

## Eating Disorder Diagnostic Scale: Additional Evidence of Reliability and Validity

Eric Stice, Melissa Fisher, and Erin Martinez  
University of Texas at Austin

The authors conducted 4 studies investigating the reliability and validity of the Eating Disorder Diagnostic Scale (EDDS; E. Stice, C. F. Telch, & S. L. Rizvi, 2000), a brief self-report measure for diagnosing anorexia nervosa, bulimia nervosa, and binge eating disorder. Study 1 found that the EDDS showed criterion validity with interview-based diagnoses, convergent validity with risk factors for eating pathology, and internal consistency. Studies 2 and 3 found that the EDDS was sufficiently sensitive to detect the effects of eating disorder prevention programs. Regarding predictive validity, Studies 3 and 4 found that the EDDS predicted response to a prevention program and future onset of eating pathology and depression. Results provide additional evidence of the reliability and validity of this scale and suggest it may be useful in clinical and research applications.

Eating disorders are a common psychiatric problem faced by adolescent girls and young women and are marked by chronicity, relapse, and functional impairment (Fairburn, Cooper, Doll, Norman, & O'Connor, 2000; Lewinsohn, Striegel-Moore, & Seeley, 2000). Eating disorders are also associated with serious medical complications and high rates of inpatient hospitalization, suicide attempts, and mortality (Newman et al., 1996; Wilson, Heffernan, & Black, 1996). In addition, eating pathology increases the risk for future onset of obesity, depressive disorders, suicide attempts, anxiety disorders, substance abuse, and health problems (Johnson, Cohen, Kasen, & Brook, 2002; Stice, Cameron, Killen, Hayward, & Taylor, 1999; Stice, Hayward, Cameron, Killen, & Taylor, 2000).

Research on the etiology, prevention, and treatment of eating disorders has made remarkable advances, but progress has been hampered by the absence of a brief self-report diagnostic measure of eating disorders. Although structured interviews for ascertaining the *Diagnostic and Statistical Manual of Mental Disorders*' (4th ed.; *DSM-IV*; American Psychiatric Association, 1994) diagnoses of anorexia nervosa, bulimia nervosa, and binge eating disorder exist (e.g., Eating Disorder Examination [EDE; Fairburn & Cooper, 1993] and Structured Clinical Interview for *DSM-III-R* [SCID; Spitzer, Williams, Gibbon, & First, 1990]), most self-report scales do not assess the diagnostic criteria for these disorders. Scales such as the Eating Attitudes Test (EAT; Garner, Olmsted, Bohr, & Garfinkel, 1982), the Eating Disorder Inventory

(Garner, Olmsted, & Polivy, 1983), the Bulimia Test—Revised (Thelen, Farmer, Wonderlich, & Smith, 1991), the Kids Eating Disorders Survey (Childress, Brewerton, Hodges, & Jarrell, 1992), and the Eating Disorder Examination—Questionnaire (Fairburn & Beglin, 1994) do not yield *DSM-IV* diagnoses for these eating disorders. The Questionnaire for Eating Disorder Diagnoses (Mintz, O'Halloran, Mulholland, & Schneider, 1997) does generate diagnoses for these three eating disorders but is a lengthy measure, and the agreement between these diagnoses and those from validated structured interviews has not been established.

A brief self-report questionnaire would be useful for etiologic research, because psychiatric interviews are time consuming and expensive to conduct with large samples. Such a questionnaire might also be useful when researchers need frequent measures of eating pathology (e.g., in prevention or treatment trials) because it would minimize response burden. Finally, such a questionnaire might prove useful in clinical settings where a brief screening measure is needed to identify individuals with eating pathology.

In a previous article, Stice, Telch, and Rizvi (2000) described the development of a brief self-report scale for diagnosing anorexia nervosa, bulimia nervosa, and binge eating disorder (the Eating Disorder Diagnostic Scale [EDDS]). The EDDS contains 22 items that assess *DSM-IV* symptoms for all three eating disorders that were adapted from validated structured psychiatric interviews: the EDE (Fairburn & Cooper, 1993) and the eating disorder module of the SCID (Spitzer et al., 1990). Sample items include "Has your weight influenced how you think about (judge) yourself as a person?"; "During the past 6 months have there been times when you felt you have eaten what other people would regard as an unusually large amount of food (e.g., a quart of ice cream) given the circumstances?"; "How many times per week on average over the past 3 months have you made yourself vomit to prevent weight gain or counteract the effects of eating?"; "Over the past 3 months, how many menstrual periods have you missed?" The EDDS generates diagnoses for all three *DSM-IV* eating disorders and a continuous eating disorder symptom composite. The scale can be scored by hand or with a computer algorithm (see the Appendix).

---

Eric Stice, Melissa Fisher, and Erin Martinez, Department of Psychology, University of Texas at Austin.

This study was supported by a career award (MH1708) and research grants (MH/DK61957 and MH64560) from the National Institute of Mental Health. We thank Molly Forester, Kacey Little, Kendra Orjada, and Kathryn Whitenton for their assistance in managing the various data collection projects. We also thank Cristy Telch for her central role in the original development of the EDDS.

Correspondence concerning this article should be addressed to Eric Stice, who is now at the Oregon Research Institute, 1715 Franklin Boulevard, Eugene, OR 97403. E-mail: estice@ori.org

The initial report (Stice, Telch, & Rizvi, 2000) presented preliminary evidence of the reliability and validity of the EDDS. A panel of eating disorder experts verified the content validity of the EDDS by indicating that this measure assessed all relevant symptoms and contained no extraneous items. Data from individuals with and without eating disorders ( $N = 367$ ) provided evidence of the test–retest reliability (mean  $\kappa = .80$ ) and criterion validity with interview diagnoses (mean  $\kappa = .83$ ) for EDDS diagnoses. In support of convergent validity, EDDS-identified individuals with eating disorders showed elevations on validated measures of eating disturbances relative to EDDS-identified individuals without an eating disorder. The overall eating disorder symptom composite showed test–retest reliability ( $r = .87$ ), internal consistency (mean  $\alpha = .89$ ), and convergent validity with extant eating pathology scales.

Although these preliminary data were encouraging, it seemed prudent to conduct further research on the reliability and validity of the EDDS. In the studies reported here, we attempted to replicate the evidence that the EDDS possesses criterion validity, convergent validity, and internal consistency and sought to provide new evidence concerning the sensitivity and predictive validity of this questionnaire.

## Study 1

### Method

#### Participants and Procedure

Participants were 728 adolescent girls and young women recruited from a large southwest metropolitan area. Because eating disorders are rare among boys and men, all of the samples in this report focused on data from girls and women. Participants included (a) a middle school-recruited sample of adolescent girls in a longitudinal risk factors study ( $n = 220$ ), (b) a high school- and university-recruited randomized trial of an eating disorder prevention program ( $n = 219$ ), (c) a university-recruited sample of women with bulimia nervosa and a weight-matched control sample without an eating disorder from a lab experiment ( $n = 101$ ), and (d) a university-recruited sample of undergraduates from two evaluations of an eating disorder prevention trial ( $n = 188$ ).

Participants from the first source were originally recruited from middle schools and received a gift certificate for completing each annual assessment. Participants from the second source were recruited from high schools (62%) and universities (38%) through direct mailings, fliers, and handbills. They were paid \$20 for completing the questionnaire and interview. Participants from the third source were recruited from the research participant pool at a university and received course credit for completing the project. The EDDS was used as an initial screening measure to identify individuals with probable bulimia nervosa and weight-matched control individuals, and the EDE was used to confirm these diagnoses. Participants from the fourth source were recruited from upper division undergraduate psychology classes. A drawing was held for gift certificates to compensate these participants.

Participants completed the EDDS at the baseline phase of the studies with the exception of those from the first source, who completed this scale at the third annual assessment. Along with the EDDS, participants completed the additional self-report scales used in the convergent validity analyses. Participants from the first three sources then completed the EDE, and those from the first source also completed an interview assessing diagnostic symptoms of major depression. After the interview, clinical assessors measured height and weight. Among the 540 participants from the first three sources who completed the EDE, 1 met criteria for anorexia

nervosa, 34 met criteria for bulimia nervosa, and 1 met criteria for binge eating disorder.

Assessors had either a bachelor's, master's, or doctoral degree in psychology. They attended 24 hr of training, during which structured interview skills were taught, diagnostic criteria for the disorders were reviewed, simulated interviews were observed, and interviews were role-played. They had to demonstrate an interrater agreement ( $\kappa > .80$ ) with experts using tape-recorded interviews before collecting data. Interviews were recorded periodically throughout the study to ensure that assessors continued to demonstrate acceptable interrater agreement ( $\kappa > .80$ ) with experts.

Participants in the combined sample ranged in age from 13 to 55 years ( $M = 17.9$ ,  $SD = 3.4$ ). The sample was composed of 8% Asians/Pacific Islanders, 6% Blacks, 17% Hispanics, 1% Native Americans, 65% Whites, and 3% who specified "other" or mixed racial heritage. The mean body mass index ( $BMI = \text{kg/m}^2$ ; Garrow & Webster, 1985), which reflects relative weight, was 22.6 ( $SD = 4.2$ ).

### Measures

**EDDS.** The EDDS (Stice, Telch, & Rizvi, 2000) contains items assessing the *DSM-IV* diagnostic criteria for anorexia nervosa, bulimia nervosa, and binge eating disorder. Responses can be used to generate *DSM-IV* diagnoses for the three eating disorders. Items can also be standardized (to control for the different response formats) and summed (except the height and birth control pill use items) to create an overall eating disorder symptom composite. The preliminary psychometric study (Stice, Telch, & Rizvi, 2000) provided evidence that the EDDS was reliable and valid within a sample containing both adolescents and adults.

**EDE.** The EDE (Fairburn & Cooper, 1993) was the gold standard against which the EDDS was compared for the test of the criterion validity. The EDE is a semistructured psychiatric interview assessing current diagnostic criteria for all three *DSM-IV* eating disorders. Diagnostic items were also standardized and averaged to form an overall eating disorder symptom composite. Because the symptom composite was skewed, a square root transformation was used to normalize the distribution. Studies with community and clinical samples have reported test–retest reliability correlations for diagnostic items ranging from .83 to .97 and interrater reliability kappa coefficients ranging from .83 to .99 (Fairburn & Cooper, 1993; Rizvi, Peterson, Crow, & Agras, 2000; Williamson, Anderson, Jackman, & Jackson, 1995; Wilson & Smith, 1989). Studies have also found that the EDE diagnostic items and subscales are able to discriminate individuals with eating disorders from overweight individuals, dieters, and normal control individuals (Cooper, Cooper, & Fairburn, 1989; Fairburn & Cooper, 1993; Wilson & Smith, 1989). To assess the interrater reliability of the EDE in our adolescent samples, we had a randomly selected subset of participants (5%) interviewed within a 3-day period by a second assessor who was blinded to the first diagnosis; this resulted in high interrater agreement ( $\kappa = .88$ ). Another randomly selected subset of participants (5%) completed a second diagnostic interview with the same assessor 1 week later, resulting in high test–retest reliability ( $\kappa = .97$ ). The EDE eating disorder symptom composite showed acceptable internal consistency in our sample (Cronbach's  $\alpha = .82$ ).

**Thin-ideal internalization.** Internalization of the thin ideal was assessed with the Ideal-Body Stereotype Scale—Revised (Stice, 2001). This scale asks participants to indicate their level of agreement with statements concerning what constitutes physical attractiveness in women using 5-point subscales that range from *strongly disagree* to *strongly agree*. This scale was found to have acceptable internal consistency ( $\alpha = .89$ ), temporal reliability (10-month test–retest  $r = .63$ ), and predictive validity in adolescent samples (Stice, 2001). Pilot testing revealed a 2-week test–retest coefficient of .80 for this scale.

**Body dissatisfaction.** A short form of the Satisfaction and Dissatisfaction with Body Parts Scale (Berscheid, Walster, & Bohrnstedt, 1973) assessed body dissatisfaction. This scale asks participants to indicate their level of satisfaction with nine body parts (e.g., waist, thighs) using a

5-point response format ranging from *extremely satisfied* to *extremely dissatisfied*. This scale was found to have acceptable internal consistency ( $\alpha = .94$ ), 3-week test-retest reliability ( $r = .90$ ), and predictive validity in adolescent samples (Stice, 2001).

**Dieting.** The Dutch Restrained Eating Scale (van Strien, Frijters, Bergers, & Defares, 1986) assessed the frequency of dieting behaviors using 5-point response scales ranging from *never* to *always*. This scale possessed acceptable internal consistency ( $\alpha = .95$ ) and 2-week test-retest reliability ( $r = .82$ ) and correlated negatively with self-reported caloric intake in adolescent and adult samples (Stice, 2001; van Strien et al., 1986).

**Negative emotionality.** Buss and Plomin's (1984) Emotionality Scale assessed participants' proclivity to become affectively distressed using a 5-point response format ranging from *never true of me* to *always true of me*. This scale had acceptable internal consistency ( $\alpha = .82$ ), temporal reliability (10-month test-retest  $r = .69$ ), and predictive validity in adolescent and adult samples (Buss & Plomin, 1984; Stice, 2001).

**Depression.** The Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS; Puig-Antich & Chambers, 1983) diagnostic interview assessed *DSM-IV* major depression symptoms. Severity ratings for each symptom were averaged to form an overall depressive symptom composite. The K-SADS had high test-retest reliability ( $\kappa = .63$ – $1.00$ ), interrater reliability ( $\kappa = .73$ – $1.00$ ), internal consistency ( $\alpha = .68$ – $.84$ ), and discriminant validity in adolescent samples (Ambrosini, 2000; Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993). A second assessor who was blinded to the first diagnosis interviewed a randomly selected subset of girls (5%) within a 3-day period, resulting in high interrater agreement for major depression diagnoses ( $\kappa = 1.00$ ). Another randomly selected subset of girls (5%) was interviewed a second time by the same assessor 1 week later, resulting in high test-retest reliability ( $\kappa = 1.00$ ).

**Social impairment.** Items adapted from the Social Adjustment Scale (SAS; Weissman & Bothwell, 1976) assessed impaired psychosocial functioning in the family, peer group, school, and work spheres. Research found that the SAS possessed internal consistency (mean  $\alpha = .74$ ), temporal reliability (mean test-retest reliability  $r = .80$ ), convergent validity with clinician and collateral ratings (mean  $r = .72$ ), discriminant validity, and treatment sensitivity (Edwards, Yarvis, Mueller, Zingale, & Wagman, 1978; Weissman & Bothwell, 1976; Weissman, Prusoff, Thompson, Harding, & Myers, 1978). Pilot testing with adolescents revealed an alpha of .77 and a 1-week test-retest correlation of .83 for this adapted scale.

## Results

### Criterion Validity

The criterion validity of the EDDS was assessed by examining the concordance between the EDDS diagnoses and the EDE interview diagnoses for anorexia nervosa, bulimia nervosa, and binge eating disorder among the subset of 443 participants who provided complete data for both measures. The kappa coefficient, which represents the chance-corrected level of agreement between two nominal variables, was .78. This kappa is considered excellent according to the criteria proposed by Fleiss (1981). The sensitivity of .88 reflects the proportion of individuals with a positive interview diagnosis that were correctly identified by the EDDS. The specificity of .98 reflects the proportion of individuals with a negative interview diagnosis who were correctly identified by the EDDS. The positive predictive value of .74 represents the proportion of individuals who were classified as having a positive diagnosis by the EDDS who actually met criteria for that diagnosis on the structured interview. The negative predictive value of .98 represents the proportion of individuals who were classified as having a negative diagnosis by the EDDS who actually did not

meet criteria for the diagnosis on the structured interview. The overall accuracy of .96 represents the proportion of individuals for whom the negative and positive EDDS diagnoses matched the interview diagnoses. These results should be interpreted as primarily reflecting the concordance between the EDDS and EDE for a bulimia nervosa diagnosis versus an absence of an eating disorder because of the low rates of EDE-identified anorexia nervosa and binge eating disorder in the subsample that completed both scales.

The correlation between the EDDS symptom composite and the parallel symptom composite from the EDE was .82. This suggests that there is acceptable concordance between both the eating disorder diagnoses and the continuous symptom composite.

### Internal Consistency

The internal consistency of the EDDS symptom composite was assessed by calculating Cronbach's alpha for the standardized items that make up this score. Cronbach's alpha for the symptom composite was .89 in the full sample.

### Convergent Validity

To generate evidence of convergent validity, we tested whether EDDS-identified individuals with bulimia nervosa showed the expected elevations on validated measures of eating pathology, risk factors for bulimic pathology (thin-ideal internalization, body dissatisfaction, dieting, temperamental emotionality, depressive symptoms), and social impairment relative to EDDS-identified individuals without an eating disorder. Analyses solely compared EDDS-identified individuals with bulimia nervosa to control individuals without an eating disorder because among the participants who provided data for the EDDS ( $n = 728$ ), there were not sufficient numbers of individuals with EDDS-identified anorexia nervosa ( $n = 5$ ) and binge eating disorder ( $n = 9$ ) for analyses. One-way analysis of variance (ANOVA) models compared the EDDS-identified individuals with bulimia nervosa to those without an eating disorder. A Bonferroni correction was applied to minimize Type I error rate inflation ( $\alpha = .007$ ). The means and standard deviations for each variable for the two groups are presented in Table 1, along with the results of the planned contrasts and the percentages of variance accounted for each effect. EDDS-identified individuals with bulimia nervosa showed significant elevations on the validated measure of eating pathology, risk factors for eating pathology, and social impairment, relative to the EDDS-identified individuals without an eating disorder. These effects ranged from small to large in magnitude according to Cohen's (1988) criteria.

To examine the convergent validity of the EDDS symptom composite, we tested whether this composite was positively correlated with risk factors for eating pathology and social impairment (see Table 2). The EDE symptom composite was included for comparison purposes. An alpha level of .001 was used to minimize Type I error rate inflation. The EDDS symptom composite showed significant positive correlations with the risk factors for eating pathology and social impairment. It was reassuring that the EDE symptom composite showed similar correlations with these criteria. The correlations were all large in magnitude according to Cohen's criteria.

Table 1  
*Mean Differences Between EDDS-Identified Individuals With Bulimia Nervosa and Individuals Without an Eating Disorder on Eating Pathology Measures*

Measure	Individuals with bulimia nervosa ( <i>n</i> = 80)		Individuals without an eating disorder ( <i>n</i> = 538)		Variance explained (%)
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
EDE symptom composite	4.41 <sub>a</sub>	2.91	0.89 <sub>b</sub>	0.99	43
Thin-ideal internalization	3.99 <sub>a</sub>	0.41	3.50 <sub>b</sub>	0.69	6
Body dissatisfaction	3.74 <sub>a</sub>	1.04	3.16 <sub>b</sub>	0.98	4
Dieting	3.46 <sub>a</sub>	0.80	2.32 <sub>b</sub>	0.96	14
Negative affectivity	3.21 <sub>a</sub>	0.48	2.80 <sub>b</sub>	0.81	4
Depressive symptoms	19.40 <sub>a</sub>	11.60	9.00 <sub>b</sub>	8.22	20
Social impairment	3.01 <sub>a</sub>	0.41	2.35 <sub>b</sub>	0.53	6

*Note.* Row means with different subscripts are significantly different (Bonferroni corrected  $\alpha = .007$ ). Means and standard deviations for the EDE symptom composite are from the untransformed version of this variable, whereas the transformed version was used in the analyses. EDDS = Eating Disorder Diagnostic Scale; EDE = Eating Disorder Examination.

### Discussion

The results from Study 1 provide additional evidence for the criterion and convergent validity of the EDDS diagnoses and symptom composite and for the internal consistency of the symptom composite. These results generally replicate those in the original psychometric investigation (Stice, Telch, & Rizvi, 2000). The fact that this sample was younger than that used in the original investigation of the concordance between the EDDS and the EDE suggests that this scale can be used with adolescent girls as well as adult women. An important limitation, however, is that the sample contained low rates of EDE-confirmed anorexia nervosa and binge eating disorder.

It was encouraging that the findings from Study 1 replicated the evidence for the reliability and validity of the EDDS observed in the initial psychometric study (Stice, Telch, & Rizvi, 2000), but it would be useful to amass additional evidence for the validity of this scale. Thus, in Study 2 we sought to determine whether the EDDS possessed sufficient sensitivity to detect intervention effects by testing whether the EDDS could detect the effects of a psychoeducational eating disorder prevention intervention. Data were drawn from a past evaluation of this psychoeducational interven-

tion (Stice & Ragan, 2002) and from an unpublished replication of this trial (Orjada & Stice, 2003).

### Study 2

#### Method

#### Participants and Procedure

Participants were 150 women undergraduates who enrolled in one of two sections of an eating disorder class or other upper division seminars offered concurrently in the psychology department at a university. The mean age of participants was 21.6 years (range = 19–55). The sample was composed of 14% Asians/Pacific Islanders, 2% Blacks, 14% Hispanics, 68% Whites, and 2% who specified “other.” The mean BMI of participants was 22.3 ( $SD = 3.8$ ).

The study was described as an evaluation of the effects of a course on students’ attitudes and behaviors. Students in the eating disorder class were asked to complete pretest and posttest surveys to help the instructor gain a more complete understanding of the effects of the class and to provide guidance for refinements to the course. All students in both classes volunteered to complete the surveys. Students in six other upper division psychology seminars were asked to complete parallel pretest and posttest

Table 2  
*Correlations Between the EDDS Symptom Composite, the EDE Symptom Composite, Risk Factors for Eating Pathology, and Social Maladjustment*

Variable	1	2	3	4	5	6	7	8
1. EDDS symptom composite	—	.82	.46	.52	.69	.39	.45	.40
2. EDE symptom composite		—	.40	.46	.63	.35	.50	.41
3. Thin-ideal internalization			—	.27	.39	.32	.23	.19
4. Body dissatisfaction				—	.49	.31	.22	.22
5. Dieting					—	.35	.28	.32
6. Negative affectivity						—	.49	.43
7. Depressive symptoms							—	.43
8. Social maladjustment								—

*Note.* Absolute correlations greater than .18 are significant at  $p < .001$ . EDDS = Eating Disorder Diagnostic Scale; EDE = Eating Disorder Examination.

surveys. Approximately 70% of these students agreed to complete the surveys. Students were recruited from other upper division courses for a matched comparison group in an effort to control for the amount of time spent in a seminar, contact with other students and an instructor, and other demographic factors (e.g., age). Participants completed the 30-min pretest survey on the first day of classes and an identical posttest survey on the final day of classes 4 months later. Drawings were held for gift certificates to compensate students for completing the surveys.

### Intervention

The intervention consisted of an upper division seminar on eating disorders that met twice weekly for 1.5 hr each session over the 15-week semester. The course focused on descriptive pathology, epidemiology, etiologic models, risk factors, preventive interventions, and treatments for eating disorders and obesity (details of the class are presented in Stice & Ragan, 2002). Students were required to do a 20–30 min class presentation that represented a critical synthesis of a topic of their choice at the end of the course and to write a 10-page paper on the same subject. They also completed three written essay exams.

The EDDS (Stice, Telch et al., 2000) was used as the sole measure of intervention effects. EDDS items were summed to form a continuous measure of overall eating disorder symptoms (as described in Study 1). EDDS responses were also used to generate threshold and subthreshold *DSM-IV* diagnoses of current anorexia nervosa, bulimia nervosa, and binge eating disorder. Because of the modest sample size in this trial, we focused on both threshold and subthreshold diagnoses to maximize statistical power. Subthreshold diagnoses required the presence of all of the symptoms of the disorder but allowed at least one of these symptoms to be of subdiagnostic severity (e.g., binge eating only once per week). The eating disorder symptom composite had an alpha of .90 at pretest.

### Results

Preliminary analyses indicated that the rates of threshold and subthreshold eating disorders (as assessed by the EDDS) were significantly higher,  $\chi^2(4, N = 165) = 10.6, p = .031$ , for the participants who signed up for the eating disorder classes (33%) than for the students in the other upper-division courses (15%). The two groups also differed in terms of dieting but not thin-ideal internalization, body dissatisfaction, negative emotionality, body mass, age, year in school, or ethnicity at pretest. Accordingly, control participants were matched to intervention participants to reduce the possibility that regression to the mean could account for any putative intervention effects. To ensure an adequate sample size, we matched up to 3 control participants to each intervention participant (the control participant's eating disorder symptom scores had to be within 0.5 standard deviations of the intervention participant's score to be matched). This process resulted in a final sample size of 150 (42 intervention participants and 108 control participants). Analyses verified that the intervention and matched control groups did not differ in terms of eating disorder symptoms, eating disorder diagnoses, or any other study variable. Moreover, all analyses testing for intervention effects controlled for baseline levels of the outcome to rule out the possibility that any initial differences between groups (even if nonsignificant) could account for apparent intervention effects.

Because two of the instructors from which we recruited our matched controls did not have time to allow us to collect follow-up data at the end of the term during the class periods, we could not collect posttest data from 34% of the participants. We were able to collect complete pretest and posttest data from only 35 interven-

tion participants and 63 control participants. Analyses indicated that those who did not provide posttest data did not differ from those who did provide data at both assessments on any study variables.

A repeated measures ANOVA model tested whether intervention participants showed significantly greater reductions in the continuous EDDS eating disorder symptom composite from pretest to posttest than matched controls (condition was a two-level between-subjects factor and time was a two-level within-subjects factor). A significant Time  $\times$  Condition interaction,  $F(1, 96) = 7.00, p = .010$ , indicated that the reduction in the EDDS eating disorder symptom scale for intervention participants was significantly larger than the decrease observed in the control group. This effect accounted for 7% of the variance in change in eating disorder symptoms over the 4-month study. Paired *t* tests indicated that there was a significant reduction in EDDS eating disorder symptoms in the intervention condition but not in the control condition. Means and standard deviations for the intervention and control groups at pretest and posttest are reported in Table 3, along with results from follow-up paired *t* tests (Bonferroni corrected  $\alpha = .025$ ).

Because of concerns that the high attrition rate might have biased our results, we reestimated the model reported above using missing values replicated through multiple imputation (Sinharay, Stern, & Russell, 2001). Multiple imputation models the statistical uncertainty in the missing data by generating multiple data sets with imputed values. The existing data in the models are used to predict the values of missing data, ensuring that the imputed data are similar to related variables and thus a viable value. The uncertainty of the value of the missing data is represented by the multiple values generated for each missing data point. As multiple imputation has been shown to be an efficient algorithm, only five iterations are necessary (Rubin, 1996). Hence, five data sets were imputed, and five identical repeated measures ANOVA models that parallel the model used in the above analyses were conducted on each of the five data sets. Inferential results were produced by combining results from the five analyses conducted with the five data sets. The imputation method used to generate data sets was the Markov Chain Monte Carlo (MCMC) method, which uses Bayesian inference to generate a posterior distribution. The expectation-maximization algorithm, a technique for maximum likelihood estimation, was used to set the initial values for the MCMC process. The results generated from the combined parameter estimates with adjusted degrees of freedom showed a significant Time  $\times$  Condi-

Table 3  
*Pretest and Posttest Means and Standard Deviations for the Intervention and Matched Control Groups on the Eating Disorder Diagnostic Scale (EDDS) Symptom Composite From Study 2*

	Pretest		Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
EDDS symptom composite				
Intervention group ( <i>n</i> = 35)	3.76 <sub>a</sub>	11.96	0.75 <sub>b</sub>	13.76
Matched control group ( <i>n</i> = 63)	−0.23	10.10	0.89	11.69

*Note.* Means within the same group with different subscripts were significantly different (Bonferroni corrected  $\alpha = .025$ ).

tion interaction,  $F(1, 61.98) = 7.08, p < .01$ . Thus, similar to the analyses of the completer data set, results indicated that the reduction in the EDDS eating disorder symptom scale was significantly larger for intervention participants relative to the control group. Paired  $t$  tests generated from combined parameters of multiply imputed MCMC data indicated that there was a significant reduction in EDDS eating disorder symptoms in the intervention condition but not in the control condition.

To provide an index of the clinical importance of this intervention, we tested whether the rate of threshold and subthreshold eating disorders at the posttest was lower in the intervention condition than in the control condition. We regressed the dichotomous variable that reflected the presence of a threshold or subthreshold diagnosis of eating pathology at posttest on a dummy coded vector representing treatment condition and the baseline eating disorder symptom composite score (to control for initial levels of eating pathology). Participants in the intervention condition were marginally,  $Wald(1) = 2.58, p = .055$ , less likely to meet criteria for any eating pathology at posttest (9%) than control participants (25%).

We again reestimated this analyses using multiple imputation to generate values for the missing cases. A logistic regression technique was used that first fits a regression model and then simulates a second logistic regression model from the posterior predictive distribution of the parameters, which is then used to impute missing observations (Rubin, 1987). Logistic regression models were then used to estimate models on the five imputed data sets. Analysis of the combined parameter estimates showed that participants in the intervention condition were marginally less likely to meet criteria for any eating pathology at posttest than control participants,  $t(188.21) = 1.80, p = .07$ , again replicating the effects from the analyses that used the completer data.

### Discussion

The results of Study 2 provide preliminary evidence that the continuous symptom measure and the diagnoses from the EDDS were sufficiently sensitive to detect intervention effects in a controlled evaluation of an eating disorder prevention intervention. However, these findings would be more compelling if there was also evidence of the intervention's effectiveness from a gold standard, if the data were drawn from a randomized prevention trial to reduce the possibility that a confound accounts for the apparent intervention effects, and if the effect replicated in a larger sample. Accordingly, Study 3 tested whether the EDDS was sufficiently sensitive to detect intervention effects from a randomized eating disorder prevention trial and compared the sensitivity of the EDDS to the EDE. Data were drawn from an ongoing randomized trial of a dissonance eating disorder prevention program that has been evaluated previously (Stice, Trost, & Chase, 2003).

An additional aim of Study 3 was to investigate the predictive validity of the EDDS. Treatment and prevention trials have found that intervention effects are often stronger for participants with initial elevations in symptomatology (e.g., Buddeberg-Fischer, Klaghofer, Gnam, & Buddeberg, 1998; Bulik, Sullivan, Carter, McIntosh, & Joyce, 1999). The use of the EDE as the gold standard outcome measure of this eating disorder prevention trial allowed us to test whether initial elevations on the EDDS symptom composite would predict a heightened response to the intervention

as assessed by the EDE. This was an important aim because the predictive validity of this new measure has not been previously tested.

## Study 3

### Method

#### Participants and Procedure

Participants were 181 adolescent girls who ranged in age from 13 to 20 years (mean age = 17.1). The sample was composed of 10% Asians/Pacific Islanders, 3% Blacks, 18% Hispanics, 63% Whites, and 6% who specified "other." The mean BMI of participants was 22.9 ( $SD = 4.1$ ).

Participants were recruited with direct mailing and fliers inviting high school and college girls and young women between the ages of 13 and 20 with body image concerns to participate in a study evaluating interventions aimed at helping them improve their body satisfaction. Participants were randomly assigned to the dissonance intervention, an expressive writing alternative-intervention control condition, or a measurement-only control condition. The dissonance intervention consisted of three 1-hr sessions that occurred weekly (groups contained 4–12 participants). Sessions were facilitated by a psychology graduate student and cofacilitated by an undergraduate. Participants in the expressive writing control condition were asked to write about their feelings during three weekly 1-hr sessions. Participants completed a structured diagnostic interview and a survey at pretest and at posttest (1 month later) and were compensated \$20 for each assessment.

#### Dissonance Intervention

*Session 1.* First, an overview of the purpose of the study was provided. Participants were told that this intervention was based on the idea that the act of discussing how to help other women avoid body image problems can help improve their own body satisfaction. Next, the group was asked to collectively define the thin ideal. The remainder of the session was spent critically discussing (a) the origin of the thin ideal; (b) how it is perpetuated; (c) the impact of messages about the thin ideal from family, peers, dating partners, and the media; and (d) who benefits from the thin ideal. Finally, participants were asked if they would be willing to write a one-page (counterattitudinal) essay about the costs associated with pursuit of the thin ideal.

*Session 2.* After an overview of the previous session, discussion focused on participants' experiences of writing the essay and the costs associated with pursuing the thin ideal. Second, the group discussed who benefits from society's pursuit of the thin ideal, with a focus on corporations that profit from it. Third, a counterattitudinal role play was conducted, in which participants attempted to dissuade the group leaders from pursuing the thin ideal. Finally, participants were asked to engage in a body acceptance exercise at home in which they examine their reflection in a mirror and record positive aspects of themselves (physical, behavioral, emotional, and social characteristics). They were asked not to record any negative thoughts.

*Session 3.* After an overview of the previous session, each participant was asked to share her responses to the body acceptance homework and the feelings and thoughts she had while completing it. Next, the group discussed difficulties participants might encounter in resisting the thin ideal and techniques to overcome these obstacles. Participants were asked to role-play, making counter thin-ideal statements to resist pressure from peers, and the group explored ways in which they might unwittingly promote the thin ideal. Group leaders then asked for specific suggestions to help younger girls accept their bodies and avoid pursuit of the thin ideal. Finally, participants were encouraged to challenge themselves if they

engaged in thin-ideal thinking and to enact behavioral challenges related to body image concerns they may have in the future.

### *Expressive Writing Alternative-Intervention Control Condition*

In the expressive writing control condition, participants were asked to write about issues of emotional significance in three weekly sessions. The scripted instructions informed participants that the exercise was a novel application of research that suggests body image concerns are a product of emotional issues, which are often resolved through expressive writing. The instructions offered examples of potential topics, such as relationships or goals, as guidelines for this writing exercise. Participants were informed that their work would not be read and that the only requirement was that they write continuously for the duration of the session about an emotionally important topic.

### *Measures*

**EDE.** The EDE (Fairburn & Cooper, 1993) was used as the primary outcome. Items were summed to form overall eating disorder symptom composites at pretest and posttest (see Study 1). Responses to the EDE interview were also used to classify participants as meeting threshold or subthreshold diagnostic criteria for anorexia nervosa, bulimia nervosa, and binge eating disorder at pretest and posttest. Subthreshold diagnoses required the presence of all of the symptoms of the disorder but also that at least one of these symptoms was of subdiagnostic severity. Past studies that investigated the reliability and validity of this measure are reported in Study 1. The eating disorder symptom composite had an alpha of .92 in this sample at pretest.

**EDDS.** The EDDS (Stice, Telch et al., 2000) was used as the secondary outcome. EDDS items were summed to form a continuous measure of overall eating disorder symptoms and were used to generate threshold and subthreshold *DSM-IV* diagnoses of current anorexia nervosa, bulimia nervosa, and binge eating disorder (as described in Studies 1 and 2). The eating disorder symptom composite had an alpha of .90 at pretest.

### *Results*

The intervention, alternative-intervention control, and measurement-only control groups did not differ in terms of eating disorder symptoms, age, ethnicity, or parental education at pretest,

suggesting that randomization created generally equivalent groups. Only 5 participants (3%) did not provide complete data at posttest. Attrition did not appear to have introduced systematic bias, because those who did not provide posttest data did not differ from those who provided complete data on any study variables. There was no significant difference in the expected benefits of the dissonance intervention and the expressive writing condition by the end of the first session, suggesting that the alternative-intervention control condition was perceived as credible.

We first tested whether the EDE detected any intervention effects. A repeated measures ANOVA model indicated that there were significantly greater reductions in EDE eating disorder symptoms from pretest to posttest in the intervention condition than in the measurement-only condition,  $F(1, 115) = 10.35, p = .002$ , 8% variance explained. There were also significantly greater reductions in EDE eating disorder symptoms from pretest to posttest in the intervention condition than in the alternative-intervention control condition,  $F(1, 127) = 16.70, p < .001$ , 12% variance explained. Paired *t* tests indicated that there was a significant reduction in EDE eating disorder symptoms in the intervention condition but not in the two control conditions (Bonferroni corrected  $\alpha = .008$ ). Means and standard deviations for the intervention and control groups at pretest and posttest are reported in Table 4, along with results from follow-up paired *t* tests. The EDE indicated that this intervention produced statistically reliable effects from pretest to posttest compared with both the alternative-intervention and measurement-only control conditions.

Next, we tested whether the EDDS was sufficiently sensitive to detect the intervention effects. There were significantly greater reductions in EDDS eating disorder symptoms from pretest to posttest in the intervention condition than in the measurement-only condition,  $F(1, 114) = 4.77, p = .031$ , 4% variance explained. There were also significantly greater reductions in EDDS eating disorder symptoms from pretest to posttest in the intervention condition than in the alternative-intervention control condition,  $F(1, 124) = 9.53, p < .002$ , 7% variance explained. Paired *t* tests indicated that there were significant reductions in EDDS eating disorder symptoms in the intervention condition but not in the two

Table 4  
*Pretest and Posttest Means and Standard Deviations for Study 3 Groups on the EDE and EDDS Symptom Composites*

Symptom composite and group	Pretest		Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<b>EDE</b>				
Intervention group ( <i>n</i> = 70)	14.31 <sub>a</sub>	14.43	7.84 <sub>b</sub>	8.09
Alternative-intervention control group ( <i>n</i> = 59)	12.59	11.57	12.25	12.58
Measurement-only control group ( <i>n</i> = 47)	14.08	13.37	13.00	12.36
<b>EDDS</b>				
Intervention group ( <i>n</i> = 69)	0.43 <sub>a</sub>	11.80	-1.67 <sub>b</sub>	10.27
Alternative-intervention control group ( <i>n</i> = 57)	-1.55	10.71	1.71	12.78
Measurement-only control group ( <i>n</i> = 47)	0.01	12.25	2.31	15.16

*Note.* Means within the same group with different subscripts were significantly different (Bonferroni corrected  $\alpha = .008$ ). Because the EDDS symptom composite was positively skewed at posttest, it was necessary to apply a normalizing transformation. However, the untransformed variable means and standard deviations are reported here to facilitate comparison across studies. EDE = Eating Disorder Examination; EDDS = Eating Disorder Diagnostic Scale.

control conditions. Means and standard deviations for the intervention and control groups at pretest and posttest are reported in Table 4, along with results from follow-up paired *t* tests.

To provide an index of the clinical importance of this intervention, we tested whether the rate of threshold and subthreshold eating disorders at the posttest was lower in the intervention condition than in the control conditions. We regressed the dichotomous variable that reflected the presence of a threshold or subthreshold diagnosis of eating pathology at posttest on a dummy coded vector representing treatment condition and the baseline eating disorder symptom composite (to control for initial variation in eating pathology). Participants in the intervention condition were significantly less likely to meet criteria for any eating pathology at posttest (6%) compared with the measurement-only controls (15%) and the alternative-intervention controls (15%) on the EDE, Wald(1) = 3.41,  $p = .032$ , and Wald(1) = 4.81,  $p = .014$ , respectively. Participants in the intervention condition were similarly less likely to meet criteria for any eating pathology at posttest (7%) compared with measurement-only controls (13%) and alternative-intervention controls (10%) on the EDDS, but neither effect was statistically significant, Wald(1) = 1.41,  $p = .117$ , and Wald(1) = 0.85,  $p = .178$ , respectively.

Finally, we investigated the predictive validity of the EDDS by testing whether initial elevations on this measure predicted a stronger response to the intervention as assessed by the EDE (i.e., whether the pretest EDDS symptom composite moderated the effects of the intervention). Multiple regression models tested whether pretest EDDS scores interacted with treatment condition to predict posttest EDE scores controlling for pretest EDE scores. Separate models compared the intervention to the measurement-only control condition and to the alternative-intervention control condition. In both models, the EDDS  $\times$  Treatment condition interaction was significant ( $p < .001$ , 4% variance explained; and  $p < .001$ , 9% variance explained; respectively). Follow-up analyses verified that the intervention effects were larger for participants in the upper half of the EDDS symptom composite than those in the lower half of the scale when the intervention was compared with measurement-only condition (11% vs. 8% variance explained) and when the intervention was compared with the alternative-intervention condition (16% vs. 6% variance explained).

### Discussion

In summary, the EDE data suggested that the intervention produced significant reductions in eating disorder symptoms and rates of threshold and/or subthreshold eating disorder diagnoses. Findings suggested that the EDDS symptom composite was sufficiently sensitive to detect interventions effects but that the EDDS diagnoses were less sensitive. These findings imply that the EDDS should be used only as a general measure of eating disorder symptoms but that the EDE is a more sensitive instrument that should be used as the outcome variable when change in eating disorder diagnoses will be the central outcome measure in prevention and treatment trials. Results also provided evidence that the EDDS has predictive validity, in that participants with higher scores on the EDDS symptom composite showed a stronger response to the eating disorder prevention program relative to those with lower initial scores. These results converge with similar

findings observed in past prevention and treatment trials (e.g., Bulik et al., 1999). This finding is noteworthy because it is the first evidence that this measure possesses predictive validity.

Although Study 3 provided preliminary evidence for the predictive validity of the EDDS, it would be useful if there was additional evidence. Thus, Study 4 tested whether elevated EDDS scores predicted subsequent onset of eating pathology and major depression in a prospective study of adolescent girls. Theoretically, girls initially without an eating disorder with elevated scores on the EDDS would be more likely to show future onset of eating disorder symptoms. In addition, past research has found that eating pathology is a risk factor for onset of major depression (Johnson et al., 2002), presumably because the shame, guilt, and frustration associated with eating disturbances contributes to the development of depression.

## Study 4

### Method

#### Participants and Procedure

Participants were 220 adolescent girls from public and private high schools in a southwestern city. Data were drawn from a larger study of 496 adolescent girls who have been followed annually for 4 years with structured interviews and surveys. The 220 girls represented a subsample that was asked to complete the EDDS at the second follow-up interview of the study (T3). Participants ranged in age from 14 to 18 years ( $M = 15.8$ ) at T3 and included 2% Asians, 11% Blacks, 62% Whites, 21% Hispanics, 1% Native Americans, 3% other/mixed). The mean BMI of the girls was 22.4 ( $SD = 4.8$ ).

The study was described as an investigation of adolescent mental and physical health. An active consent procedure via direct mailings resulted in a mean participation rate of 56%. This rate was similar to that of other school-recruited samples that required active consent and structured interviews (e.g., 61% for Lewinsohn et al., 1993). The ethnic composition of the sample was representative of the schools from which we sampled (2% Asians, 8% Blacks, 65% Whites, 21% Hispanics, 4% other/mixed). This sample has been found to be representative in terms of socioeconomic status and prevalence of psychiatric disorders (Stice, Ragan, & Randall, in press).

In the present study, girls completed a survey and an interview at baseline and at three annual follow-ups (T1, T2, T3, and T4, respectively), but this report focused on the T3 and T4 data. Assessors, who were women with a bachelor's, master's, or doctoral degree in psychology, attended 24 hr of training, in which they learned interview skills, reviewed diagnostic criteria, observed interviews, and role-played interviews. Assessors had to demonstrate an interrater agreement ( $\kappa > .80$ ) with experts using tape-recorded interviews before collecting data. Interviews were randomly recorded to ensure that assessors continued to demonstrate acceptable interrater agreement ( $\kappa > .80$ ) with experts. Assessments took place during or after school hours on the school campus or at the girls' homes. Each girl received a \$15 gift certificate at each assessment.

#### Measures

**EDDS.** The EDDS (Stice, Telch, & Rizvi, 2000) was administered at T3. EDDS items were summed to form a continuous measure of overall eating disorder symptoms (as described in Study 1). The eating disorder symptom composite had an alpha of .86 at T3.

**EDE.** The EDE (Fairburn & Cooper, 1993) was used to assess diagnostic criteria for *DSM-IV* anorexia nervosa, bulimia nervosa, and binge eating disorder at T3 and T4. Past studies that investigated the reliability

and validity of this measure are reported in Study 1. Because only 1–3 girls showed onset of each of the three eating disorders (i.e., threshold or subthreshold diagnoses) from T3 to T4 in this subsample, analyses focused on predicting onset of at least two episodes of *DSM-IV* defined binge eating and compensatory behaviors from T3 to T4 ( $n = 8$  and  $n = 16$ , respectively). Insufficient numbers of girls reported onset of the *DSM-IV* criteria for a medically dangerous body weight (the cardinal feature of anorexia nervosa) for analyses ( $n = 1$ ).

**Depression.** The K-SADS (Puig-Antich & Chambers, 1983) diagnostic interview assessed *DSM-IV* major depression symptoms at T3 and T4. Girls were diagnosed with current major depression at each assessment. Analyses predicted onset of major depression ( $n = 14$ ) from T3 to T4. Past studies that investigated the reliability and validity of this measure are reported in Study 1, along with evidence that the K-SADS showed high interrater agreement ( $\kappa = 1.0$ ) and high test–retest reliability ( $\kappa = 1.0$ ) in our study.

### Results

To assess the predictive validity of the EDDS, we first estimated logistic regression models that tested whether the EDDS symptom composite at T3 predicted subsequent onset of eating disorder symptoms by T4 among those free of the symptom at T3. Separate models predicted onset of binge eating and compensatory behaviors. We standardized the EDDS symptom composite to facilitate interpretation of the odds ratios (ORs) from the logistic models. Results indicated that participants with elevated EDDS symptom composite scores at T3 were significantly more likely to show onset of both binge eating, Wald(1) = 3.42,  $p = .035$ , OR = 1.69, and compensatory behaviors, Wald(1) = 6.78,  $p = .005$ , OR = 2.00. The ORs indicate that for each standard deviation increase in the EDDS, there was a corresponding 69% increase in the probability for binge eating onset and a 100% increase for the probability for compensatory behavior onset.

We then tested whether the EDDS symptom composite at T3 predicted subsequent onset of major depression by T4 among the girls who were initially not depressed at T3. Participants with elevated EDDS symptom composite scores at T3 were significantly more likely to show onset of major depression, Wald(1) = 4.20,  $p = .020$ , OR = 1.57. The OR indicates that for each standard deviation increase in the EDDS, there was a corresponding 57% increase in the probability for major depression onset.

### Discussion

Study 4 provided evidence that the EDDS shows predictive validity, in that elevated scores on this measure predicted increased risk for subsequent onset of binge eating and compensatory behaviors and onset of major depression. The former findings are consistent with the expectation that asymptomatic girls with the highest scores on the EDDS would show the greatest risk for onset of clinically significant eating disorder symptoms. The latter findings accord with the theory that elevated eating pathology increases the risk for onset of clinically significant mood disturbances, a finding that has been previously observed (Johnson et al., 2002). These results are important because they provide additional evidence for the predictive validity of this new scale.

### General Discussion

The first aim of our study was to replicate the evidence that the EDDS possesses criterion validity, convergent validity, and inter-

nal consistency. Study 1 found that the EDDS eating disorder diagnoses and continuous symptom measure evidences criterion validity with the EDE, which is considered the current gold standard. The overall agreement between the EDDS and the EDE for *DSM-IV* eating disorder diagnoses was 96%, which replicates the overall agreement rate of 96% observed in the initial psychometric evaluation of the EDDS (Stice, Telch, & Rizvi, 2000). Findings also provide additional support for the convergent validity of the EDDS. Study 1 indicated that the EDDS eating disorder diagnoses and continuous symptom composite show the expected positive correlations with risk factors for eating pathology and social impairment, which replicates the findings observed in the initial psychometric study. The magnitude of these relations ranged from medium to large effect sizes according to the Cohen criteria. There is also support for the internal consistency of the EDDS continuous symptom composite. The results from all four studies suggest that the EDDS possesses adequate internal consistency (mean  $\alpha = .89$ ), which replicates the average internal consistency observed in the first psychometric study (mean  $\alpha = .89$ ).

The second aim of our study was to provide new evidence concerning the sensitivity and predictive validity of the EDDS. Study 2 found that the EDDS symptom composite and diagnoses are sufficiently sensitive to detect intervention effects. Study 3 similarly found that the EDDS symptom composite is sufficiently sensitive to detect intervention effects but suggests that the EDE is more sensitive for detecting change in eating disorder diagnoses. Although these results generally provide support for the sensitivity of the EDDS, they suggest that it is preferable to use the EDE in trials focusing on changes in eating disorder diagnoses.

Results also generated evidence for the predictive validity of the EDDS. Study 3 found that the EDDS symptom composite predicts response to a preventive intervention. Consistent with past findings, our results indicated that participants with the highest EDDS eating disorder symptoms show the greatest response to the eating disorder prevention program. Study 4 found that elevated EDDS symptom scores predict an increased risk for subsequent onset of binge eating and compensatory behaviors, as well as subsequent onset of major depression. The latter finding converges with past evidence that eating pathology is a risk factor for onset of depression.

It is important to note the limitations of these studies when interpreting the findings. First, the evidence in support of the criterion validity in Study 1 should be interpreted with caution because of the low numbers of participants with anorexia nervosa and binge eating disorder. Although the initial report included large numbers of individuals with these diagnoses, it would be useful if additional research examined the criterion validity of EDDS diagnoses of anorexia nervosa and binge eating disorder. Second, it may be illuminating for future studies to directly compare the psychometric properties (e.g., criterion validity) of the EDDS relative to other widely used measures of eating pathology. Third, the fact that only girls and women were included in the present studies precludes generalization to boys and men. It will be necessary for future studies to establish the reliability and validity of the EDDS for men. Fourth, the sample size used in Study 4 to investigate the predictive validity of the EDDS was relatively small given the base rates of the outcomes, which suggests that these results should be interpreted with care. It will also be important for future studies to conduct an in-depth study of the

factor structure of the EDDS to determine whether reliable subscales can be extracted from this measure that correspond to the three eating disorders articulated in the *DSM-IV*. Not only would such analyses provide information regarding the factor structure of the EDDS but they may also yield more sensitive factor-score composites.

These results provide additional evidence of the criterion validity, convergent validity, and internal consistency of the EDDS diagnoses and symptom composite. The fact that these findings replicate the evidence from the first psychometric investigation of the EDDS implies that these estimates of the reliability and validity are relatively robust. These studies also provide new evidence suggesting that the EDDS diagnoses and symptom composite are sufficiently sensitive to detecting intervention effects and show predictive validity. However, findings imply that structured psychiatric interviews are more sensitive measures of change in eating disorder diagnoses in controlled trials. These findings, taken in conjunction with those from the initial psychometric evaluation, provide mounting support for the reliability and validity of the EDDS. The benefits of this scale are that it can be completed quickly and easily relative to interview measures of eating pathology. Thus, it appears that this scale might prove useful for the assessment of eating disorders in etiologic, prevention, and treatment research applications, as well as in clinical settings where structured psychiatric interviews are difficult to conduct.

## References

- Ambrosini, P. J. (2000). Historical development and present status of the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS). *Journal of the American Academy of Child Psychiatry*, 39, 49–58.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Berscheid, E., Walster, E., & Bohrnstedt, G. (1973). The happy American body: A survey report. *Psychology Today*, 7, 119–131.
- Buddeberg-Fischer, B., Klaghofer, R., Gnam, G., & Buddeberg, C. (1998). Prevention of disturbed eating behaviour: A prospective intervention study in 14- to 19-year old Swiss students. *Acta Psychiatrica Scandinavica*, 98, 146–155.
- Bulik, C. M., Sullivan, P. F., Carter, F. A., McIntosh, V. V., & Joyce, P. R. (1999). Predictors of rapid and sustained response to cognitive-behavioral therapy for bulimia nervosa. *International Journal of Eating Disorders*, 26, 137–144.
- Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Erlbaum.
- Childress, A. C., Brewerton, T. D., Hodges, E. L., & Jarrell, M. P. (1992). The Kid's Eating Disorders Survey (KEDS): A study of middle school students. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32, 843–850.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cooper, Z., Cooper, P. J., & Fairburn C. G. (1989). The validity of the Eating Disorder Examination and its subscales. *British Journal of Psychiatry*, 154, 807–812.
- Edwards, D. W., Yarvis, R. M., Mueller, D. P., Zingale, H. C., & Wagman, W. J. (1978). Test-taking and the stability of adjustment scales. *Evaluation Quarterly*, 2, 275–291.
- Fairburn, C. G., & Beglin, S. J. (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. Fairburn & G. Wilson (Eds.), *Binge eating: Nature, assessment, and treatment* (pp. 317–360). New York: Guilford Press.
- Fairburn, C. G., Cooper, Z., Doll, H. A., Norman, P., & O'Connor, M. (2000). The natural course of bulimia nervosa and binge eating disorder in young women. *Archives of General Psychiatry*, 57, 659–665.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.
- Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. (1982). The Eating Attitudes Test: Psychometric features and clinical correlates. *Psychological Medicine*, 12, 871–878.
- Garner, D. M., Olmsted, M. P., & Polivy, J. (1983). The development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders*, 2, 15–34.
- Garrow, J. S., & Webster, J. (1985). Quetlet's index (W/H<sup>2</sup>) as a measure of fatness. *International Journal of Obesity*, 9, 147–153.
- Johnson, J. G., Cohen, P., Kasen, S., & Brook, J. S. (2002). Eating disorders during adolescence and the risk for physical and mental disorders during early adulthood. *Archives of General Psychiatry*, 59, 545–552.
- Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R., & Andrews, J. A. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology*, 102, 133–144.
- Lewinsohn, P. M., Striegel-Moore, R. H., & Seeley, J. R. (2000). Epidemiology and natural course of eating disorders in young women from adolescence to young adulthood. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39, 1284–1292.
- Mintz, L. B., O'Halloran, M. S., Mulholland, A. M., & Schneider, P. A. (1997). Questionnaire for eating disorder diagnoses: Reliability and validity of operationalizing DSM-IV criteria into a self-report format. *Journal of Counseling Psychology*, 44, 63–79.
- Newman, D. L., Moffitt, T. E., Caspi, A., Magdol, L., Silva, P. A., & Stanton, W. R. (1996). Psychiatric disorder in a birth cohort of young adults: Prevalence, comorbidity, clinical significance, and new case incidence from ages 11–21. *Journal of Consulting and Clinical Psychology*, 64, 552–562.
- Orjada, K. S., & Stice, E. (2003). *A controlled evaluation of an eating disturbance psychoeducational intervention for college students: A replication*. Manuscript in preparation.
- Puig-Antich, J., & Chambers, W. J. (1983). *Schedule for Affective Disorders and Schizophrenia for School-Age Children (6–18 years)*. Pittsburgh, PA: Western Psychiatric Institute.
- 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.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.
- Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association*, 90, 822–828.
- Sinharay, S., Stern, H. S., & Russell, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological Methods*, 6, 317–329.
- Spitzer, R. L., Williams, J. B., Gibbon, M., & First, M. B. (1990). *Structured clinical interview for DSM-III-R (SCID)*. Washington, DC: American Psychiatric Association.
- Stice, E. (2001). A prospective test of the dual pathway model of bulimic pathology: Mediating effects of dieting and negative affect. *Journal of Abnormal Psychology*, 110, 124–135.
- Stice, E., Cameron, R., Killen, J. D., Hayward, C., & Taylor, C. B. (1999). Naturalistic weight reduction efforts prospectively predict growth in relative weight and onset of obesity among female adolescents. *Journal of Consulting and Clinical Psychology*, 67, 967–974.
- Stice, E., Hayward, C., Cameron, R., Killen, J. D., & Taylor, C. B. (2000).

- Body image and eating related factors predict onset of depression in female adolescents: A longitudinal study. *Journal of Abnormal Psychology*, 109, 438–444.
- Stice, E., & Ragan, J. (2002). A controlled evaluation of an eating disturbance psychoeducational intervention. *International Journal of Eating Disorders*, 31, 159–171.
- Stice, E., Ragan, J., & Randall, P. (in press). Prospective relations between social support and depression: Differential direction of effects for parent and peer support? *Journal of Abnormal Psychology*.
- Stice, E., Telch, C. F., & Rizvi, S. L. (2000). Development and validation of the Eating Disorder Diagnostic Scale: A brief self-report measure of anorexia, bulimia, and binge-eating disorder. *Psychological Assessment*, 12, 123–131.
- Stice, E., Trost, A., & Chase, A. (2003). Healthy weight control and dissonance-based eating disorder prevention programs: Results from a controlled trial. *International Journal of Eating Disorders*, 33, 10–21.
- Thelen, M., Farmer, J., Wonderlich, S., & Smith, M. (1991). A revision of the Bulimia Test: The BULIT-R. *Psychological Assessment*, 3, 119–124.
- van Strien, T., Frijters, J. E., Bergers, G. P., & Defares, P. B. (1986). The Dutch eating behavior questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. *International Journal of Eating Disorders*, 5, 295–315.
- Weissman, M. M., & Bothwell, S. (1976). Assessment of social adjustment by patient self-report. *Archives of General Psychiatry*, 33, 1111–1115.
- Weissman, M. M., Prusoff, B. A., Thompson, W. D., Harding, P. S., & Myers, J. K. (1978). Social adjustment by self-report in a community sample and in psychiatric outpatients. *Journal of Nervous and Mental Disorders*, 166, 317–326.
- Williamson, D. A., Anderson, D. A., Jackman, L. P., & Jackson, S. R. (1995). Assessment of eating disordered thoughts, feelings, and behaviors. In D. B. Allison (Ed.), *Handbook of assessment methods for eating behaviors and weight-related problems: Measures, theory, and research* (pp. 347–386). Thousand Oaks, CA: Sage.
- Wilson, G. T., Heffernan, K., & Black, C. M. (1996). Eating disorders. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (pp. 541–571). New York: 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.

## Appendix

### SPSS Computer Code for Scoring the Eating Disorders Diagnostic Scale

The SPSS code for scoring the Eating Disorder Diagnostic Scale (EDDS) is provided below. Items are labeled edds1 through edds22. The EDDSDX variable provides tentative diagnoses of full threshold anorexia nervosa (1), full threshold bulimia nervosa (2), full threshold binge eating disorder (3), subthreshold anorexia nervosa (5), subthreshold bulimia nervosa (6), subthreshold binge eating disorder (7), and no diagnoses (0). This variable can easily be recoded to allow a focus on full threshold diagnoses or a particular diagnostic category. This code also generates an overall eating disorder symptom composite.

Originally, we recommended taking *z* scores of each item before averaging the items (see code below for EDDSZSYM). To do this, one must first save out the standardized scores for the EDDS items (zedds1–zedds21). However, we have subsequently found that the overall eating disorder symptom composite often evidences satisfactory internal consistency if one sums the raw items (see code below for EDDSSYM). Note, however, that the height and weight data and the birth control items are omitted from both symptom composite variables.

```
If edds5=0 and edds6=0 edds9=0.
If edds5=0 and edds6=0 edds10=0.
If edds5=0 and edds6=0 edds11=0.
If edds5=0 and edds6=0 edds12=0.
If edds5=0 and edds6=0 edds13=0.
If edds5=0 and edds6=0 edds14=0.
Compute eddsBMI=(Edds19 / 2.205)/((Edds20 / 39.37)
** 2).
If edds19 < 50 eddsBMI=999.
If edds20 < 40 eddsBMI=999.
RECODE eddsBMI (999=SYSMIS).
Compute feature=sum.1(edds9, edds10, edds11, edds12,
edds13).
Compute bedbing=999.
If edds5=0 bedbing=0.
If edds5=1 and edds6=1 and edds7 >=1 and feature
>=3 and edds14=1 bedbing=1.
```

```
If edds5=1 and edds6=1 and edds7 >=2 and feature
>=3 and edds14=1 bedbing=2.
Compute bulbing=999.
If edds5=0 bulbing=0.
If edds5=1 and edds6=1 and edds8 >=1 bulbing=1.
If edds5=1 and edds6=1 and edds8 >=2 bulbing=2.
RECODE bedbing bulbing (999=SYSMIS).
Compute compsum=999.
Compute
compsum=sum.2(edds15,edds16,edds17,edds18).
RECODE compsum (999=SYSMIS).
Compute compen=999.
If compsum < 1 compen=0.
If compsum >=1 compen=1.
If compsum >=2 compen=2.
RECODE compen (999=SYSMIS).
Compute wtshap=999.
If edds3 < 6 or edds4 < 6 wtshap=0.
If edds3 >=1 or edds4 >=1 wtshap=1.
If edds3 >=4 or edds4 >=4 wtshap=2.
Compute LowBMI=999.
If eddsBMI > 10 lowBMI=0.
If eddsBMI <= 18.5 LowBMI=1.
If eddsBMI <= 17.5 LowBMI=2.
Compute fearwt=999.
If edds2 <7 fearwt=0.
If edds2 >=1 fearwt=1.
If edds2 >=4 fearwt=2.
Compute amen=999.
If edds21 >=0 amen=0.
If edds21 >=1 amen=1.
If edds21 >=0 and edds22=1 amen=1.
If edds21 >=3 amen=2.
If edds21 >=0 and edds22=1 amen=2.
RECODE wtshap lowbmi fearwt amen (999=SYSMIS).
Compute eddsdx=999.
```

If edds1 < 8 eddsdx=0.  
 If bedbing=1 and compen=0 eddsdx=7.  
 If bulbing >=1 and compen >=1 and wtshap >=1  
 eddsdx=6.  
 If bedbing=2 and compen=0 eddsdx=3.  
 If bulbing =2 and compen =2 and wtshap =2 eddsdx=2.  
 If lowBMI >=1 and amen >=1 eddsdx=5.  
 If lowBMI =2 and amen =2 eddsdx=1.  
 RECODE eddsdx (999=SYSMIS).  
 Compute eddszsym=mean.8(zedds1 to zedds18, zedds21).  
 Compute eddssym=sum.8(edds1 to edds18, edds21).

Because individuals with anorexia nervosa often show limited insight, investigators often exclude the requirements that they (a) report an intense fear of gaining weight or becoming fat even though underweight or (b) report a disturbance in the way they experience their body shape and/or weight, report undue influence of weight and/or shape on self-evaluation,

or deny the seriousness of their low body weight. In the Eating Disorder Examination (EDE), interviewers are instructed to override asymptomatic responses if the patient shows limited insight. The code above takes this approach if the participant meets the criteria for low body weight and amenorrhea. Individuals who use the EDDS are free to alter the code if they wish to include these other symptoms.

A score of 4 or greater on EDDS Item 1 is indicative of body image disturbance and a denial of the seriousness of low body mass if the participant meets the low-body-mass criteria. A score of 4 or greater on EDDS Item 2 is indicative of an intense fear of gaining weight or becoming fat if the participant meets the low-body-mass criteria. A score of 4 or greater on either EDDS Item 3 or 4 is indicative of undue influence of weight and/or shape on self-evaluation.

Received April 7, 2003  
 Revision received July 23, 2003  
 Accepted October 29, 2003 ■

## ORDER FORM

Start my 2004 subscription to *Psychological Assessment!* ISSN: 1040-3590

\_\_\_ \$51.00, APA MEMBER/AFFILIATE \_\_\_\_\_  
 \_\_\_ \$104.00, INDIVIDUAL NONMEMBER \_\_\_\_\_  
 \_\_\_ \$254.00, INSTITUTION \_\_\_\_\_  
*In DC add 5.75% / In MD add 5% sales tax* \_\_\_\_\_  
**TOTAL AMOUNT ENCLOSED** \$ \_\_\_\_\_

**Subscription orders must be prepaid.** (Subscriptions are on a calendar year basis only.) Allow 4-6 weeks for delivery of the first issue. Call for international subscription rates.



AMERICAN  
 PSYCHOLOGICAL  
 ASSOCIATION

**SEND THIS ORDER FORM TO:**  
 American Psychological Association  
 Subscriptions  
 750 First Street, NE  
 Washington, DC 20002-4242

Or call (800) 374-2721, fax (202) 336-5568.  
 TDD/TTY (202) 336-6123.  
 For subscription information, e-mail:  
**subscriptions@apa.org**

Send me a **FREE Sample Issue**  
 Check enclosed (make payable to **APA**)  
**Charge my:**  VISA  MasterCard  American Express

Cardholder Name \_\_\_\_\_  
 Card No. \_\_\_\_\_ Exp. Date \_\_\_\_\_  
 \_\_\_\_\_  
 Signature (Required for Charge)

**BILLING ADDRESS:** \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip \_\_\_\_\_  
 Daytime Phone \_\_\_\_\_  
 E-mail \_\_\_\_\_

**SHIP TO:**  
 Name \_\_\_\_\_  
 Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip \_\_\_\_\_  
 APA Member # \_\_\_\_\_ PASA14