45 | (4.32) | $F(2,89) = 1.19$ | .309 | .026 |
| Female, no (%) | 30 | | (100%) | 25 | | (100%) | 37 | (97.4%) | $\chi^2(2, N = 93) = 1.46$ | .481 | .125 |
| White, no (%)[a] | 27 | | (90%) | 21 | | (84%) | 28 | (73.7%) | $\chi^2(2, N = 93) = 3.22$ | .212 | .183 |
| Body mass index[b], mean, (SD) | 23.19 | | (4.54) | 22.38 | | (2.61) | 21.01 | (2.22) | $F(2,83) = 3.65$ | .030 | .081 |
| | **Binge-eating disorder** | | | | | | | | | | |
| | Mild N = 157 | | Moderate N = 142 | | Severe N = 27 | | Extreme N = 17 | | Test statistic | p value | Effect size[c] |
| Age in years, mean (SD) | 46.17 | (9.96) | 47.76 | (10.50) | 44.93 | (12.66) | 45.71 | (10.35) | $F(3,336) = 0.95$ | .419 | .008 |
| Female, no (%) | 115 | (73.2%) | 98 | (69%) | 24 | (88.9%) | 10 | (58.8%) | $\chi^2(3, N = 343) = 6.04$ | .110 | .133 |
| White, no (%)[a] | 102 | (65%) | 90 | (63.4%) | 21 | (77.8%) | 6 | (35.3%) | $\chi^2(3, N = 343) = 8.37$ | .039 | .156 |
| Body mass index[b], mean, (SD) | 38.64 | (5.50) | 38.49 | (5.81) | 36.71 | (4.90) | 40.93 | (7.23) | $F(3,339) = 1.96$ | .120 | .017 |

*Note.* For anorexia nervosa, severity categories are based on Body Mass Index as follows: Mild: $\geq 17.0$ kg/m², Moderate: 16–16.99 kg/m², Severe: 15–15.99 kg/m², Extreme: $< 15$ kg/m². For bulimia nervosa, severity categories are based on frequency of inappropriate compensatory behaviors as follows: Mild: 1–3 weekly episodes, Moderate: 4–7 weekly episodes, Severe: 8–13 weekly episodes, Extreme: 14 or more weekly episodes. For binge-eating disorder, severity categories are based on frequency of binge episodes as follows: Mild: 1–3 weekly episodes, Moderate: 4–7 weekly episodes, Severe 8–13 weekly episodes, Extreme: 14 or more weekly episodes. [a]$\chi^2$ analysis for White versus non-White. [b]Body Mass Index = kg/m². [c]Effect size for ANOVA = partial $\eta^2$; Effect size for $\chi^2$ analysis = Cramer's V.

differed on the distribution of ethnicity/race, with the extreme group having about half as many White participants. Given that BMI is the variable used to distinguish the AN severity groups, as expected, the groups significantly differed on BMI, with progressively lower mean BMIs across the severity rankings. The BN severity groups also differed by BMI, with the extreme group having a significantly lower mean BMI than the mild/moderate group ($df = 60$, $t = 2.49$, $p = .016$).

## 3.1 | Anorexia nervosa: DSM-5 severity categories

Thirty-nine (24.1%) of the 162 participants with AN were classified as mild, 38 as moderate (23.5%), 27 (16.7%) as severe, and 32 (19.7%) as extreme. Table 2 summarizes descriptive statistics and statistical analyses for participants with AN comparing *DSM-5* severity groups on clinical measures. There were no significant linear trends across severity groups on BDI, CIA, the four EDE subscales (Restraint, Eating Concerns, Shape and Weight Concerns), or the component scores of the MOS SF-36. There was a significant linear trend for duration of illness, indicating that duration of illness increased across the severity groups

from mild to extreme. A significant linear trend was also found in the number of previous inpatient hospitalizations with those in the extreme category reporting approximately two more previous inpatient hospitalizations than those in the mild category. For both of these significant results, effect sizes were small.

## 3.2 | Bulimia nervosa: DSM-5 severity categories

Seven (7.5%) of the 93 participants with BN were classified as mild, 23 (24.7%) as moderate, 25 (26.9%) as severe, and 38 (40.9%) as extreme. Given the small number of participants classified as mild, the mild and moderate groups were combined for further analyses. Table 2 summarizes descriptive statistics and severity analyses for participants with BN. The BN study group was comprised of individuals from two sites, NYSPI, which was primarily comprised of outpatients, and McLean, which is a residential program. We empirically tested whether there might be a difference by site on our variables of interest. The group from McLean had higher scores than the group from NYSPI on the BDI ($F(1,86) = 58.43$, $p < .001$, partial $\eta^2 = .405$), CIA ($F(1,57) = 4.08$,

**TABLE 2** Comparison of clinical characteristics of participants with anorexia nervosa, bulimia nervosa, and binge-eating disorder across DSM-5 severity groups

| | Anorexia nervosa | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Mild N = 39 | | Moderate N = 38 | | Severe N = 27 | | Extreme N = 32 | | ANOVA | p value | Partial η² |
| | M | sd | M | sd | M | sd | M | sd | | | |
| Duration of illness (months) | 93.78 | (71.31) | 94.59 | (110.03) | 107.65 | (65.75) | 140.81 | (94.17) | 5.35 | .022 | .046 |
| Number of eating disorder hospitalizations | 1.51 | (1.68) | 2.24 | (2.43) | 1.74 | (1.89) | 3.56 | (3.91) | 6.10 | .017 | .059 |
| BDI | 24.51 | (11.53) | 25.58 | (12.70) | 27.00 | (11.82) | 28.26 | (10.53) | 1.90 | .171 | .016 |
| CIA | 37.62 | (8.06) | 35.81 | (9.39) | 34.64 | (8.62) | 37.97 | (8.79) | 0.00 | .986 | .000 |
| EDE restraint | 3.19 | (1.71) | 3.06 | (1.94) | 3.04 | (1.77) | 3.28 | (1.75) | 0.03 | .857 | .000 |
| EDE eating concern | 2.23 | (1.32) | 2.32 | (1.38) | 2.63 | (1.52) | 2.83 | (1.73) | 3.25 | .074 | .026 |
| EDE shape concern | 3.72 | (1.71) | 3.46 | (1.90) | 3.67 | (2.06) | 3.04 | (1.99) | 1.55 | .216 | .013 |
| EDE weight concern | 3.71 | (1.65) | 3.36 | (1.75) | 3.49 | (1.93) | 2.79 | (1.67) | 3.79 | .054 | .031 |
| MOS SF-36 | | | | | | | | | | | |
| Physical health T-score | 38.38 | (9.58) | 36.14 | (11.90) | 39.39 | (10.27) | 36.24 | (13.48) | 0.11 | .739 | .001 |
| Mental health T-score | 25.65 | (10.72) | 26.93 | (10.18) | 26.32 | (10.16) | 25.80 | (10.48) | 0.00 | .988 | .000 |

| | Bulimia nervosa | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Mild/Moderate N = 30 | | Severe N = 25 | | Extreme N = 38 | | ANOVA | p value | Partial η² |
| | M | sd | M | sd | M | sd | | | |
| BDI[a] | 23.51 | (17.13) | 19.96 | (13.34) | 23.08 | (12.89) | 0.01 | .918 | .000 |
| CIA[a] | 32.94 | (9.90) | 32.85 | (10.76) | 38.68 | (7.44) | 4.51 | .038 | .076 |
| EDE-Q restraint | 3.91 | (0.94) | 3.64 | (1.36) | 3.66 | (1.62) | 0.16 | .689 | .002 |
| EDE-Q eating concern | 3.51 | (1.06) | 3.64 | (0.93) | 4.56 | (1.15) | 15.83 | <.001 | .154 |
| EDE-Q shape concern[a] | 4.77 | (0.87) | 4.37 | (1.37) | 4.72 | (1.21) | 0.00 | .984 | .000 |
| EDE-Q weight concern[a] | 4.62 | (0.94) | 3.89 | (1.59) | 4.16 | (1.63) | 0.32 | .859 | .000 |
| RSES | 23.50 | (5.40) | 25.93 | (7.29) | 23.24 | (5.78) | 0.02 | .902 | .000 |

| | Binge-eating disorder | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Mild N = 157 | | Moderate N = 142 | | Severe/Extreme N = 44 | | ANOVA | p value | Partial η² |
| | M | sd | M | sd | M | sd | | | |
| BDI | 13.33 | (7.97) | 15.76 | (9.02) | 16.68 | (11.38) | 3.35 | .073 | .014 |
| EDE restraint | 1.68 | (1.23) | 1.80 | (1.31) | 1.66 | (1.33) | 0.00 | .948 | .000 |
| EDE eating concern | 1.76 | (1.18) | 2.07 | (1.28) | 2.82 | (1.48) | 19.36 | <.001 | .067 |
| EDE shape concern | 3.25 | (1.15) | 3.64 | (1.19) | 4.11 | (1.10) | 18.86 | <.001 | .053 |
| EDE weight concern | 2.95 | (.99) | 3.28 | (1.10) | 3.49 | (1.16) | 8.84 | .003 | .025 |
| RSES | 20.72 | (6.43) | 19.23 | (6.08) | 18.82 | (7.06) | 2.36 | .127 | .020 |
| MOS SF-36 | | | | | | | | | |
| Physical health T-score | 45.19 | (9.24) | 43.59 | (11.16) | 42.17 | (12.14) | 2.87 | .091 | .008 |
| Mental health T-score | 44.58 | (10.82) | 41.55 | (11.48) | 40.72 | (12.78) | 3.96 | .047 | .012 |

*Note.* BDI = Beck Depression Inventory, CIA = Clinical Impairment Assessment, EDE = Eating Disorder Examination, EDE-Q = Eating Disorder Examination-Questionnaire, RSES = Rosenberg Self-Esteem Scale. F ratios: AN df ranged from (1,102) to (1,128), BN df ranged from (1,46) to (1,85), BED df ranged from (1,113) to (1,340). [a] site entered as a covariate.

$p = .048$, partial $\eta^2 = .067$), EDE-Q Shape Concern subscale ($F(1,88) = 10.71$, $p = .002$, partial $\eta^2 = .109$, and EDE-Q Weight Concern subscale ($F(1,88) = 7.41$, $p = .008$, partial $\eta^2 = .078$). Therefore, site was entered as a covariate when conducting analyses examining these variables.

The distribution of individuals across the DSM 5 severity specifier categories did not differ by site ($\chi^2(1, n = 93) = 0.09$, $p = .955$, Cramer's $V = .031$). There were no significant linear trends across severity groups on BDI, Restraint, Shape, and Weight Concern subscales of the

**TABLE 3** Comparison of clinical characteristics of participants with bulimia nervosa across purging method groups

| | Bulimia nervosa | | | | | | | |
| | Single purging method $N = 65$ | | Multiple purging methods $N = 28$ | | ANOVA | | ANCOVA Covary purging frequency | ANCOVA Covary severity |
| | M | sd | M | sd | F | $\eta^2$ | $\eta^2$ | $\eta^2$ |
|---|---|---|---|---|---|---|---|---|
| BDI[a] | 21.23 | (14.63) | 24.69 | (13.86) | 0.14 | .002 | .002 | .000 |
| CIA[a] | 33.55 | (9.86) | 39.13 | (7.59) | 3.75 | .063 | .058 | .046 |
| EDE-Q restraint | 3.68 | (1.40) | 4.21 | (1.34) | 2.76 | .030 | .038 | .027 |
| EDE-Q eating concern | 3.79 | (1.08) | 4.41 | (1.24) | 5.63* | .060 | .046 | .020 |
| EDE-Q shape concern[a] | 4.75 | (1.20) | 5.13 | (0.75) | 1.43 | .016 | .014 | .016 |
| EDE-Q weight concern[a] | 4.25 | (1.49) | 5.00 | (0.87) | 4.82* | .052 | .058 | .065 |
| RSES | 24.63 | (6.76) | 22.71 | (4.05) | 0.97 | .020 | .021 | .019 |

*Note.* BDI = Beck Depression Inventory, CIA = Clinical Impairment Assessment, EDE-Q = Eating Disorder Examination-Questionnaire, RSES = Rosenberg Self-Esteem Scale. F ratios: df ranged from (1,47) to (1,88) for ANOVAs, and from (1,46) to (1,87) for ANCOVAs adjusting for purging frequency and for purging severity. [a] site also entered as a covariate. *$p < .05$, **$p < .01$, ***$p < .001$

EDE-Q, or RSES. There were significant linear trends for the CIA, with a small effect size, and the Eating Concerns subscale of the EDE-Q, with a medium effect size, such that greater eating disorder-related impairment and severity of eating concerns increased across severity specifier groups.

## 3.3 | Bulimia nervosa: Purging method groups

In this group of 93 participants with BN, 65 (69.9%) were categorized as reporting a single method of purging, while 28 (30.1%) were categorized as reporting multiple methods of purging. The single method of purging group had a higher mean age (M = 23.95 years, SD = 4.82) compared to the multiple methods of purging group (M = 20.96 years, SD = 3.59), $F(1,90) = 8.56$, $p = .004$, partial $\eta^2 = .087$. The groups did not differ significantly on BMI, (Single Method group M = 22.31, SD = 3.66; Multiple Method group M = 21.55, SD = 2.33), $F(1,84) = 0.94$, $p = .335$, partial $\eta^2 = .001$; sex (Single Method group % female = 98.46%, Multiple Method group % female = 100%), $\chi^2(1, n = 93) = 0.43$, $p = .509$, Cramer's V = .068; or ethnicity/race (Single Method group % white = 81.54%, Multiple Method group % white = 82.14%), $\chi^2(1, n = 93) = 0.01$, $p = .945$, Cramer's V = .007. Table 3 summarizes descriptive statistics and severity analyses for participants with BN comparing the purging method groups. The multiple purging methods group had scores reflecting significantly greater severity of eating disorder-related impairment, general eating pathology, and eating and weight concerns, than the single purging method group, though all effect sizes were in the small range. Furthermore, the majority of effect sizes did not change meaningfully after including the *DSM-5* severity categories and purging frequency as covariates.

## 3.4 | Binge-eating disorder: DSM-5 severity categories

In this group of 343 participants with BED, 157 (45.8%) were classified as mild, 142 (41.4%) as moderate, 27 (7.9%) as severe, and 17 (5%) as extreme. Given the small number of participants classified as severe and extreme, these groups were combined for further analyses. Table 2

summarizes descriptive statistics and severity analyses for participants with BED. A linear trend was not detected across severity groups on the BDI, the Restraint subscale of the EDE, RSES, or the Physical Health component t-score of the MOS SF-36. There was a significant linear trend indicating an increase on the Eating, Shape, and Weight Concerns subscales of the EDE. Additionally, a significant linear trend was found indicating greater impairment across severity groups on the Mental Health component t-score of the MOS SF-36. All effect sizes were in the small range.

## 3.5 | Binge-eating disorder: Overvaluation of shape and weight groups

In this group of 343 participants with BED, 239 (69.7%) were categorized as having clinically significant levels of overvaluation of shape and weight, while 104 (30.3%) were categorized as not having clinical levels of overvaluation of shape and weight. The overvaluation groups did not differ significantly on age (No Overvaluation group M = 47.55 years, SD = 11.04; Overvaluation group M = 46.34 years, SD = 10.04), $F(1,338) = 0.99$, $p = .320$, partial $\eta^2 = .003$, BMI (No Overvaluation group M = 38.20, SD = 5.50, Overvaluation group M = 38.75, SD = 5.82), $F(1,339) = 0.68$, $p = .409$, partial $\eta^2 = .002$, sex, (No Overvaluation group % female = 72.1%, Overvaluation group % female = 71.2%), $\chi^2(1, n = 343) = 0.00$, $p = .977$, Cramer's V = .002 or ethnicity/race, (No Overvaluation group % white = 58.7%, Overvaluation group % white = 66.1%), $\chi^2(1, n = 343) = 1.75$, $p = .187$, Cramer's V = .071. Table 4 summarizes descriptive statistics and severity analyses for participants with BED comparing the overvaluation groups. The overvaluation group had scores reflecting significantly greater severity than the group without overvaluation on all but one of eight variables. Effect sizes were in the large range for the Shape and Weight Concerns subscales of the EDE, and in the small range for the other five variables. Furthermore, the majority of effect sizes did not change meaningfully after including the DSM-5 severity categories and binge frequency as covariates.

**TABLE 4** Comparison of clinical characteristics of participants with binge-eating disorder across overvaluation groups

| | Binge-eating disorder | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | No overvaluation N = 104 | | Overvaluation N = 239 | | ANOVA | | ANCOVA Covary OBE frequency | ANCOVA Covary severity |
| | M | sd | M | sd | F | $\eta^2$ | $\eta^2$ | $\eta^2$ |
| BDI | 10.47 | (7.55) | 16.63 | (8.93) | 37.46*** | .100 | .098 | .095 |
| EDE restraint | 1.52 | (1.17) | 1.82 | (1.31) | 4.23* | .012 | .012 | .012 |
| EDE eating concern | 1.41 | (1.06) | 2.29 | (1.31) | 36.14*** | .096 | .094 | .090 |
| EDE shape concern | 2.54 | (1.05) | 3.95 | (0.98) | 142.58*** | .295 | .296 | .293 |
| EDE weight concern | 2.29 | (0.76) | 3.54 | (0.97) | 136.15*** | .285 | .284 | .281 |
| RSES | 32.31 | (5.77) | 28.81 | (5.80) | 9.34** | .076 | .081 | .081 |
| MOS SF-36 | | | | | | | | |
|   Physical health T-score | 44.89 | (10.15) | 43.81 | (10.63) | 0.77 | .002 | .002 | .002 |
|   Mental health T-score | 47.80 | (9.37) | 40.65 | (11.61) | 30.37*** | .083 | .081 | .079 |

*Note.* BDI = Beck Depression Inventory, CIA = Clinical Impairment Assessment, EDE = Eating Disorder Examination, RSES = Rosenberg Self-Esteem Scale.. F ratios: df ranged from (1, 114) to (1,341) for ANOVAs, and from (1,113) to (1,340) for ANCOVAs adjusting for binge eating frequency and for binge eating severity. *$p < .05$, **$p < .01$, ***$p < .001$.

## 3.6 | Transdiagnostic severity: Overvaluation of shape and weight groups

The overvaluation groups differed significantly by diagnosis, $\chi^2(2, n = 552) = 27.09$, $p < .001$, Cramer's V = .222. In the AN group, 102 (85.7%) were categorized as having clinically significant levels of overvaluation of shape and weight, while 17 (14.3%) were categorized as not having clinical levels of overvaluation of shape and weight. In the BN group, 83 (92.2%) were categorized as having clinically significant levels of overvaluation of shape and weight, while 7 (7.8%) were categorized as not having clinical levels of overvaluation of shape and weight. As noted above, in BED, 239 (69.7%) were categorized as having clinically significant levels of overvaluation of

shape and weight, while 104 (30.3%) were categorized as not having clinical levels of overvaluation of shape and weight. The overvaluation groups also differed significantly on BMI (No Overvaluation group M = 34.61, SD = 9.41, Overvaluation group M = 30.56, SD = 11.07), $F(1,530) = 13.82$ $p < .001$, partial $\eta^2 = .025$, age (No Overvaluation group M = 43.61, SD = 13.55, Overvaluation group M = 36.77, SD = 14.05), $F(1,535) = 23.11$, $p < .001$, partial $\eta^2 = .041$, and ethnicity/race, (No Overvaluation group % white = 62.99%, Overvaluation group % white = 76.07%), $\chi^2(1, n = 549) = 8.47$, $p = .004$, Cramer's V = .124. The overvaluation groups did not differ significantly on sex (No Overvaluation group % female = 75.78%, Overvaluation group % female = 82.78%), $\chi^2(1, n = 552) = 3.15$, $p = .076$, Cramer's V = .076.

**TABLE 5** Transdiagnostic comparison of clinical characteristics of participants across overvaluation groups controlling for diagnosis and body mass index

| | No overvaluation N = 128 | | Overvaluation N = 424 | | ANCOVA Covary diagnosis, age, BMI | | ANCOVA Covary diagnosis, age, BMI, severity |
|---|---|---|---|---|---|---|---|
| | M | sd | M | Sd | F | $\eta^2$ | $\eta^2$ |
| BDI[a] | 10.88 | (8.25) | 20.52 | (11.64) | 50.54*** | .091 | .088 |
| CIA[b] | 27.17 | (10.98) | 37.86 | (8.25) | 22.58*** | .141 | .146 |
| EDE restraint[c] | 1.44 | (1.20) | 2.27 | (1.59) | 17.17*** | .039 | .039 |
| EDE eating concern[c] | 1.37 | (1.02) | 2.40 | (1.35) | 49.53*** | .104 | .100 |
| EDE shape concern[c] | 2.35 | (1.19) | 3.94 | (1.18) | 156.28*** | .267 | .264 |
| EDE weight concern[c] | 2.14 | (0.91) | 3.57 | (1.16) | 165.88*** | .245 | .243 |
| MOS SF-36 | | | | | | | |
|   Physical health T-score[c] | 44.60 | (10.25) | 42.07 | (11.22) | 1.63 | .004 | .003 |
|   Mental health T-score[c] | 46.76 | (9.98) | 36.92 | (13.08) | 41.08*** | .091 | .089 |
| RSES[d] | 32.24 | (5.85) | 26.82 | (6.23) | 17.46*** | .100 | .109 |

*Note.* BDI = Beck Depression Inventory, CIA = Clinical Impairment Assessment, EDE = Eating Disorder Examination, RSES = Rosenberg Self-Esteem Scale.. F ratios: df ranged from (1,138) to (1,506) for ANCOVAs adjusting for diagnosis, age, and BMI, and from (1,137) to (1,505) for ANCOVAs adjusting for diagnosis, BMI, and DSM 5 severity category. [a]available for AN, BN, and BED, [b]available for AN and BN, [c]available for AN and BED, and [d] available for BN and BED. ***$p < .001$

Table 5 summarizes severity analyses comparing the overvaluation groups transdiagnostically while covarying for diagnosis, BMI, and age. The overvaluation group had scores reflecting significantly greater severity than the group without overvaluation on all but one of nine variables. Effect sizes were in the large range for the Shape and Weight Concerns subscales of the EDE, the medium range for the CIA, and in the small range for the other 5 significant variables. Furthermore, the majority of effect sizes did not change meaningfully after including the DSM-5 severity categories as covariates.

## 4 | DISCUSSION

The current study is among the first to assess the validity and utility of the recently proposed DSM-5 measures of severity across multiple eating disorder diagnoses. In addition, it examined a range of psychological and clinical outcomes and also examined existing proposed alternative severity classification schemes for the purging form of BN and BED, and a potential transdiagnostic severity classification scheme.

Ideally, a measure of severity would provide an index not only of the intensity of defining psychopathological features of the illness, but also of the level of functional impairment and the need for treatment services. The latter, for example, is likely to be of interest to patients, clinicians, and insurance providers. Our findings suggest that the DSM-5 severity ratings for AN may provide information about both aspects of severity; BMI is a direct measure of the low body weight required by Criterion A, and also was related to indicators of need for greater services, such as number of hospitalizations, which is consistent with other literature finding that low BMI predicted poor long-term outcomes and mortality in AN (Button, Chadalavada, & Palmer, 2010; Löwe et al., 2001; Misra et al., 2003). The DSM-5 severity measures for BN and BED are based on frequency of purging and binging, respectively—defining behavioral characteristics of these disorders. Our findings for the DSM-5 severity specifiers for the purging form of BN and BED suggest that they also show some statistically significant associations with cross-sectional measures of psychopathology, replicating results found in other samples (Grilo et al., 2015a, 2015b, 2015c). However, existing proposed alternative severity classifications for BN (based upon number of purging methods) and BED (based on overvaluation of shape and weight) had more robust patterns of statistically significant associations with cross-sectional measures of psychopathology than the DSM-5 severity specifiers. The transdiagnostic overvaluation classification scheme also showed a more robust pattern of associations between measures of pathology than the DSM-5 severity classification system, though it is notable that only a very small proportion of the AN and BN groups did not endorse overvaluation of shape and weight, and in our current sample we only had one assessment given across all 3 diagnoses (BDI), thus reducing the utility and applicability of this classification system. It is also important to acknowledge these alternative classification schemes are dichotomous in nature, thus making their comparison to a dimensional classification scheme, such as the DSM-5 categories, somewhat limited.

Furthermore, effect sizes were generally in the small range among the significant results, including those found in examining alternative classification schemes, and a potential transdiagnostic classification scheme. DSM 5 notes that clinicians have discretion to alter the severity specifier classification for a patient based upon functional impairment, need for supervision, and intensity of symptoms. In the current analyses, we based severity specifier classification solely on BMI, purging, and binge eating frequency, and were not able to include potential clinician-judged alterations to these categorizations. It is therefore possible that the DSM 5 severity specifiers may have greater ability to predict symptom severity and other cross-sectional measures of pathology and impairment when clinical judgment is incorporated. It is also possible that the severity specifiers may have greater utility predicting other outcomes, such as prognosis, treatment response, and medical complications, which were not assessed in the current sample. Thus, future research will be needed to fully determine the clinical utility of these classification schemes.

Future research, in addition to replicating these findings, should include other potential validators, especially those related to course of illness, functional impairment, and treatment outcomes. Notably, in the current study self-reported psychological and physical impairment were compared among the different severity specifier groups; however, objective assessments of physical symptoms and medical conditions were not available. Therefore, we cannot determine whether the current severity specifier system might have utility in distinguishing groups based upon objectively-measured medical impairment. Given that the eating disorders, in general, and AN, in particular, are known to be associated with substantial medical morbidity, this is an important area for future study.

Additionally, the number of participants with AN or BN in the current study groups might have limited our ability to detect small group differences. Very few participants with BN were categorized as "mild," and very few participants with BED were categorized as "extreme," thus limiting power to detect significant differences between all groups. In future samples, more equal distribution of individuals across severity groups will be important. At the same time, that there were relatively few people in some of these severity groups provides valuable information about the frequency with which individuals in the various severity specifier groups are seen in a treatment-seeking clinical facility. In both the BN and BED study groups, more severe groups tended to be more ethnically diverse, though this difference was only significant in the BED group. These differences may be due to a true effect, such that individuals of color are more likely to have more severe eating disorder presentations, or it is possible that it may be due to treatment-seeking bias, such that in ethnically diverse groups, more severe pathology is needed for treatment-seeking to occur. Future research with larger samples will be needed to see if this finding is replicated.

In the current study, we did not assess severity in non-treatment seeking individuals and all participants with AN sought out and received inpatient treatment, suggesting that this study group was likely more impaired than a group composed of both inpatients and outpatients. All participants with BED were obese and sought out and

received outpatient treatment, and the BN group included a mix of participants receiving inpatient, outpatient, and residential care; therefore, it is possible that *DSM-5* severity categorization could have been confounded by level of care. Additionally, the BN group was comprised of individuals who endorsed binge eating and compensatory behaviors at a frequency of at least 2 times per week, and as a result does not capture the full spectrum of individuals who would meet for a *DSM-5* BN diagnosis. Due to limitations of the assessment measures used, we were not able to included excessive exercise as one of the compensatory behaviors to derive the *DSM 5* severity categories. Future research examining the severity specifiers among non-treatment seeking samples, across levels of care, across weight categories, including excessive exercise as a compensatory behavior, and incorporating individuals with a lower frequency of compensatory behaviors would further illuminate the validity and utility of this classification scheme.

In addition to more thorough exploration of the utility of the existing severity specifiers, continued examination of the utility of alternative approaches to severity classification, such as single vs. multiple methods of purging in BN and overvaluation of shape and weight in BED are important. Further exploration of a transdiagnostic severity classification system such as overvaluation of shape and weight may be fruitful, and it may also be useful to develop a measure that can demonstrate more discrimination within eating disorder groups. As noted above, there are many cognitive and behavioral similarities across AN, BN, and BED, as well as crossover between these diagnoses over time, suggesting that a transdiagnostic approach to severity criteria may be warranted. In the current study groups we had limited overlap of measures between the diagnoses thus limiting our ability to fully test overvaluation of shape and weight as a transdiagnostic classification scheme across the diagnoses. Given that overvaluation of shape and weight has been proposed as a common "core pathology" across several eating disorder diagnoses, and our preliminary results with a limited transdiagnostic sample demonstrated that those who endorsed overvaluation tended to also report greater severity of many cross-sectional measures of psychopathology, it may be a useful lens through which to continue to examine disorder severity (Fairburn et al., 2003). Continued work is needed to examine how to most effectively operationalize severity in the eating disorders.

## REFERENCES

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77–100.

Beck, A. T., Ward, C., & Mendelson, M. (1961). Beck depression inventory (BDI). *Archives of General Psychiatry*, 4(6), 561–571.

Berg, K. C., Peterson, C. B., Frazier, P., & Crow, S. J. (2012). Psychometric evaluation of the eating disorder examination and eating disorder examination-questionnaire: A systematic review of the literature. *International Journal of Eating Disorders*, 45(3), 428–438.

Bohn, K., Doll, H. A., Cooper, Z., O'Connor, M., Palmer, R. L., & Fairburn, C. G. (2008). The measurement of impairment due to eating disorder psychopathology. *Behaviour Research and Therapy*, 46(10), 1105–1110.

Button, E. J., Chadalavada, B., & Palmer, R. L. (2010). Mortality and predictors of death in a cohort of patients presenting to an eating disorders service. *International Journal of Eating Disorders*, 43(5), 387–392.

Castellini, G., Sauro, C. L., Mannucci, E., Ravaldi, C., Rotella, C. M., Faravelli, C., & Ricca, V. (2011). Diagnostic crossover and outcome predictors in eating disorders according to DSM-IV and DSM-V proposed criteria: A 6-year follow-up study. *Psychosomatic Medicine*, 73(3), 270–279.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Cooper, Z., & Fairburn, C. (1987). The eating disorder examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders*, 6(1), 1–8.

Dakanalis, A., Bartoli, F., Caslini, M., Crocamo, C., Zanetti, M. A., Riva, G., . . . , Carrà, G. (2016). Validity and clinical utility of the DSM-5 severity specifier for bulimia nervosa: Results from a multisite sample of patients who received evidence-based treatment. *European Archives of Psychiatry and Clinical Neuroscience*, doi:10.1007/s00406-016-0712-7.

Dakanalis, A., Colmegna, F., Riva, G., & Clerici, M. (2017). Validity and utility of the DSM-5 severity specifier for binge-eating disorder. *International Journal of Eating Disorders*.

Eddy, K. T., Crosby, R. D., Keel, P. K., Wonderlich, S. A., Le Grange, D., Hill, L. , . . . , Mitchell, J. E. (2009). Empirical identification and validation of eating disorder phenotypes in a multisite clinical sample. *The Journal of Nervous and Mental Disease*, 197(1), 41–49.

Edler, C., Haedt, A. A., & Keel, P. K. (2007). The use of multiple purging methods as an indicator of eating disorder severity. *The International Journal of Eating Disorders*, 40(6), 515–520.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *The International Journal of Eating Disorders*, 16(4), 363–370.

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. *Behaviour Research and Therapy*, 41(5), 509–528.

Favaro, A., & Santonastaso, P. (1996). Purging behaviors, suicide attempts, and psychiatric symptoms in 398 eating disordered subjects. *International Journal of Eating Disorders*, 20(1), 99–103.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2007). Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P, 1/2007 revision). New York: Biometrics Research, New York State Psychiatric Institute.

Grilo, C. M., Hrabosky, J. I., White, M. A., Allison, K. C., Stunkard, A. J., & Masheb, R. M. (2008). Overvaluation of shape and weight in binge eating disorder and overweight controls: Refinement of a diagnostic construct. *Journal of Abnormal Psychology*, 117(2), 414.

Grilo, C. M., Ivezaj, V., & White, M. A. (2015a). Evaluation of the DSM-5 severity indicator for bulimia nervosa. *Behaviour Research and Therapy*, 67, 41–44.

Grilo, C. M., Ivezaj, V., & White, M. A. (2015b). Evaluation of the DSM-5 severity indicator for binge eating disorder in a clinical sample. *Behaviour Research and Therapy*, 71, 110–114.

Grilo, C. M., Ivezaj, V., & White, M. A. (2015c). Evaluation of the DSM-5 severity indicator for binge eating disorder in a community sample. *Behaviour Research and Therapy*, 66, 72–76.

Grilo, C. M., Masheb, R. M., & Wilson, G. T. (2001a). A comparison of different methods for assessing the features of eating disorders in

patients with binge eating disorder. *Journal of Consulting and Clinical Psychology*, *69*(2), 317–322.

Grilo, C. M., Masheb, R. M., & Wilson, G. T. (2001b). Different methods for assessing the features of eating disorders in patients with binge eating disorder: A replication. *Obesity Research*, *9*(7), 418–422.

Jenkinson, C., Wright, L., & Coulter, A. (1994). Criterion validity and reliability of the SF-36 in a population sample. *Quality of Life Research*, *3*(1), 7–12.

Löwe, B., Zipfel, S., Buchholz, C., Dupont, Y., Reas, D., & Herzog, W. (2001). Long-term outcome of anorexia nervosa in a prospective 21-year follow-up study. *Psychological Medicine*, *31*(05), 881–890.

Machado, P. P., Grilo, C. M., & Crosby, R. D. (2017). Evaluation of the DSM-5 Severity Indicator for Anorexia Nervosa. *European Eating Disorders Review*, *25*(3), 221–223.

Misra, M., Miller, K. K., Bjornson, J., Hackman, A., Aggarwal, A., Chung, J., . . . Klibanski, A. (2003). Alterations in growth hormone secretory dynamics in adolescent girls with anorexia nervosa and effects on bone metabolism. *The Journal of Clinical Endocrinology & Metabolism*, *88*(12), 5615–5623.

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. (2004). Validity of the eating disorder examination questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour Research and Therapy*, *42*(5), 551–567.

Ojserkis, R., Sysko, R., Goldfein, J. A., & Devlin, M. J. (2012). Does the overvaluation of shape and weight predict initial symptom severity or treatment outcome among patients with binge eating disorder?. *The International Journal of Eating Disorders*, *45*(4), 603–608.

Peterson, C. B., Crosby, R. D., Wonderlich, S. A., Joiner, T., Crow, S. J., Mitchell, J. E., . . . , le Grange, D. (2007). Psychometric properties of the eating disorder examination-questionnaire: Factor structure and internal consistency. *The International Journal of Eating Disorders*, *40*(4), 386–389.

Reas, D. L., Grilo, C. M., & Masheb, R. M. (2006). Reliability of the eating disorder examination-questionnaire in patients with binge eating disorder. *Behaviour Research and Therapy*, *44*(1), 43–51.

Rosenberg, M. (1965). *Society and the adolescent self-image* (Vol. 11, p. 326). Princeton, NJ: Princeton University Press.

Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg self-esteem scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology*, *89*(4), 623.

Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, *30*(6), 473–483.

Wilfley, D. E., Schwartz, M. B., Spurrell, E. B., & Fairburn, C. G. (1997). Assessing the specific psychopathology of binge eating disorder patients: Interview or self-report? *Behaviour Research and Therapy*, *35*(12), 1151–1159.

Wilson, G. T., & Sysko, R. (2009). Frequency of binge eating episodes in bulimia nervosa and binge eating disorder: Diagnostic considerations. *International Journal of Eating Disorders*, *42*(7), 603–610.