

# A Comparison of the Binge Eating Scale, Questionnaire for Eating and Weight Patterns-Revised, and Eating Disorder Examination Questionnaire with Instructions with the Eating Disorder Examination in the Assessment of Binge Eating Disorder and its Symptoms

Angela A. Celio,<sup>1,2</sup> Denise E. Wilfley,<sup>2,3\*</sup> Scott J. Crow,<sup>4</sup> James Mitchell,<sup>5</sup>  
and B. Timothy Walsh<sup>6</sup>

<sup>1</sup>Joint Doctoral Program in Clinical Psychology, San Diego State University/  
University of California at San Diego, San Diego, California

<sup>2</sup>Center for Eating and Weight Disorders, San Diego, California

<sup>3</sup>Department of Psychiatry, Washington University School of Medicine,  
St. Louis, Missouri

<sup>4</sup>Department of Psychiatry, University of Minnesota, Minneapolis, Minnesota

<sup>5</sup>Department of Neuroscience, University of North Dakota School of Medicine  
and Health Sciences, North Dakota

<sup>6</sup>New York State Psychiatric Institute, Columbia-Presbyterian Medical Center,  
New York, New York

Accepted 11 March 2004

**Abstract: Objective:** *The current study assesses concordance between self-administered measures and a diagnostic standard for assessment of binge frequency and diagnosis of binge eating disorder (BED) in a sample of binge eaters. Method:* The Questionnaire for Eating and Weight Patterns-Revised (QEWP-R), Binge Eating Scale (BES), two items from the Eating Disorder Examination Questionnaire with Instructions (EDE-Q-I), and the Eating Disorder Examination (EDE) were administered. Participants were 157 adults volunteering for a clinical study, of whom 129 (79%) were diagnosed with BED using the EDE as the diagnostic standard. **Results:** *In the identification of BED, the QEWP-R yielded a sensitivity value of .74 and a specificity value of .35. The BES yielded a sensitivity value of .85 and a specificity value of .20. Frequency of binge eating days and episodes on the EDE-Q-I correlated highly with the EDE (.65 and .48, respectively;  $p < .001$ ). Discussion:* *The accuracy of diagnosis and symptomatology among*

---

Supported by Abbott Laboratories.

Portions of this manuscript were presented at the annual meeting of the Academy of Eating Disorders in Denver, CO, May 29–31, 2003.

\*Correspondence to: Denise E. Wilfley, PhD, Department of Psychiatry, Washington University School of Medicine, 660 South Euclid, Campus Box 8134, St. Louis, MO 63110. E-mail: wilfleyd@msnotes.wustl.edu

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/eat.20057

© 2004 by Wiley Periodicals, Inc.

*self-administered questionnaires is variable. The BES and the QEWP-R performed satisfactorily as initial screens for the diagnosis of BED, but were less accurate in identifying non-BED individuals and the frequency of binge eating. The EDE-Q-I most accurately assessed the frequency of binge eating.* © 2004 by Wiley Periodicals, Inc. *Int J Eat Disord* 36: 434–444, 2004.

**Key words:** *binge eating assessments; binge frequency; binge eating disorder*

## INTRODUCTION

The accurate assessment of binge eating disorder (BED) and the hallmark feature of this disorder, binge eating, are essential to determining the course and treatment of individuals with BED. Researchers have turned to self-administered measures of BED and its associated behavioral features due to the cost and participant burden of interview-based methods. Unfortunately, this overreliance on self-administered measures of overeating is problematic because of low convergence with diagnostic standards, such as the interview-based Eating Disorder Examination (EDE; Fairburn & Cooper, 1993). In addition, commonly used self-report assessments fail to assess binge eating frequency, an essential component of diagnosis and a measure of treatment outcome. As a result, concern has been raised regarding the usefulness of these measures as screening and diagnostic tools and the ways that in which their widespread use may limit our understanding of BED (Wilfley, Wilson, & Agrad, 2003).

In the existing literature, three self-administered measures have been compared with a clinical interview in samples of individuals who experience recurrent binge eating without regular compensatory behaviors: the Binge Eating Scale (BES), the Questionnaire on Eating and Weight Patterns-Revised (QEWP-R), and the Eating Disorder Examination Questionnaire (EDE-Q). The BES (Gormally, Black, Daston, & Rardin, 1982) assesses binge eating severity on a continuous scale with scores  $>27$  indicating severe binge eating. In a study with treatment-seeking obese individuals, concordance between the BES and the EDE differed significantly by the degree of symptomatology (Greeno, Marcus, & Wing, 1995). The BES correctly identified 93.5% of binge eaters (cutoff value  $\geq 27$ ), but correctly identified only 49.4% of non-binge eaters (cutoff value  $\leq 17$ ). Therefore, severe binge eating was attributed to approximately one half of the individuals who were identified by the EDE as not meeting the binge eating severity criterion for a BED diagnosis. Participants who scored between 17 and 27 were excluded from the analysis, even though they endorsed a moderate level of binge eating on the BES. In another study using an unvalidated clinical interview (i.e., developed for the study and supplemented by questions based on the QEWP-R), agreement between the BES and the interview in identifying binge eaters among treatment-seeking obese individuals was only fair ( $\kappa = .31-.50$ ; Brody, Walsh, & Devlin, 1994). In the Brody et al. study, the BES correctly identified 38.5% of BED individuals using a cutoff value of  $\geq 27$  and 98.2% of non-BED individuals using a cutoff value  $< 27$ .

The QEWP-R (Spitzer et al., 1992), which obtains binge eating frequency as well as other criteria for the diagnosis of BED, has not been compared with the EDE in the assessment of BED, but has been compared with clinical interviews based on the criteria for BED in the 4th ed. of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association [APA], 1994). When comparing results from the self-administered QEWP-R (Spitzer, Yanovski, & Marcus, 1993) and a clinical interview based on this same measure, adequate agreement was obtained ( $\kappa = .60$ ; Spitzer,

Yanovski, Wadden, et al., 1993). In the de Zwaan et al. (1993) study of the QEWP-R compared with a modified structured clinical interview for DSM-III-R (SCID), 77% of individuals with BED and 80% of individuals without BED were correctly identified. Nevertheless, nearly one fourth of individuals were misdiagnosed using the QEWP-R.

Comparisons between commonly used self-administered assessments highlight disparities that may subsequently make an understanding of the nature of binge eating problematic. An evaluation of the BES versus the QEWP-R, which was supplemented with clarification questions by an interviewer, revealed only moderate agreement ( $\kappa = .45-.64$ ; Gladis, Wadden, Foster, Vogt, & Wingate, 1998). When classifying respondents as binge eaters or non-binge eaters, 50% of individuals identified as having BED on the QEWP-R had BES scores categorizing them as moderate or non-binge eaters. Gladis et al. explain the discordance between the two measures as resulting from the differing foci of the measures—the BES has a greater focus on psychopathology associated with binge eating, whereas the QEWP-R specifically assesses criteria necessary to make the diagnosis of BED.

In addition to the diagnosis of BED, an important focus of self-administered assessments is the frequency of binge eating, the key feature of BED. The EDE-Q (Fairburn & Beglin, 1994) does not assess all diagnostic features of BED or provide information beyond the past 28 days, but it does obtain information regarding binge eating frequency. The EDE-Q has been compared in several studies with the EDE and these comparisons have yielded generally positive results. Several studies report statistically significant correlations between binge episode frequencies on the two measures (Pearson's  $r$  or Kendall's  $\tau$ -b = .28-.53,  $p < .01$ ). Further, differences between the reported frequency of binge episodes on the EDE versus the EDE-Q are not statistically significant, suggesting that the estimates provided by respondents on the EDE-Q closely approximate those obtained via an expert-rated interview (Black & Wilson, 1996; Grilo, Masheb, & Wilson, 2001a, 2001b; Kalarchian, Wilson, Brolin, & Bradley, 2000). In another study, an examination of a clinical and a community sample revealed statistically significant discrepancies on reports of binge episodes using these measures, particularly in the community sample, but correlations between the estimates were also evident (Kendall's  $\tau$ -b = .45-.60,  $p < .001$ ; Fairburn & Beglin, 1994). Wilson (1993) suggested assessing binge eating days in addition to episodes for BED patients due to the lack of discrete binge eating episodes often characterizing individuals with BED. Days may be easier to remember and quantify, thus providing a more accurate measure of frequency. However, an initial study of BED patients yielded a nonsignificant correlation between binge days on the EDE-Q and the EDE (Kendall's  $\tau$ -b = .20,  $p < .06$ ; Wilfley, Schwartz, Spurrell, & Fairburn, 1997).

The studies that have reported significant differences between interviews and questionnaires on binge frequencies attribute the difference, in part, to the lack of a shared definition of bingeing between the investigator and the respondent on the self-administered measure. Wilfley et al. (1997) suggested the development of a set of instructions with detailed definitions of a large amount of food and loss of control to aid the respondent while answering the questions on the EDE-Q. A preliminary study of 84 individuals with BED comparing two versions of the EDE-Q—one with instructions (EDE-Q-I) and one without instructions—with the EDE demonstrated that the addition of instructions improves the concordance of mean binge frequencies between the self-administered questionnaire and clinical interview (EDE-Q-I:  $\tau$ -b = .467,  $p = .000$ ; EDE-Q:  $\tau$ -b = .244,  $p = .102$ ; Goldfein et al., 2002). These findings suggest that this modest modification to the EDE-Q can aid in the accurate assessment of recurrent binge eating. Additional studies are needed to assess the utility of the EDE-Q-I in larger samples and in samples that include individuals with subdiagnostic frequencies of binge eating.

To our knowledge, no studies to date have attempted to compare multiple self-administered measures of binge eating and BED within one sample or to assess the level of agreement between the EDE and these self-administered measures on diagnostic criteria and for the overall diagnosis of BED. As evidenced by the studies mentioned, there is some concern regarding the diagnostic utility of existing self-administered assessments of binge eating and BED. There is a need to establish the accuracy of these self-administered assessments, as compared with diagnostic interviews, to determine the reliability of past findings of studies on BED. The concomitant evaluation of multiple self-administered assessments also has important implications for the assessment of binge eating, the hallmark feature of BED. An extension of the Goldfein et al. (2002) evaluation of the EDE-Q-I would enhance the existing literature as well.

The current study utilized the QEWP-R, the BES, the overeating section from the EDE-Q-I, and a version of the EDE specifically developed for the diagnosis of BED (Fairburn & Cooper, 2000). The aims of the current study were twofold. First, we assessed the level of agreement between the self-administered and expert rating for the overall diagnosis of BED in a large sample of binge eaters. Second, we evaluated the level of agreement between the self-administered and interview-based methods in the assessment of the frequency and severity of binge eating.

## METHOD

### Participants

Participants included 157 adults (86.6% female) selected from the initial screening phase of a multisite, randomized, controlled trial that studied the effectiveness of sibutramine in the treatment of BED (Knoll Pharmaceutical Company, 1999, now Abbott Laboratories). Three of the 15 sites (located in the Northeast, Midwest, and West Coast) participated in this substudy. The mean age of the participants was 41.9 years ( $SD = 11.3$ ). In addition, and 70.3% of the participants were White, 11.0% were Hispanic, 6.5% were Black, 4.5% were Asian, and 7.6% were of other ethnicity. Most of the sample reported having at least some college education (87.9%). Body mass index (BMI) was calculated using self-report data and ranged from 21.6 to 52.7, with a mean of 33.6 ( $SD = 5.3$ ). Study participants were self-referred and were initially screened over the phone to exclude non-binge eaters and minimal-frequency binge eaters (i.e., in the past 28 days, participants binged less than once per week on average or less than four times total per month; in the past 6 months, participants binged less than twice per week or less than eight times per month on average). Individuals gave their informed consent in-person before completing a battery of self-administered questionnaires, the EDE, and a physical examination. After these assessments, inclusion was determined for the pharmaceutical trial.

### Procedures

The EDE, QEWP-R, BES, and binge frequency items from the EDE-Q-I were administered on the same day to ensure overlapping time periods. The order of the assessments differed by site, with one site administering the EDE before the self-administered questionnaires ( $n = 20$ ) and two sites administering the EDE last (combined  $n = 137$ ). The data were reanalyzed without the 20 participants who received the EDE first and the results remained unchanged. Interviewers were blind to respondents' answers on the self-administered instruments.

A diagnosis of BED was made based on responses to the EDE. Standardized training on the EDE occurred before the initiation of the study and included a training seminar with an author of the interview, Dr. Fairburn, coding of at least two audiotapes, co-rating at least two EDEs administered by expert interviewers, and observation of two EDEs by an experienced interviewer. Interviewers were closely monitored until they were able to administer and code the interview reliably and accurately. Ongoing monitoring occurred during the study to ensure the continued high quality of interviewer ratings.

## Measures

### EDE-BED Diagnostic Version

The BED diagnostic version of the EDE is an investigator-based, semistructured clinical interview designed to provide a comprehensive assessment of the specific psychopathology of BED. It was derived from the 12th ed. of the EDE (Fairburn & Cooper, 1993), which is considered to be the gold standard of eating disorder assessment. Psychometric studies of the EDE have described good internal consistency (Cooper, Cooper, & Fairburn, 1989), discriminative validity (Cooper et al., 1989; Rosen, Vara, Wendt, & Leitenberg, 1990; Wilson & Smith, 1989), concurrent validity (Rosen et al., 1990), and test-retest reliability (Rizvi, Peterson, Crow, & Agras, 2000). Among patients with BED, the EDE has also demonstrated good interrater reliability and test-retest reliability (Grilo, Masheb, Lozano-Blanco, & Barry, 2004). The EDE has been updated to include items necessary for the diagnosis of BED. Diagnostic behavioral items assess overeating episodes (including objective bulimic episodes [OBE], subjective bulimic episodes [SBE], and objective overeating), features associated with binge eating (e.g., eating much more rapidly than normal), self-induced vomiting, laxative misuse, diuretic misuse, dietary restriction outside of bulimic episodes, and intense exercising to control shape or weight. The frequencies (episodes and days) of these behaviors for the past 6 months are recorded. Attitudinal items consist of Importance of Weight and Importance of Shape, which are scored in terms of their severity (0 = no importance to 6 = supreme importance—nothing is more important in the subject's scheme for self-evaluation) for the most recent 3 months. The EDE requires intensive training of interviewers and typically takes 60 min to administer.

### QEWPR

The QEWPR (Spitzer, Yanovski, & Marcus, 1993) is a self-administered measure designed to assess the presence or absence of binge episodes, the frequency of such episodes, and additional required features for the diagnosis of BED criteria, as defined in DSM-IV (APA, 1994). The test-retest reliability of the QEWPR in the diagnosis of BED within a sample of self-referred binge eaters and a comparison sample was adequate ( $\kappa = .58$ ; Nangle, Johnson, Carr-Nangle, & Engler, 1994).

### BES

The BES (Gormally et al., 1982) was originally developed to identify binge eaters within an obese population. It does not specify a time frame and presents a series of differently weighted statements for each item, from which respondents select the statement that best describes their attitudes and behaviors. This yields a continuous measure of binge eating pathology of 0–46. Scores of  $\geq 27$  have conventionally served as a cutoff value for identifying the presence of severe binge eating and  $\leq 17$  as a cutoff value for mild or no binge eating (Greeno et al., 1995). The BES has good test-retest reliability ( $r = .87, p < .001$ ) and moderate associations with binge eating severity as measured by food records ( $r = .20-.40, p < .05$ ; Timmerman, 1999).

### EDE-Q-I

The EDE-Q (Fairburn & Beglin, 1994) was developed as a self-administered version of the EDE and takes approximately 15 min to complete. The EDE-Q cannot be used for the diagnosis of BED because it does not specifically assess the full range of diagnostic criteria over a 6-month period. For the current study, a modified version of the EDE-Q-I (Goldfein et al., 2002) was used, which provides definitions of “unusually large amount of food” and “sense of loss of control.” Only two items of the EDE-Q-I were administered: number of days on which participants reported having OBE in the past 4 weeks and number of OBE during the same period.

### Statistical Analysis

Pearson chi-square and *t* tests were used to assess site differences. For the evaluation of BED diagnoses, the EDE served as the diagnostic standard by which the self-administered assessments were compared. Sensitivity, defined as the probability of a positive test among patients with a disease, and specificity, defined as the probability of a negative test among patients without a disease, were calculated for the BES and the QEWP-R. For the comparisons of binge frequencies using continuous measures (i.e., EDE-Q-I and BES), Pearson’s correlations and Student’s paired *t* tests were calculated. To compare the average number of binge days per week over the past 6 months on the EDE and the QEWP-R, EDE responses were grouped to have a scale equivalent to the QEWP-R (i.e., less than once per week, once per week, two or three times per week, four to five times per week, nearly every day) and Kendall’s  $\tau$ -b for ordinal data was calculated.

The BES does not assess the amount of food consumed in an episode. Therefore, the BES may be measuring eating episodes characterized simply by feelings of loss of control. As a result, EDE reports of binge eating and SBE (i.e., a normal amount of food with loss of control) were combined to assess the association with the BES. They were also examined individually. EDE reports of objective overeating episodes (i.e., a large amount of food without a loss of control) were also compared with the BES.

## RESULTS

No differences were found between sites on age, sex, self-report BMI, or level of education. Sites differed slightly by ethnic distribution, with the Midwest site having a larger percentage of Whites than the other two sites,  $\chi^2(8) = 17.3, p < .05$ .

### Comparison 1: Diagnosis of BED

The sample included 129 (79%) individuals diagnosed with BED using the EDE. In comparison, the QEWP-R identified 111 (70.7%) participants as having BED and the BES identified 130 participants (82.8%) who scored  $>27$ , indicating serious binge eating.

Using the EDE as the diagnostic standard, the QEWP-R yielded a sensitivity value of .74 and a specificity value of .35 (Table 1). Using a score of  $\geq 27$  to indicate serious binge eating, the BES yielded a sensitivity value of .85 and a specificity value of .20 in identifying individuals with severe binge eating (Table 2).

Table 1. QEWP-R vs. EDE diagnoses ( $N = 143$ )

QEWP-R Diagnosis	EDE Diagnosis		Total
	BED	Non-BED	
BED	91	13	104
Non-BED	32	7	39
Total	123	20	143

Note: Of the total sample, 143 participants provided adequate information to obtain diagnoses on both assessments. Sensitivity is calculated by dividing the number of true positives on the self-administered assessment by the number of individuals with BED. Specificity is calculated by dividing the number of true negatives on the self-administered assessment by the number of individuals without BED. QEWP-R = Questionnaire for Eating and Weight Patterns-Revised; EDE = Eating Disorder Examination; BED = binge eating disorder.

### Comparison 2: Binge Frequency

To assess the level of agreement between self-administered and interview-based methods in the assessment of the frequency of binge eating, the frequencies of OBE days and episodes for the past 4 weeks as reported on the EDE were compared with the EDE-Q-I and the QEWP-R. The EDE-Q-I showed slightly higher reports of binge eating, but these were not significantly different than the frequencies obtained with the EDE (Table 3).

The QEWP-R correctly identified the frequency of binge days over the past 6 months for 63.2% of participants and the correlation with the EDE was good ( $r = .53, p < .001$ ). Data were also dichotomized into high and low binge frequency (i.e., two or more binges per week vs. less than 2 days per week on average for the past 6 months) for both measures. This comparison revealed that 95.3% of individuals with a high binge frequency were correctly identified by the QEWP-R. None of the three individuals identified as low frequency binge eaters were positively identified by the QEWP-R.

Correlations between the BES as a continuous measure of severity of binge eating and frequency of binge eating as measured by the EDE (i.e., OBES) were also calculated

Table 2. BES vs. EDE diagnosis ( $N = 141$ )

BES Score	EDE Diagnosis		Total
	BED	Non-BED	
27+	103	16	119
<27	18	4	22
Total	121	20	141

Note: Of the total sample, 141 participants provided adequate information to obtain diagnoses on both assessments. Sensitivity is calculated by dividing the number of true positives on the self-administered assessment by the number of individuals with BED. Specificity is calculated by dividing the number of true negatives on the self-administered assessment by the number of individuals without BED. BES = Binge Eating Scale; EDE = Eating Disorder Examination; BED = binge eating disorder.

Table 3. OBE days and episodes in the past month

	EDE	EDE-Q-I	Correlation	Paired-Sample <i>t</i> Test	95% CI
OBE days ( <i>n</i> = 154) <sup>a</sup>	16.8 ± 8.0	17.1 ± 7.0	0.65*	<i>t</i> (153) = -.584, <i>p</i> = .56	-1.31, .71
OBE episodes ( <i>n</i> = 153) <sup>a</sup>	23.0 ± 15.4	25.6 ± 19.7	0.48*	<i>t</i> (152) = -1.771, <i>p</i> = .08	-5.54, .30

Note: OBE = objective bulimic episodes; EDE = Eating Disorder Examination; EDE-Q-I = Eating Disorder Examination Questionnaire with Instructions.

<sup>a</sup>154 and 153 of the total sample of participants provided information for items on both assessments.

\**p* < .001.

(Table 4). The BES was correlated with frequency of OBE and days (*p* < .01), but no other type of overeating experience, including SBE (i.e., normal amount of food plus feeling of loss of control), objective overeating episodes (i.e., a large amount of food without a loss of control), or OBE and SBE combined.

## DISCUSSION

The current study compared results from three self-administered assessments for binge eating and BED with the current diagnostic standard, the EDE. The QEWP-R and the BES provided reasonable sensitivity in identifying individuals with BED (.74 and .85, respectively) and considerably less specificity in identifying non-BED individuals (.35 and .20, respectively). Although the QEWP-R and BES may be adequate for screening purposes, sole reliance on these measures for the diagnosis of BED is not recommended due to the high percentage of misclassifications.

Consistent with other studies using the EDE-Q, the EDE-Q-I minimally overestimated binge frequency as compared with the EDE. However, of the instruments examined, this measure had the best agreement with the EDE, supporting the Goldfein et al. (2000) findings and demonstrating the relative strength of the EDE-Q-I in the domain of assessing binge eating frequency. It is important to note that the language of the EDE-Q-I is based on the EDE. Therefore, agreement between these two measures will logically be greater than that of the other measures.

The frequency of binge eating days as reported on the EDE-Q-I and the EDE was more highly associated than episodes of binge eating, which is consistent with the observation that BED patients may recall binge days more accurately than they do binge episodes (Rossiter, Agras, Telch, & Bruce, 1992; Wilson, 1993). The binge eating episodes of BED

Table 4. Binge frequency: BES vs. EDE

	Episodes	Days
OBE	<i>r</i> = .24, <i>p</i> = .003	<i>r</i> = .24, <i>p</i> = .004
OBE+SBE	<i>r</i> = .14, <i>p</i> = .10	<i>r</i> = .11, <i>p</i> = .20
SBE	<i>r</i> = .11, <i>p</i> = .18	<i>r</i> = .09, <i>p</i> = .29
OO	<i>r</i> = .01, <i>p</i> = .89	<i>r</i> = .03, <i>p</i> = .72

Note: OBE = objective bulimic episode; SBE = subjective bulimic episode; OO = objective overeating; BES = Binge Eating Scale; EDE = Eating Disorder Examination.



patients are believed to be more difficult to recall because, unlike bulimic patients, the eating episodes of BED patients are not punctuated by vomiting or other compensatory behaviors and their eating behaviors tend to be more chaotic (Marcus, 1997). Using days as the unit of measurement in self-administered measures of binge eating is likely to yield more accurate data.

A key criterion for the diagnosis of BED is the average number of binge days per week over the past 6 months. The overall agreement between the QEWP-R and the EDE on this item was only 63.2% and none of the low-frequency binge eaters were positively identified by the QEWP-R. However, when diagnostic threshold levels of binge eating were examined (i.e., two or more binge days per week), agreement increased to 95.3%. This suggests that the QEWP-R is quite effective in identifying individuals who binge eat at levels severe enough to warrant a diagnosis of BED. As a result, the QEWP-R might be particularly useful in screening for individuals with severe binge eating.

Frequency of binge eating as measured by the BES evidenced a modest association with the EDE only when using OBE (i.e., a large amount of food with a loss of control), with greater accuracy found in the positive identification of individuals with BED (i.e., sensitivity) than those without BED (i.e., specificity). Our findings were inconsistent with those of past investigators (Brody et al., 1994), who reported lower accuracy of the BES in identifying recurrent binge eaters (.39) and greater accuracy in identifying non-binge eaters (.98). The findings of Greeno et al. (1995) showed greater sensitivity and specificity compared with the current study (.94 and .49, respectively). One reason for this could be the different samples used (e.g., individuals seeking treatment for weight loss vs. individuals seeking treatment for BED), which directly affect calculations of sensitivity and specificity, or the different cutoff values used on the BES for determining binge eaters (e.g., 17 vs. 27). Discrepancies with past findings may also be due to the broader focus of the BES. The global scale of binge eating severity on the BES assesses the presence of behaviors and attitudes associated with recurrent binge eating, but may not adequately classify individuals as having or not having BED.

One of the limitations of the current study is the order of administration of the EDE and the self-administered measures. However, we conducted the analyses separately with and without the individuals who completed the EDE first and the results were not altered. Another limitation is that the study was conducted with individuals presenting for treatment of binge eating. This has a direct impact on calculations of specificity and sensitivity, such that inclusion of greater numbers of non-binge eaters would increase specificity, thus improving the overall assessment profile. It is important to be able to discriminate between BED and non-BED individuals within samples of binge eaters such as this. An additional community-based study testing the screening and diagnostic strengths of these measures would aid researchers working with more diverse populations.

In the current study, we compared frequently used assessments of binge eating and BED. It appears that the BES and the QEWP-R perform satisfactorily as initial screens for the diagnosis of BED, but are less accurate in identifying non-BED individuals and the frequency of binge eating. An exception to this is the QEWP-R's identification of high-frequency binge eating over a 6-month period. The EDE-Q-I most capably assessed the frequency of binge eating. Researchers must continue to strive to utilize measures that provide the most accurate assessment of the behavior or attitudes of interest and minimize the burden on both the participant and investigator. Of equal importance, the use of common assessments of binge eating and BED is essential during this period of formative research. Thoughtful consideration of measures is likely to improve our ability to describe, understand, and treat individuals who struggle with binge eating and BED.

## REFERENCES

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Black, C.M.D., & Wilson, G.T. (1996). Assessment of eating disorders: Interview versus questionnaire. *International Journal of Eating Disorders*, 20(1), 43–50.
- Brody, M.L., Walsh, B.T., & Devlin, M.J. (1994). Binge eating disorder: Reliability and validity of a new diagnostic category. *Journal of Consulting and Clinical Psychology*, 62(2), 381–386.
- Cooper, Z., Cooper, P., & Fairburn, C. (1989). The validity of the Eating Disorder Examination and its subscales. *British Journal of Psychiatry*, 154, 807–812.
- de Zwaan, M., Mitchell, J.E., Specker, S.M., Pyle, R.L., Mussell, M.P., & Seim, H.C. (1993). Diagnosing binge eating disorder: Level of agreement between self-report and expert-rating. *International Journal of Eating Disorders*, 14(3), 289–295.
- Fairburn, C., & Beglin, S. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16, 363–370.
- Fairburn, C.G., & Cooper, Z. (1993). *The Eating Disorder Examination* (12th ed.). In C.G. Fairburn & G.T. Wilson (Eds.), *Binge eating: Nature, assessment, and treatment* (pp. 317–360). New York, NY: Guilford Press.
- Fairburn, C.G., & Cooper, Z. (2000). Interview for the diagnosis of binge eating disorder. Unpublished interview.
- Gladis, M.M., Wadden, T.A., Foster, G.D., Vogt, R.A., & Wingate, B.J. (1998). A comparison of two approaches to the assessment of binge eating in obesity. *International Journal of Eating Disorders*, 23, 17–26.
- Goldfein, J., Devlin, M., Kamenetz, C., Wolk, S., Raizman, P., & Dobrow, I. (2002, May). EDE-Q with and without instruction to assess binge eating in patients with binge eating disorder. Poster presented at the annual meeting of the Academy of Eating Disorders, Boston.
- Gormally, J., Black, S., Daston, S., & Rardin, D. (1982). The assessment of binge eating severity among obese persons. *Addictive Behaviors*, 7(1), 47–55.
- Greeno, C.G., Marcus, M.D., & Wing, R.R. (1995). Diagnosis of binge eating disorder: Discrepancies between a questionnaire and clinical interview. *International Journal of Eating Disorders*, 17(2), 153–160.
- Grilo, C.M., Masheb, R.M., Lozano-Blanco, C., & Barry, D.T. (2004). Reliability of the Eating Disorder Examination in patients with binge eating disorder. *International Journal of Eating Disorders*, 35, 80–85.
- 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., and 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.
- Kalarchian, M.A., Wilson, G.T., Brolin, R.E., & Bradley, L. (2000). Assessment of eating disorders in bariatric surgery candidates: Self-report questionnaire versus interview. *International Journal of Eating Disorders*, 28, 465–469.
- Knoll Pharmaceutical Company. (1999). A multicenter, double blind, randomized, placebo controlled, parallel group study to examine the efficacy and safety of Meridia—sibutramine hydrochloride— in binge eating disorder [Clinical protocol]. Mount Olive, NJ.
- Marcus, M.D. (1997). Adapting treatment for patients with binge eating disorder. In D.M. Garner & P. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 484–493). New York, NY: Guilford Press.
- Nangle, D.W., Johnson, W.G., Carr-Nangle, R.E., & Engler, L.B. (1994). Binge eating disorder and the proposed DSM-IV criteria: Psychometric analysis of the questionnaire of eating and weight patterns. *International Journal of Eating Disorders*, 16(2), 147–157.
- Rizvi, S.L., Peterson, C.B., Crow, S.J., & Agras, W.S. (2000). Test-retest reliability of the Eating Disorder Examination. *International Journal of Eating Disorders*, 28, 311–316.
- Rosen, J., Vara, L., Wendt, S., & Leitenberg, H. (1990). Validity studies of the Eating Disorder Examination. *International Journal of Eating Disorders*, 9, 519–528.
- Rossiter, E., Agras, S., Telch, C., & Bruce, B. (1992). The eating patterns of nonpurging bulimic subjects. *International Journal of Eating Disorders*, 11, 111–120.
- Spitzer, R.L., Devlin, M., Walsh, B.T., Hasin, D., Wing, R., Marcus, M., Stunkard, A., Wadden, T., Yanovski, S., Agras, S., Mitchell, J., & Nonas, C. (1992). Binge eating disorder: A multisite field trial of the diagnostic criteria. *International Journal of Eating Disorders*, 11, 191–203.
- Spitzer, R.L., Yanovski, S., Wadden, T., Wing, R., Marcus, M.D., Stunkard, A., Devlin, M., Mitchell, J., Hasin, D., & Horne, R.L. (1993). Binge eating disorder: Its further validation in a multisite study. *International Journal of Eating Disorders*, 13(2), 137–153.
- Spitzer, R.L., Yanovski, S.Z., & Marcus, M.D. (1993). *The Questionnaire on Eating and Weight Patterns-Revised (QEWP-R, 1993)*. (Available from the New York State Psychiatric Institute, 722 West 168th Street, New York, NY 10032).
- Timmerman, G.M. (1999). Binge Eating Scale: Further assessment of validity and reliability. *Journal of Applied Biobehavioral Research*, 4(1), 1–12.

- 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? *Behavior Research and Therapy*, *35*, 1151–1159.
- Wilfley, D.E., Wilson, G.T., & Agras, W.S. (2003). The clinical significance of binge eating disorder. *International Journal of Eating Disorders*, *34*, S96–S106.
- Wilson, G.T. (1993). Assessment of binge eating. In C.G. Fairburn & G.T. Wilson (Eds.), *Binge eating: Nature, assessment, and treatment* (pp. 227–249). New York, NY: Guilford Press.
- Wilson, G.T., & Smith, D. (1989). Assessment of bulimia nervosa: An evaluation of the Eating Disorders Examination. *International Journal of Eating Disorders*, *8*, 173–179.