

**BRIEF REPORT**

# Longitudinal follow-up of a comprehensive operationalization of eating disorder recovery: Concurrent and predictive validity

Anna M. Bardone-Cone PhD | Alexandra Alvarez MSW | Jenna Gorlick BA |  
Katherine A. Koller BA | Katherine A. Thompson MA | Alexandra J. Miller MA

Department of Psychology & Neuroscience,  
University of North Carolina at Chapel Hill,  
Chapel Hill, North Carolina

**Correspondence**

Anna Bardone-Cone, PhD, CB #3270 Davie  
Hall, Department of Psychology and  
Neuroscience, University of North Carolina,  
Chapel Hill, NC 27599.  
Email: bardonecone@unc.edu

**Present address**

Alexandra Alvarez, Department of Social  
Work, University of North Carolina at Chapel  
Hill, Chapel Hill, NC. Jenna Gorlick,  
Department of Nutritional Sciences, University  
of Michigan, Ann Arbor, MI.  
Katherine A. Koller, College of Education,  
Lehigh University, Bethlehem, PA.

**Funding information**

UNC Department of Psychology and  
Neuroscience Stephenson and Lindquist  
Award; UNC Research Council Small Grant

**Abstract**

**Objective:** We aimed to replicate the concurrent validity of a comprehensive definition of eating disorder recovery (physical, behavioral, and cognitive indices) in a sample followed up 7–8 years from baseline, and to examine, for the first time with this comprehensive definition, predictive validity.

**Method:** Participants were 66 women with a history of an eating disorder and 31 age-matched controls who completed an online survey and phone interview.

**Results:** In general, women who were fully recovered were statistically indistinguishable from controls and had significantly less eating disorder attitudes and behaviors than the partially recovered and eating disorder groups. Being fully recovered at baseline was a robust predictor of stability: of those fully recovered at baseline, 80% remained fully recovered at follow-up. One-third of those with an eating disorder and one-half of those in partial recovery at baseline attained full recovery at follow-up.

**Discussion:** These findings support the current operationalization of eating disorder recovery, encompassing physical, behavioral, and cognitive indices, as valid and highlight that full recovery is not only possible but predicts full recovery years later. Future research should examine this operationalization in diverse samples and study trajectories of recovery to identify predictors.

**KEYWORDS**

concurrent validity, eating disorder cognitions, eating disorders, predictive validity, recovery

## 1 | INTRODUCTION

The eating disorder (ED) field has yet to arrive at a consensus definition of recovery (Bardone-Cone, Hunt, & Watson, 2018). This is problematic since it complicates comparisons across treatment outcome studies and limits the ability to identify reliable predictors of recovery. Historically, recovery models for EDs have conceptualized recovery as physical and behavioral changes (e.g., a minimum body mass index [BMI]; absence of purging). More recently, recovery models have

highlighted the need for a cognitive index of recovery (e.g., absence of ED-related thoughts), with evidence for the added value of this aspect of recovery (Bachner-Melman, Zohar, & Ebstein, 2006; Bardone-Cone et al., 2010; Couturier & Lock, 2006). For example, Bachner-Melman et al. (2006) found that women who recovered both behaviorally and cognitively had significantly lower levels across an array of disordered eating measures when compared to those recovered behaviorally but not cognitively.

In order to better understand the meaningfulness of these more comprehensive models of recovery, existing models need to be examined for construct and predictive validity. We focus on the recovery model proposed by Bardone-Cone et al. (2010), which includes

Results of this study were presented, in part, at the Eating Disorders Research Society Annual Meeting, September 2017, Leipzig, Germany.

physical, behavioral, and cognitive indices of recovery and has been recommended by researchers comparing different definitions (Ackard, Richter, Egan, & Cronemeyer, 2014). In the current work, we use data from a 7 to 8 year follow-up of a sample with a history of an ED with two aims: (a) to replicate concurrent validity findings of this comprehensive recovery model and (b) to, for the first time, examine predictive validity of this model. Evaluating whether recovery criteria identify individuals who may sustain a recovered state (predictive validity) is important for distinguishing between remission and recovery and identifying targets of intervention to minimize the likelihood of relapse.

## 2 | METHODS

### 2.1 | Participants and procedure

This study was a follow-up of 96 females with an ED history who participated in a study in 2007–2008; two participants were deceased at follow-up (2014–2015) and three did not provide permission to be recontacted, resulting in 91 possible participants. (See Supporting Information for details about the original study.) Sixty-six participated in the follow-up study, representing 73% of the 91 possible participants and 85% of the 78 we were able to recontact. Additionally, 31 female control participants were recruited through fliers and a university-wide listserv posting and enrolled if, based on a phone screen, they reported no history of eating pathology and fit an age band represented in the ED history sample.

Participants completed an online survey and, about 2 weeks later, a phone interview. Interviewers were the first author and several advanced undergraduate and postbaccalaureate research assistants, all extensively trained on ED diagnostic criteria and the diagnostic interview used. Participants received an Amazon.com gift card as remuneration. All aspects of this study were approved by the university's Institutional Review Board.

### 2.2 | Measures

#### 2.2.1 | Defining ED recovery status

Measures used to categorize individuals in terms of recovery status (ED, partial recovery, full recovery—defined below) were the ones used in the original recovery work by Bardone-Cone et al. (2010).

For current ED diagnosis, the *Structured Clinical Interview for DSM-IV, Patient Edition* (First, Spitzer, Gibbon, & Williams, 2002) was administered in the interview to diagnose anorexia nervosa (AN), bulimia nervosa (BN), binge eating disorder (BED), and eating disorder not otherwise specified (EDNOS). DSM-IV criteria were assessed due to baseline data collection using these criteria and the replication aim. As in the original work, for the diagnosis of AN, amenorrhea was not required. Based on a random subset (~10%) of diagnostic interviews of those with a history of an ED,  $\kappa$  was .67 for current DSM-IV EDs, reflecting substantial agreement (Landis & Koch, 1977).

For physical recovery, a BMI  $\geq 18.5$  kg/m<sup>2</sup> was required; this aligns with the World Health Organization's recommendation of a BMI  $< 18.5$  reflecting "underweight" (Björntorp, 2002). We used self-reported weight and height (from the survey), which are reasonable proxies for measured reports in both general and ED samples (Craig & Adams, 2008; McCabe, McFarlane, Polivy, & Olmsted, 2001). Behavioral recovery was assessed during the interview; presence of binge eating, vomiting, laxative use, or fasting was assessed over the past 3 months using annotated calendars. Absence of all four ED behaviors was required to meet behavioral recovery criteria. Cognitive recovery was assessed in the survey with the Eating Disorder Examination-Questionnaire (EDE-Q; Fairburn & Beglin, 1994) which contains four subscales providing broad coverage of ED cognitions over the past 28 days: restraint, eating concern, weight concern, shape concern. Obtaining scores within 1 SD of age-matched community norms for each of the EDE-Q subscales (Mond, Hay, Rodgers, & Owen, 2006) was required for cognitive recovery. In this study, coefficient alphas for these subscales were .85–.94.

Following the operationalization in Bardone-Cone et al. (2010), full recovery required: absence of an ED diagnosis; physical recovery, operationalized as a BMI  $\geq 18.5$  kg/m<sup>2</sup>; behavioral recovery, operationalized as no binge eating, vomiting, laxatives, or fasting in the past 3 months; and cognitive recovery, operationalized as all four EDE-Q subscales within 1 SD of age-matched community norms (Mond et al., 2006). Partial recovery required: absence of an ED diagnosis, physical recovery, and behavioral recovery, but the absence of cognitive recovery (i.e., 1+ EDE-Q subscale  $> 1$  SD of norms).

#### 2.2.2 | Measures for concurrent validity

Validation measures assessed constructs related to disordered eating but not part of the recovery operationalization. Measures were those used in Bardone-Cone et al. (2010) given the replication aim. In brief, the following well-established measures with strong psychometric support were administered in the survey: (a) Eating Attitudes Test (EAT-26; Garner, Olmsted, Bohr, & Garfinkel, 1982); (b) Appearance Schemas Inventory-Revised (ASI-R; Cash, Melnyk, & Hrabosky, 2004)—psychological investment in appearance; (c) body shame subscale of the Objectified Body Consciousness Scale (McKinley & Hyde, 1996); (d) brief version of the Thinness and Restricting Expectancy Inventory (TREI; Davis, Guller, & Smith, 2016; Hohlstein, Smith, & Atlas, 1998)—expectations for overgeneralized life improvement due to thinness or dieting; (e) thin-ideal internalization subscale of the Sociocultural Attitudes Toward Appearance Questionnaire-3 (Thompson, van den Berg, Roehrig, Guarda, & Heinberg, 2004). All measures had good reliability in the current sample ( $\alpha$ : .88–.96).

We assessed obsessiveness in the manner developed by Bardone-Cone et al. (2010), assessing (a) amount of time participants typically spent thinking about their weight, part of their body, and food/eating ( $\alpha = .84$ ), and (b) difficulty stopping thinking about these topics ( $\alpha = .89$ ). Participants were directed to respond in terms of recent thinking patterns. We also replicated asking about strength of urges to engage in ED behaviors in the past week (binge eating, vomiting,

laxative use, food restriction;  $\alpha = .75$ ) using items developed by Bardone-Cone et al. (2010).

### 2.3 | Analytic strategy

To test the concurrent validity aim, we performed one-way analysis of variance with recovery status as the independent variable and disordered eating measures as dependent variables. Significant omnibus *F*-statistics were followed up by Tukey's tests for pairwise comparisons among the four groups (ED, partial recovery, full recovery, controls). The Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) was applied to correct for multiple comparisons, with the false discovery rate (the proportion of significant results that are in reality false positives) set to .05. The Benjamini-Hochberg procedure is preferable to the Bonferroni correction since it provides greater statistical power (Thissen, Steinberg, & Kuang, 2002). To test the predictive validity aim, we examined the contingency table of recovery status groups at baseline and follow-up.

## 3 | RESULTS

### 3.1 | Attrition analyses

We compared the 66 participants who completed the survey and interview to the 25 noncompleters on baseline indicators of ED severity (Supporting Information Table S1). Groups did not differ in age, BMI at start of treatment, percentage with a lifetime history of DSM-IV AN, any of the EDE-Q subscales, or presence of any of the following at baseline: an ED, binge eating, vomiting, laxative use, or fasting ( $ps > .236$ ). Thus, completers and noncompleters were similar in terms of ED severity, minimizing attrition concerns.

### 3.2 | Classification by recovery status

Nineteen of the participants met criteria for an ED: 3 for AN without the amenorrhea requirement, 1 for BN, and 15 for EDNOS (subthreshold AN,  $n = 4$ ; subthreshold BN,  $n = 1$ ; subthreshold BED,  $n = 4$ ; purging disorder,  $n = 6$ ). Twenty-eight met criteria for full recovery (physical, behavioral, and cognitive), and 11 met criteria for partial recovery (physical and behavioral, but not cognitive). Demographically, the four groups (ED, partial recovery, full recovery, controls) did not differ on age (means of 31.79, 34.09, 31.11, and 32.35, respectively) or race (93–100% identified as White in each group;  $ps \geq .464$ ). Eight participants did not fit any of these a priori recovery status groups mainly due to binge eating and/or vomiting in the prior 3 months (e.g., typically 1–2 episodes). Thus, of the 66 participants with a history of an ED, 29% currently had an ED, 17% were partially recovered, 42% were fully recovered, and 12% could not be classified.

### 3.3 | Concurrent validity

Table 1 contains the results of the concurrent validity analyses. For the measure of broad eating pathology (EAT-26), the fully recovered group was statistically indistinguishable from controls and significantly less

disordered than both the partially recovered and ED groups. Furthermore, the mean EAT-26 score of the fully recovered group (6.11) was well below the clinical cut-off of 20 (in contrast to the partially recovered and ED groups). For thinness and restricting expectancies, the fully recovered group was statistically indistinguishable from controls and had significantly lower levels than the partially recovered group.

For body image-focused constructs, findings were mixed. For body shame, the fully recovered group was statistically indistinguishable from controls and had significantly less body shame than both the partially recovered and ED groups. However, for physical investment in appearance (ASI-R) and thin-ideal internalization, the fully recovered group was not statistically different from the partially recovered or ED group. The only instance where the fully recovered group and controls differed was investment in physical appearance.

For the measures of obsessionality, as well as urges, the fully recovered group was statistically indistinguishable from controls and significantly less disordered than the ED group (and for difficulty stopping thoughts and urges, less disordered than the partially recovered group). The partially recovered group was significantly different from the ED group only for urges.

For an alternative analytic approach using hierarchical regression to examine the added value of cognitive recovery that is a criterion of full recovery, see Supporting Information Table S2. In brief, there was a consistent pattern of cognitive recovery accounting for significant variance in the disordered eating variables above and beyond physical and behavioral recovery. Unique variance ranged from 6% (thin ideal internalization) to 31% (body shame, EAT-26), with a mean of an additional 19% of the variance in disordered eating levels explained by cognitive recovery. Thus, this analytic approach provides additional support for cognitive recovery providing “value added” in the conceptualization of recovery.

### 3.4 | Predictive validity

In order to examine how well recovery status at baseline predicted recovery status at follow-up, we focused on women classified into a recovery status group at both baseline and follow-up: 29 women with an ED diagnosis at baseline, 10 in partial recovery at baseline, and 15 in full recovery at baseline, for a total of 54. The change in the distribution across recovery status groups over time was significant (Fisher's Exact Test = 16.09,  $p = .001$ ; Figure 1). Of those with an ED at baseline, a little over half had an ED 7–8 years postbaseline, and about one-third achieved full recovery. Of those in partial recovery at baseline, about the same number had moved to full recovery as had remained in partial recovery (40–50%). Of those in full recovery at baseline, a substantial majority (80%) remained fully recovered 7–8 years later.

## 4 | DISCUSSION

This follow-up study generally replicated findings from earlier work (Bardone-Cone et al., 2010) demonstrating concurrent validity (aim 1) of an operationalization of recovery based on physical, behavioral,

**TABLE 1** Comparison of disordered eating cognitions and behaviors across eating disorder recovery status groups

Measure/ construct	Current ED (n = 19)	Partially recovered (n = 11)	Fully recovered (n = 28)	Controls (n = 31)	Significance	Pair-wise comparisons
EAT-26	20.47 (15.35)	17.55 (11.44)	6.11 (6.11)	3.77 (4.34)	$F(3, 85) = 17.24, p < .001,$ partial $\eta^2 = .38$	C < CED ( $p < .001$ ) C < PRED ( $p < .001$ ) FRED < CED ( $p < .001$ ) FRED < PRED ( $p = .004$ )
TREI	37.21 (13.58)	42.64 (10.42)	30.64 (9.67)	26.26 (10.92)	$F(3, 85) = 7.64, p < .001,$ partial $\eta^2 = .21$	C < CED ( $p = .006$ ) C < PRED ( $p < .001$ ) FRED < PRED ( $p = .017$ )
Body shame	4.36 (1.27)	4.56 (1.42)	2.81 (1.02)	2.10 (0.76)	$F(3, 85) = 26.16, p < .001,$ partial $\eta^2 = .48$	C < CED ( $p < .001$ ) C < PRED ( $p < .001$ ) FRED < CED ( $p < .001$ ) FRED < PRED ( $p < .001$ )
ASI-R	3.84 (0.58)	4.05 (0.54)	3.52 (0.71)	2.93 (0.67)	$F(3, 85) = 12.31, p < .001,$ partial $\eta^2 = .30$	C < CED ( $p < .001$ ) C < PRED ( $p < .001$ ) C < FRED ( $p = .004$ )
Thin-ideal internalization	32.21 (10.12)	32.36 (9.22)	26.89 (11.23)	21.58 (9.25)	$F(3, 85) = 5.69, p = .001,$ partial $\eta^2 = .17$	C < CED ( $p = .003$ ) C < PRED ( $p = .016$ )
Obsession— amount of time	2.68 (1.15)	2.67 (1.13)	1.88 (0.67)	1.66 (0.52)	$F(3, 85) = 8.69, p < .001,$ partial $\eta^2 = .24$	C < CED ( $p < .001$ ) C < PRED ( $p = .004$ ) FRED < CED ( $p = .007$ )
Obsession— difficulty stopping thoughts	3.02 (1.28)	3.00 (1.12)	1.80 (0.93)	1.57 (0.63)	$F(3, 85) = 13.31, p < .001,$ partial $\eta^2 = .32$	C < CED ( $p < .001$ ) C < PRED ( $p < .001$ ) FRED < CED ( $p < .001$ ) FRED < PRED ( $p = .003$ )
Urges	3.49 (1.48)	2.50 (0.99)	1.54 (0.72)	1.16 (0.30)	$F(3, 85) = 31.12, p < .001,$ partial $\eta^2 = .52$	C < CED ( $p < .001$ ) C < PRED ( $p < .001$ ) FRED < CED ( $p < .001$ ) FRED < PRED ( $p = .016$ ) PRED < CED ( $p = .020$ )

Note. Means and (standard deviations) are presented above by group. For body shame, the possible range was 1–7; for each of the obsession measures, 1–5; and for urges, 1–7. In all cases, higher scores reflect greater levels of the constructs. Effect sizes of partial  $\eta^2$  are reported where .01 = small, .06 = medium, and .14 = large (Cohen, 1988). The pairwise comparisons listed are those with significant  $p$ -values after applying the Benjamini–Hochberg procedure (Benjamini & Hochberg, 1995) which controls the false discovery rate that is otherwise inflated with multiple tests.

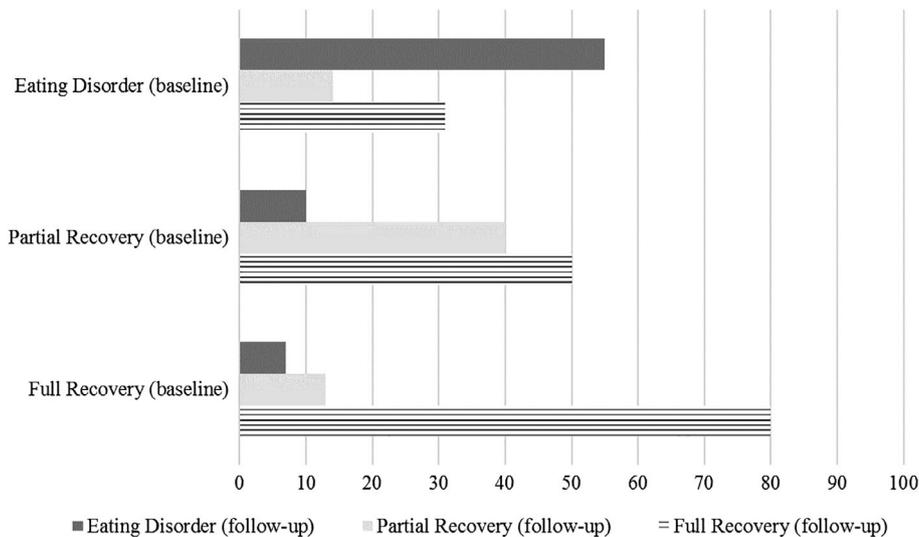
Abbreviations: ASI-R, Appearance Schemas Inventory-Revised (possible range: 1–5); C, controls; CED, current eating disorder; EAT-26, Eating Attitudes Test-26 (possible range: 0–78); ED, eating disorder; FRED, fully recovered; PRED, partially recovered; TREI, Thinness and Restricting Expectancy Inventory (possible range: 8–56).

and cognitive criteria. There was a strong pattern of those in full recovery being statistically indistinguishable from controls on an array of ED constructs, in addition to being less disordered than those with an ED and often those in partial recovery (thus highlighting the added value of cognitive recovery). Interestingly, group comparisons on two constructs that reflect “buying into” the importance of appearance (psychological investment in appearance, thin-ideal internalization) were inconsistent with previous work. Examination of the means for these two constructs at both time points suggests that the different pattern of findings may be due in part to a reduction in the investment in appearance among the ED and partially

recovered groups at follow-up. Interestingly, there is some support for decreased importance placed on appearance with age (e.g., Pliner, Chaiken, & Flett, 1990).

The predictive validity analyses (aim 2) also yielded support for the operationalization of full recovery. Conceptually, a robust model of recovery means that being recovered at one point in time will predict recovery at a later point in time. Indeed, full recovery demonstrated the greatest stability; 80% of those in full recovery at baseline also met full recovery criteria 7–8 years later.

Strengths of the current work include the longitudinal component with the relatively high retention rate and the first examination of the



**FIGURE 1** Predictive validity: change and stability (in percentiles) in eating disorder recovery status over a 7–8 year period from baseline to follow-up

predictive validity of a comprehensive operationalization of recovery. Limitations include the small sample size and limits to generalizability to males and individuals of different races/ethnicities. It is important to acknowledge that the baseline and follow-up data collection are “snapshots,” reflecting discrete time points in recovery. This limits our ability to understand the various trajectories that may have occurred in the interim.

Future work should consider how to best represent the small minority who do not fit any of the recovery status groups, often due to infrequent binge eating or purging. These individuals may fit best with the partial recovery or full recovery group (depending on cognitive recovery status) or perhaps a distinct category is needed characterized by partial (or full) recovery with the rare behavioral “slip.” Another aim for future research is to better understand for whom partial recovery is a transition point between an ED and full recovery and for whom partial recovery may be a final recovery stage. Interestingly, examination of the bidirectional, inverse relationship between quality of life and ED symptoms suggests that targeting quality of life may influence ED symptoms and could be helpful in moving some individuals from partial to full recovery (Mitchison, Morin, Mond, Slewa-Younan, & Hay, 2015).

More longitudinal work is needed, including examinations of trajectories. For example, of the 80% who were recovered at both time points, did some of those individuals regress to partial recovery in between or even relapse? Recovery research with more frequent assessments would paint a more detailed picture of trajectories and predictors of movement across ED recovery groups. Additionally, future work should apply this operationalization of recovery to DSM-5 ED diagnoses, consider including excessive exercise as part of behavioral recovery, and compare this comprehensive model of recovery to other models in order to move toward a consensus definition.

In conclusion, replication of concurrent validity findings supports comprehensive recovery from an ED along physical, behavioral, and cognitive indices as possible, with individuals in full recovery looking like those who have never had an ED on ED measures. A key novel

finding is that attaining full recovery is a robust predictor of full recovery 7–8 years later, suggesting the potential for sustainability.

## ACKNOWLEDGMENTS

This research was supported in part by a University of North Carolina at Chapel Hill (UNC) Research Council Small Grant and a UNC Department of Psychology and Neuroscience Stephenson and Lindquist Award.

## DATA ACCESSIBILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

Anna M. Bardone-Cone  <https://orcid.org/0000-0002-1855-6328>

## REFERENCES

- Ackard, D. M., Richter, S. A., Egan, A. M., & Cronemeyer, C. L. (2014). What does remission tell us about women with eating disorders? Investigating applications of various remission definitions and their associations with quality of life. *Journal of Psychosomatic Research, 76*, 12–18. <https://doi.org/10.1016/j.jpsychores.2013.10.002>
- Bachner-Melman, R., Zohar, A. H., & Ebstein, R. P. (2006). An examination of cognitive versus behavioral components of recovery from anorexia nervosa. *The Journal of Nervous and Mental Disease, 194*, 697–703. <https://doi.org/10.1097/01.nmd.0000235795.51683.99>
- Bardone-Cone, A. M., Harney, M. B., Maldonado, C. R., Lawson, M. A., Robinson, D. P., Smith, R., & Tosh, A. (2010). Defining recovery from an eating disorder: Conceptualization, validation, and examination of psychosocial functioning and psychiatric comorbidity. *Behaviour Research and Therapy, 48*, 194–202. <https://doi.org/10.1016/j.brat.2009.11.001>
- Bardone-Cone, A. M., Hunt, R. A., & Watson, H. J. (2018). An overview of conceptualizations of eating disorder recovery, recent findings, and

- future directions. *Current Psychiatry Reports*, 20, 79. <https://doi.org/10.1007/s11920-018-0932-9>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 57, 289–300.
- Björntorp, P. (2002). Definition and classification of obesity. In C. G. Fairburn & K. D. Brownell (Eds.), *Eating disorders and obesity* (2nd ed., pp. 377–381). New York, NY: Guilford.
- Cash, T. F., Melnyk, S. E., & Hrabosky, J. I. (2004). The assessment of body image investment: An extensive revision of the appearance schemas inventory. *International Journal of Eating Disorders*, 35, 305–316. <https://doi.org/10.1002/eat.10264>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Couturier, J., & Lock, J. (2006). What is recovery in adolescent anorexia nervosa? *International Journal of Eating Disorders*, 39, 550–555. <https://doi.org/10.1002/eat.20309>
- Craig, B. M., & Adams, A. K. (2008). Accuracy of body mass index categories based on self-reported height and weight among women in the United States. *Maternal and Child Health Journal*, 13, 489–496. <https://doi.org/10.1007/s10995-008-0384-7>
- Davis, H. A., Guller, L., & Smith, G. T. (2016). Developmental trajectories of boys' driven exercise and fasting during the middle school years. *Journal of Abnormal Child Psychology*, 44, 1309–1319. <https://doi.org/10.1007/s10802-015-0119-3>
- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16, 363–370. [https://doi.org/10.1002/1098-108X\(199412\)16:4<363::AID-EAT2260160405>3.0.CO;2-#](https://doi.org/10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-#)
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (2002). *Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition*. New York, NY: Biometrics Research, New York State Psychiatric Institute.
- Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The eating attitudes test: Psychometric features and clinical correlates. *Psychological Medicine*, 12, 871–878. <https://doi.org/10.1017/S0033291700049163>
- Hohlstein, L. A., Smith, G. T., & Atlas, J. G. (1998). An application of expectancy theory to eating disorders: Development and validation of measures of eating and dieting expectancies. *Psychological Assessment*, 10, 49–58. <https://doi.org/10.1037/1040-3590.10.1.49>
- Landis, J. T., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- McCabe, R. E., McFarlane, T., Polivy, J., & Olmsted, M. P. (2001). Eating disorders, dieting, and the accuracy of self-reported weight. *International Journal of Eating Disorders*, 29, 59–64.
- McKinley, N. M., & Hyde, J. S. (1996). The objectified body consciousness scale development and validation. *Psychology of Women Quarterly*, 20, 181–215. <https://doi.org/10.1111/j.1471-6402.1996.tb00467.x>
- Mitchison, D., Morin, A., Mond, J., Slewa-Younan, S., & Hay, P. (2015). The bidirectional relationship between quality of life and eating disorder symptoms: A 9-year community-based study of Australian women. *PLoS One*, 10, e0120591.
- Mond, J. M., Hay, P. J., Rodgers, B., & Owen, C. (2006). Eating disorder examination questionnaire (EDE-Q): Norms for young adult women. *Behaviour Research and Therapy*, 44, 53–62. <https://doi.org/10.1016/j.brat.2004.12.003>
- Pliner, P., Chaiken, S., & Flett, G. (1990). Gender differences in concern with body weight and physical appearance over the life span. *Personality and Social Psychology Bulletin*, 16, 263–273. <https://doi.org/10.1177/0146167290162007>
- Thissen, D., Steinberg, L., & Kuang, D. (2002). Quick and easy implementation of the Benjamini-Hochberg procedure for controlling the false positive rate in multiple comparisons. *Journal of Educational and Behavioral Statistics*, 27, 77–83.
- Thompson, J. K., van den Berg, P., Roehrig, M., Guarda, A. S., & Heinberg, L. J. (2004). The Sociocultural Attitudes Towards Appearance Scale-3 (SATAQ-3): Development and validation. *International Journal of Eating Disorders*, 35, 293–304. <https://doi.org/10.1002/eat.10257>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Bardone-Cone AM, Alvarez A, Gorlick J, Koller KA, Thompson KA, Miller AJ. Longitudinal follow-up of a comprehensive operationalization of eating disorder recovery: Concurrent and predictive validity. *Int J Eat Disord*. 2019;52:1052–1057. <https://doi.org/10.1002/eat.23128>