Published in final edited form as:

J Consult Clin Psychol. 2008 December ; 76(6): 1022–1033. doi:10.1037/a0013887.

Subtyping Women with Bulimia Nervosa Along Dietary and Negative Affect Dimensions: Further Evidence of Reliability and Validity

Eric Stice,
Oregon Research Institute

Cara Bohon,
Department of Psychology, University of Oregon

C. Nathan Marti, and
Department of Psychology, University of Texas at Austin

Kathryn Fischer
Department of Psychology, University of Texas at Austin

Abstract

Studies have found that individuals with bulimia nervosa can be classified into dietary and dietary-negative affect subtypes and that the latter exhibit greater eating pathology, psychiatric comorbidity, functional impairment, a more protracted clinical course, and a worse treatment response. This report describes two prospective studies which found that young women with threshold ($n = 48$) and subthreshold ($n = 83$) bulimic pathology can be classified into dietary and dietary-negative affect subtypes, that two subtyping approaches produced similar results ($M \kappa = .94$), that the subtyping distinction showed 4-week test-retest reliability ($\kappa = .61$), and that the dietary-negative affect subtype showed greater eating pathology, emotional distress, functional impairment, treatment seeking, and lower likelihood of recovery over 6-month and 3-year follow-ups than the dietary subtype. The dieting-negative affect subtyping distinction evidenced greater test-retest reliability and concurrent and predictive validity than did the purging-nonpurging subtyping distinction. The additional evidence for the reliability and validity of this subtyping scheme, particularly the prognostic utility, suggests it is worth additional inquiry.

Keywords

bulimia nervosa; subtyping; dieting; negative affect

Theorists have suggested that both dieting and negative affect play a role in the development and maintenance of bulimia nervosa. Polivy and Herman (1985) argue that diet-induced hunger increases the likelihood of binge eating and that a reliance on cognitive controls over eating leaves dieters vulnerable to overeating when these cognitive processes are disrupted. Dieting

Correspondence should be addressed to Eric Stice, who is at Oregon Research Institute, 1715 Franklin Blvd., Eugene, Oregon, 97403. estice@ori.org.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/ccp

may also maintain bulimic pathology, as people who binge eat may redouble their dietary efforts to limit weight gain, which increases the risk of further binge eating (Fairburn et al., 2003). It has also been theorized that people binge eat in an effort to reduce negative affect (McCarthy, 1990). Negative affect might also maintain the binge-purge cycle because bulimic behaviors result in shame, guilt, and anxiety, which may prompt people to attempt to escape these emotions through further binge eating and purging (Waller, 2002). The dual pathway model of bulimia nervosa (Stice, 2001) proposes that both dieting and negative affect play a role in promoting bulimic pathology, in that individuals may initiate binge eating because of either dieting or negative affect, or a combination of these factors.

Adolescent girls from epidemiologic studies who report dieting, relative to their non-dieting peers, are at increased risk for future onset of bulimic symptoms (Field et al., 1999; Neumark-Sztainer et al., 2006; Stice & Agras, 1998; Stice, Killen, Hayward, & Taylor, 1998), onset of threshold and subthreshold bulimia nervosa (Killen et al., 1996) and increases in bulimic symptoms (Johnson & Wardle, 2005; Stice, 2001; Wertheim et al., 2001). Self-reported dieting predicted bulimic symptom persistence in one study (Stice & Agras, 1998), but not another (Fairburn et al., 2003). Negative affect has likewise predicted onset of bulimic symptoms (Field et al., 1999; Stice & Agras, 1998) and bulimic pathology (Killen et al., 1996) and future increases in bulimic symptoms (Cooley & Toray, 2001; Stice, 2001). Negative affect predicted bulimic symptom persistence in one study (Fairburn et al., 2003), but not another (Stice & Agras, 1998).

One hypothesis that follows from these two accounts of bulimia nervosa is that there may be subtypes of this disorder that conform to the dietary restraint and negative affect models (Stice & Agras, 1999). It is important to search for reliable subtypes of psychiatric disorders because of the implications for nosology. In addition, if the subtypes show differential treatment response and clinical course, it may suggest that there are different risk and maintenance factors for the two variants of the disorder, which could have implications for treatment planning. Identification of reliable subtypes could facilitate the development of tailored treatments that are more effective. Researchers have identified reliable subtypes of alcoholism that have been found to have distinct etiologic risk factors (e.g., unique genetic influences), psychiatric comorbidity, treatment utilization, clinical course, and response to treatments (e.g., Penick et al., 1999; Pettinati, 2001).

Studies have found that individuals with bulimia nervosa can be subtyped along dietary and negative affect dimensions. Using cluster analysis with 265 treatment-seeking women with bulimia nervosa, Stice and Agras (1999) identified a dietary-negative affect subtype (38%) and a pure dietary subtype (62%) in split halves of the sample; relative to the dietary subtype, the dietary-negative affect subtype reported greater weight, shape, and eating concerns, functional impairment, comorbid mood, anxiety, eating, impulse control, and personality disorders, and poorer response to cognitive-behavioral treatment. Grilo, Masheb, and Berman (2001) replicated the evidence of a dietary-negative affect subtype (56%) and a pure dietary subtype (44%) among 48 treatment-seeking women with bulimia nervosa, and found that the former reported greater weight, shape, and eating concerns, and body dissatisfaction and that various subtyping approaches produced converging results. Stice and Fairburn (2003) found evidence of dietary-negative affect (56%) and dietary (44%) subtypes among 82 young women from a community-recruited natural history study of bulimia nervosa and found that the former reported more frequent binge eating, compensatory behaviors, greater weight, shape, and eating concerns, functional impairment, treatment seeking, mood and anxiety disorders, and greater persistence of binge eating over a 5-year follow-up. Chen and Le Grange (2007) replicated the evidence of dietary-negative affect (62%) and dietary (38%) subtypes with 80 treatment-seeking adolescent girls and boys with bulimia nervosa and found that the former showed

greater eating and weight concerns, psychiatric comorbidity, and lower rates of abstinence from compensatory behavior after treatment.

This subtyping distinction has also emerged with individuals with binge eating disorder. Stice, Agras, Telch, Halmi, Mitchel, and Wilson (2001) identified dietary-negative affect (63%) and dietary (37%) subtypes among 159 treatment-seeking women with binge eating disorder; the former reported significantly greater binge eating, weight, shape, and eating concerns, lifetime psychiatric disorders, functional impairment, and lower abstinence rates following treatment than the latter. Grilo, Masheb, and Wilson (2001) identified dietary-negative affect (33%) and dietary (67%) subtypes among 101 treatment-seeking women and men with binge eating disorder, found that the former reported greater weight, shape, and eating concerns, body dissatisfaction and impulsivity, and mood disorders than the dietary subtype, and found that this subtyping distinction showed 4-week test-retest reliability; 82% were similarly classified on the two occasions ($\kappa = .55$).

This subtyping distinction has also emerged with individuals with subthreshold eating pathology. Grilo (2004) identified dietary-negative affect (43%) and dietary (57%) subtypes among 137 female psychiatric inpatients with eating disorder features and found that the former reported greater binge eating frequency, eating pathology, body dissatisfaction, depression, suicidal tendencies, and personality disturbances. Chen and Le Grange (2007) identified dietary-negative affect (65%) and dietary (35%) subtypes among 149 outpatient adolescent girls and boys from an eating and weight disorder treatment program and found that the former reported significantly greater binge eating, and eating, weight, and shape concerns.

In four studies the dieting-negative affect subtyping distinction showed stronger concurrent and predictive validity than the purging-nonpurging subtyping distinction from DSM-IV (American Psychiatric Association, 1994). This latter approach posits that individuals with bulimia nervosa who use vomiting or laxatives/diuretics as their primary compensatory behavior are qualitatively different than those who primarily use fasting and excessive exercise for compensatory purposes. Stice and Fairburn (2003) found evidence of concurrent and predictive validity for 14 of the 16 validation variables for the dietary-negative affect distinction, but for only 2 of the 16 variables when participants who reported purging were compared to those who did not. Grilo (2004) found evidence of concurrent and predictive validity for 14 of the 16 validation variables for the dietary-negative affect distinction, but for only 2 of the 16 variables when those reporting purging behavior were compared to those who did not. Chen and La Grange (2007) found evidence of concurrent and predictive validity for 5 of 9 validation variables for the dietary-negative affect distinction, relative to 2 of the 9 variables when those reporting frequent purging were compared to those who did not. Parenthetically, Grilo, Masheb, and Wilson (2001) found that the dietary-negative affect subtype differed significantly from the dietary subtype on 7 of the 16 validation variables, but that subtypes created based on the presence or absence of major depression differed only on 2 of the 16 variables.

The evidence that the dietary-negative affect subtype shows elevations on numerous eating pathology measures might be interpreted as suggesting that these subtypes capture a severity continuum, rather than represent true latent subtypes. However, three factors seem to argue against this interpretation. First, it is not always the case that the dietary-negative affect subtype shows elevations in core bulimic symptoms relative to the dietary subtype. For example, in the largest study to date, the two subtypes differed on frequency of laxative abuse and weight, shape, and eating concerns, but not on frequency of binge eating, vomiting, or diuretic abuse (Stice & Agras, 1999). Second, the variance explained by bulimic symptom variables is typically small relative to that explained by negative affect variables. For instance, in Stice and Agras (1999) eating disorder symptoms explained an average of 7% of the variance between

subtypes, whereas the negative affect measures explained an average of 55% of the variance. Third, Grilo, Masheb, and Wilson (2001) found that subtypes of binge eating disorder patients based on severity of binge eating differed on only 1 of the 16 validation variables, whereas those based on the dietary-negative affect distinction differed on 7 of these 16 variables.

Collectively, these studies provide considerable evidence that the dietary-negative affect subtyping distinction is reliable, in that similar results emerged from eight independent studies that examined treatment-seeking and community samples of individuals with bulimia nervosa, a mix of various eating disorders, binge eating disorder, and only eating disorder features. Moreover, similar results emerged when split half-replication was employed, when different cluster analytic algorithms were used, and when participants were subtyped repeatedly over time. These studies also provide considerable evidence for the validity of this distinction, in that the dietary-negative affect subtype typically evidenced greater eating pathology, functional impairment, psychiatric comorbidity, a poorer response to treatment, and a more protracted clinical course, relative to the dietary subtype. Further, this subtyping distinction had greater concurrent and predictive validity than the purging-nonpurging subtyping distinction. These findings suggest that dietary restraint is a central feature of bulimia nervosa, but that negative affect occurs in only a subset of cases and that the combination of dietary restraint and depressive affect signals a more severe variant of this disorder that is more difficult to treat and shows greater chronicity. Theoretically, elevations in negative affect that persist over time cause people to engage in maladaptive behaviors, such as binge eating and compensatory behaviors, to reduce negative affect, increase positive affect, or distract themselves from emotional distress (Stice & Agras, 1999). The confluence of dietary restraint and affective disturbances may be particularly problematic because both increase the odds of persistent binge eating, which may result in elevated symptoms, functional impairment, treatment utilization, a longer clinical course, and a worse treatment response.

The mounting evidence for the reliability and validity of this subtyping distinction suggests that additional studies should investigate this subtyping distinction. This is particularly vital because of the call for empirical data to inform the revision of DSM (Walsh, 2007). Thus, the first aim was to replicate this subtyping scheme in two independent samples to provide further evidence of the reliability of this distinction. Extending prior studies that have typically examined individuals with full threshold diagnoses, we examined participants with threshold and subthreshold bulimia nervosa. This is important because half of those seeking treatment for eating pathology do not meet full diagnostic criteria for anorexia and bulimia nervosa (Fairburn & Harrison, 2003; Fisher, Schneider, Burns, Symons, & Mandel, 2001) and because subthreshold bulimic pathology is associated with current and future functional impairment, emotional distress, medical problems, and treatment seeking (Lewinsohn et al., 2000; Stice, Marti, Spoor, Presnell, & Shaw, 2008; Striegel-Moore et al., 2003). Indeed, scholars have called for further research on subthreshold levels of eating pathology because this constitutes a large portion of Eating Disorder Not Otherwise Specified diagnoses (Fairburn & Harrison, 2003; Wilson, Becker, & Heffernan, 2003). The second aim was to investigate the test-retest reliability of the dietary-negative affect distinction because only one prior study examined this question. The third aim was to further investigate the concurrent validity of this subtyping distinction by replicating the evidence that the dietary-negative affect subtype shows greater eating pathology and related disturbances, social impairment, and mental health treatment than the dietary subtype. The fourth aim was to generate additional evidence of predictive validity by testing whether the subtyping scheme predicted course of illness and change in body mass, as this would have implications for treatment planning. The fifth aim was to test whether the dietary-depressive subtyping distinction has greater concurrent and predictive validity than the purging-nonpurging subtyping distinction. A key question regarding a new subtyping distinction is whether it is an improvement over extant schemes (Wonderlich, Crosby, Mitchell, & Engel, 2007).

Study 1

Participants and Procedures

Participants with threshold ($n = 41$) or subthreshold ($n = 44$) bulimia nervosa ($N = 85$) were recruited for a treatment trial. Advertisements invited women with binge eating and compensatory behavior problems to participate in a program promoting a healthier way to “achieve a healthy weight, increase body satisfaction, and break the binge-purge cycle.” For subthreshold bulimic pathology, participants were required to report at least 4 binge eating episodes and 4 compensatory behavior episodes in the past month to be eligible. Those who had received treatment for an eating disorder in the past month, a low weight ($BMI < 19$), a physical or medical condition hindering their ability to make dietary or exercise changes (e.g., an injury), current suicidal ideation, or a lifetime history of schizophrenia or bipolar disorder were excluded. The resultant sample ranged in age from 18–55 years ($M = 21$; $SD = 5.3$) and was composed of 9% Asians, 2% Blacks, 22% Hispanics, 53% Caucasians, and 13% who specified other or mixed racial heritage. Participants were randomized to either the Healthy Weight treatment or to a waitlist condition. Participants completed a survey and an interview at baseline, mid-treatment, posttest, and 3-month follow-up. They were paid \$10–\$20 for completing each assessment. See Burton and Stice (2006) for further details about the design and intervention content. For this study and the one described subsequently, participants provided informed consent and Institutional Review Board approval was secured.

Measures

Dietary restraint—The Cognitive Restraint scale (Stunkard & Messick, 1985) assesses dietary behaviors designed to produce weight loss or maintenance (sample item: *I count calories as a conscious means of controlling my weight*). Response options ranged from 1 = *never* to 5 = *always*. Items were summed ($\alpha = .78$ at baseline). This scale has shown internal consistency ($\alpha = .85 - .93$), temporal reliability (1-month test-retest $r = .98$), and increases in response to low-calorie weight loss diet interventions (French, Jeffery, & Wing, 1994; Stunkard & Messick, 1985; Williamson et al., 2007), but shows a modest relation to objectively measured caloric intake (Stice, Fisher, & Lowe, 2004).

Negative affect—The sadness, anxiety, and guilt scales from the Positive Affect and Negative Affect Scale-Revised (PANAS; Watson & Clark, 1992) asks participants to report the extent to which they had felt various negative emotions using a response format ranging from 1 = *very slightly or not at all* to 5 = *extremely*. Responses were averaged. In the present sample, the sadness, anxiety, and guilt scales showed internal consistency ($\alpha = .94, .91, \& .92$ respectively) and 1-week test-retest reliability in the waitlist control group ($r = .83, .84, \& .86$ respectively). The PANAS has shown internal consistency, convergent validity, sensitivity to detecting intervention effects, and predictive validity for bulimic symptom onset (Stice et al., 2008; Watson & Clark, 1992).

Depressive symptoms—The Beck Depression Inventory (BDI, Beck et al., 1988) asks participants to report the severity of depressive symptoms (0 = no symptom present to 3 = severe symptom present). Items were summed ($\alpha = .89$ at baseline). The BDI has shown internal consistency ($\alpha = .73 - .95$), test-retest reliability ($r = .60 - .90$), sensitivity to detecting intervention effect, and convergence with clinician ratings of depressive symptoms ($M r = .75$; Beck et al., 1988).

Bulimic pathology—The Eating Disorder Diagnostic Interview (EDDI; Stice et al., 2008), a semi-structured interview that was adapted from the Eating Disorder Examination (Fairburn & Cooper, 1993), assessed DSM-IV eating disorder symptoms over the past year or since the last assessment. Based on these items, participants were diagnosed with threshold or

subthreshold bulimia nervosa at each assessment. We also calculated continuous measures of the frequency of binge eating and compensatory behaviors (vomiting, laxative use, diuretic use, fasting, and compensatory exercise) over the past month at baseline, as well as the degree of weight and shape concerns (using a response format ranging from 0 = *no importance* to 6 = *nothing is more important than shape/weight in self-evaluation*). This diagnostic interview has shown 1-week test-retest reliability ($\kappa = .96$) and inter-rater agreement ($\kappa = .86$) for eating disorder diagnoses, sensitivity to detecting intervention effects, and predictive validity for onset of major depression (Burton & Stice, 2006; Stice, Burton, & Shaw, 2004; Stice et al., 2008).

Hunger—The Hunger scale (Stunkard & Messick, 1985) assessed subjective perception of hunger. Items used varied response options and were summed ($\alpha = .82$ at baseline). This scale has shown internal consistency ($\alpha = .63-.73$) and test-retest reliability ($r = .75$; Bond, McDowell, & Wilkinson, 2001).

Disinhibition—The Disinhibition scale (Stunkard & Messick, 1985) assessed disinhibited eating. Items used varied response options and were summed ($\alpha = .71$ at baseline). This scale has shown internal consistency ($\alpha = .60-.78$) and test-retest reliability ($r = .86$; Bond et al., 2001).

Emotional Eating—The Negative Affect scale from Eating Self-Efficacy Scale (ESES; Glynn & Ruderman, 1986) asks participants about the likelihood that they would overeat in response to various affective states. Responses ranged from 1 = *no difficulty controlling eating* to 7 = *most difficulty controlling eating*. Items were averaged ($\alpha = .93$ at baseline). The ESES has shown internal consistency ($\alpha = .94$), test-retest reliability ($r = .70$), and construct validity (Glynn & Ruderman, 1986).

Impulsivity—The Barratt Impulsivity Scale (Patton, Stanford, & Barrett, 1995) asks participants to report the extent to which various descriptions regarding impulsivity apply to them using a response format ranging from 1 = *rarely/never* to 4 = *almost always/always*. Items were summed ($\alpha = .86$ at baseline). This scale has shown internal consistency ($\alpha = .79 - .83$), 2-week test-retest reliability ($r = .88$), and discriminates between psychiatric patients and controls (Patton et al., 1995; Suris, Borman, Lind, & Kashner, 2004).

Mental health treatment—Participants answered the question: *Have you received treatment of any kind, with a doctor, a counselor, in a hospital, or with anyone else, to address problems such as depression, eating or body-image concerns, feeling anxious, substance abuse concerns, or any other emotional issues in the past year?*

Functional Impairment—Items from the Social Adjustment Scale (Weissman & Bothwell, 1976) assessed psychosocial functioning in the family, peer group, school, and work spheres. Items used varied response options and were averaged ($\alpha = .77$ at baseline). The SAS has shown convergent validity with clinician and collateral ratings ($M r = .72$), discriminates between psychiatric patients and controls, and is sensitive to treatment effects (Weissman & Bothwell, 1976). This adapted scale has shown internal consistency ($\alpha = .77$), 1-week test-retest reliability ($r = .83$) and sensitivity to intervention effects (Stice et al., 2008).

Results and Discussion

Participants with threshold or subthreshold bulimia nervosa reported an average of 11 binge eating episodes and 23 compensatory behavior episodes over the past month. These data suggest that participants were experiencing clinically meaningful bulimic pathology.

To assess whether participants with bulimic pathology could be subtyped along dieting and negative affect dimensions, their scores on the Cognitive Restraint scale, the PANAS sadness, guilt, and anxiety scales, and the BDI were submitted to a *k-means* cluster analysis using SPSS 15. Variables were converted to *z* scores prior to being cluster analyzed to eliminate the impact of disparate ranges that influence distance measures in cluster analysis. *K-means* cluster analysis selects *k* participants, where *k* is the number of cluster requested and then iteratively clusters participants into one of these groups based on squared Euclidean distances. After each case is assigned, the cluster center is updated before the next iteration. In the final step of the algorithm, each participant is reassigned to the nearest of the updated cluster centers, yielding the final clusters. Based on the fact that eight prior studies have identified dietary and dietary-negative affect subtypes, we set the number of expected clusters to two *a priori*.

The cluster analysis identified a pure dieting group ($n = 46$; 55%) and a dietary-negative affect group ($n = 38$; 45%). ANOVA models indicated that the dietary-negative affect group reported significantly more self-reported dieting, depression, sadness, anxiety, and guilt than the dietary group (Table 1), with the latter four differences accounting for considerably more variance than the former difference. The dietary and dietary-negative affect subtypes did not differ significantly on age, ethnicity, maternal education, paternal education, or Body Mass Index ($BMI = Kg/M^2$, based on directly measured weight and height) (all *p*-values $>.05$).

To investigate the concurrent validity of this subtyping distinction, we tested whether the groups differed at baseline on bulimic symptoms and variables commonly associated with eating pathology. The dietary-negative affect subtype reported significantly more binge eating and compensatory behavior in the past month, overvaluation of weight and shape, disinhibition, hunger, emotional eating, mental health treatment, and functional impairment than the dietary subtype (Table 1). The groups did not differ on impulsivity.

With regard to predictive validity, we tested whether the dietary-negative affect subtype would be less likely to show recovery from bulimic pathology during the 6-month follow-up period (i.e., not meet criteria for threshold or subthreshold bulimia nervosa at any of the post-baseline assessments in this study) and whether they were more likely to show increases in BMI over follow-up. We controlled for intervention condition in the analyses to adjust for the fact that participants were randomized to a treatment or assessment-only control condition. A logistic regression model indicated that the dietary-negative affect subtype was less likely to show recovery from bulimic pathology relative to the dietary subtype ($p < .05$, Odds ratio [*OR*] = 2.28); 29% of the dietary-negative affect participants showed recovery from bulimic pathology, relative to 46% of the dietary subtype (Table 1). We used a growth mixture model to test whether the subtypes showed differential change in BMI over the follow-up period; participants were nested within time points and the intercepts and slopes were treated as random effects. The model contained time in months, cluster, the cluster $\times$ time interaction, and intervention condition. As indicated in Table 1, there were no differences in change in BMI across groups ($t(1, 122) = .04, p = .97$).¹

We thought it prudent to use cluster analysis rather than latent profile analyses (LPA) because experts warn about the impact of small sample sizes on finite mixture models on which LPA are based (McLachlan & Peel, 2000). Nonetheless, we conducted exploratory LPA to test whether it produced effects were similar to those produced by the cluster analysis and to test whether a 2-cluster solution was superior to 1- and 3-cluster solutions. The 2-cluster model

¹Pseudo R^2 values were computed using the formula proposed by Kreft and de Leeuw (1998) in which variance explained is the difference between the unrestricted model and the restricted model divided by the variance of the unrestricted model. We defined the unrestricted model as one containing the time, group, and intervention condition effects and the restricted model as containing those terms with the addition of the group $\times$ time interaction, thus defining the derived value as the additional variance explained by the addition of the interaction term. In the event of negative values, we set variance equal to zero.

was confirmed by LPA using mixture models implemented in Mplus Version 5 (Muthén & Muthén, 2006). We used three criteria to establish the best fitting model: (1) the Lo-Mendell-Rubin likelihood ratio test (LMR-LRT), (2) the Bayesian Information Criterion (BIC), and (3) an entropy measure (for models with more than 1 cluster). LMR-LRT compares a model with k clusters with a model containing $k-1$ clusters and provides a significance test for the null hypothesis that there is no difference in the log likelihood for the $k-1$ and k -models. BIC values can be used to compare any pair of models based on the same set of data (Singer & Willet, 2003) whereby smaller values indicate greater parsimony; we use the criteria of a decrease in 10 suggested by Raftery (1995) as an indication of a very strong decrease. Entropy measures classification accuracy using a range of 0 to 1 with values closer to 1 indicating superior classification.

In comparing 2-cluster model (BIC = 1748.65, Entropy = .78) with the 1-cluster model (BIC = 1816.36), we rejected the null hypothesis that there was no improvement in a 2-cluster solution (LMR-LRT = 90.95, $p < .05$). In comparing the 2-cluster model (BIC = 1748.65, Entropy = .78) with the 3-cluster model (BIC = 1740.90, Entropy = .80), there was insufficient evidence that the 3-cluster model increased the parsimony of the model (LMR-LRT = 33.16, $p = .36$). Thus, we rejected the 3-cluster model in favor of the 2-cluster model.

There was 94% agreement between the *k-means* and LPA analyses ($\kappa = .88$) with regard to cluster membership. Further, eight out of the nine validation variables that were significant in the *k-means* model (see Table 1) were also significant using LPA groups. The significant effect for recovery from bulimic pathology became marginally significant ($p = .034$ to $p = .073$), with the *OR* changing from 2.3 to 2.1. The two non-significant effects remained non-significant. Thus, there was high agreement between the subtyping solutions generated by cluster analysis and LPA.

To examine the concurrent and predictive validity for the purging-nonpurging subtype we compared participants with bulimic pathology who reported at least two episodes of vomiting for weight control or use of laxatives or diuretics for weight control during the past month ($n = 43$) to those who did not use these compensatory behaviors ($n = 44$) on the concurrent and predictive validation variables (Table 1). The purging subtype reported significantly more binge eating and compensatory behaviors in the past month than the non-purging subtype, but did not differ on overvaluation of weight and shape. The purging subtype also reported significantly greater disinhibition, emotional eating, and mental health treatment than the non-purging subtype, but the groups did not differ on hunger, impulsivity, or functional impairment. The percent of the participants in the purging (30%) and nonpurging (43%) subtypes that showed persistent recovery from binge eating and compensatory behaviors was statistically significant ($\beta = .86$, $p = .034$, *OR* = 2.37). Change in BMI over the study period did not differ significantly across purging and nonpurging subtypes ($t(1, 124) = -0.03$, $p = .97$). Thus, whereas the dietary-negative affect subtypes differed on 9 of the 11 validation variables, the purging-nonpurging subtypes differed on only 5 of these variables; a difference that was statistically significant per a binomial test ($p < .01$). In addition, the average variance explained between the purging-nonpurging subtypes was smaller than the average variance explained between the dietary-negative affect subtypes for the continuous variables (3% vs. 20% respectively).

In sum, the dietary-negative affect subtyping distinction emerged in this clinical sample of young women with threshold or subthreshold bulimia nervosa, which is a novel contribution to the literature that extends the evidence-base for this subtyping distinction. This is important given that most individuals who seek treatment do not meet criteria for full threshold bulimia nervosa (e.g., Fisher et al., 2001). The fact that two different subtyping approaches produced results that showed high agreement provided further evidence for the reliability of this

distinction. With regard to concurrent validity, the dietary-negative affect subtype reported significant elevations in bulimic symptoms, related disturbances (disinhibition, hunger, and emotional eating), functional impairment, and mental health treatment. With regard to predictive validity, the dietary-negative affect subtype was significantly less likely to show persistent recovery from bulimic pathology over the 6-month follow-up period (though this effect was only marginal in the LPA analyses). Results also indicated that the dietary-negative affect distinction showed stronger validity than did the purging-nonpurging subtyping distinction; there were significantly fewer differences between the purging and nonpurging subtypes on the validation variables and this subtyping distinction accounted for less variance in the validation variables than was the case for the dietary-negative affect subtyping distinction.

Study 2

Participants and Procedures

Participants were drawn from an eating disorder prevention trial involving 481 high-risk adolescent girls with body dissatisfaction (M age = 17.0, SD = 1.4) (Stice et al., 2008). Although assessors attempted to exclude adolescents who met current criteria for anorexia nervosa, bulimia nervosa, and binge eating disorder, analyses revealed that 7 participants met criteria for bulimia nervosa at baseline. Another 39 participants met criteria for subthreshold bulimia nervosa at baseline. For subthreshold bulimia nervosa, we required participants to report at least 6 binge eating episodes with loss of control and 6 compensatory behavior episode over a 3-month period (an average of twice monthly for each, as opposed to twice weekly for a full threshold diagnosis), and to report that weight and shape was definitely an aspect of self-evaluation. The present report focused on the 46 participants who met criteria for threshold or subthreshold bulimia nervosa at baseline. This sample was 4% Asian/Pacific Islander, 13% Black, 11% Hispanic, 61% Caucasian, and 11% who specified other or mixed racial heritage. Participants were recruited from high schools and a university using direct mailings, flyers, and leaflets inviting females “*between the ages of 14 and 19 to participate in a research project evaluating interventions aimed at helping young women to accept their bodies.*” Participants were randomized to a dissonance-based thin-ideal internalization intervention, a healthy weight management intervention, an expressive writing control intervention, or an assessment-only control condition. They were paid for completing a survey and an interview at pretest, posttest, and at 6-month, 1-year, 2-year, and 3-year follow-ups. See Stice, et al, (2008) for details about interviewer and facilitator training and intervention content.

Measures

Dietary restraint—The Restrained Eating scale from the Dutch Eating Behavior Questionnaire (DEBQ) assessed dieting (van Strien, Frijters, van Staveren, Defares, & Deurenberg, 1986). Participants indicate the frequency of dieting behaviors using a response format ranging from 1 = *never* to 5 = *always*. Items were averaged (α = .77 at baseline). This scale has shown internal consistency (α = .95), 2-week test-retest reliability (r = .82), and predictive validity for bulimic symptom onset, but shows a weak relation with objectively measured caloric intake (Stice, Fisher et al., 2004; van Strien et al., 1986).

Negative affect—Negative affect was assessed with the sadness, anxiety, and guilt scales from the PANAS (Watson & Clark, 1992). See Study 1 for a description of these scales and evidence of the reliability and validity. At baseline, the sadness, anxiety, and guilt scales were internally consistent (α = .93, .89, & .88 respectively).

Bulimic pathology—The EDDI (Stice et al., 2008) assessed DSM-IV eating disorder symptoms. On the basis of these items, participants were diagnosed with threshold or

subthreshold bulimia nervosa at each assessment, which allowed us to predict recovery from bulimic pathology. We also calculated continuous measures of the frequency of binge eating and compensatory behaviors (vomiting, laxative use, diuretic use, fasting, and compensatory exercise) over the past month at baseline, and the degree of weight and shape concerns. See Study 1 for evidence of the reliability and validity for this measure.

Body dissatisfaction—Eight items drawn from the Satisfaction and Dissatisfaction with Body Parts Scale (Berscheid, Walster, & Bohrnstedt, 1973) assessed satisfaction with body parts that are often of concern to females (e.g., stomach, thighs, and hips). Response options range from 1 = *extremely satisfied* to 6 = *extremely dissatisfied*. Items were averaged ($\alpha = .84$ at baseline). This scale has shown internal consistency ($\alpha = .94$), 3-week test-retest reliability ($r = .90$), and predictive validity for bulimic symptom onset (Stice et al., 2008).

Distress about body image—Six items asking participants to rate their level of distress about their body image concerns were drawn from the Body Esteem Scale for Adolescents and Adults (Mendelson, White, & Mendelson, 1996). Response options ranged from 1 = *strongly disagree* to 5 = *strongly agree*. Items were averaged ($\alpha = .83$ at baseline). This scale possessed internal consistency ($\alpha = .87$) and 1-week test-retest reliability ($r = .91$) in a pilot study ($N = 42$).

Emotional Eating—The Emotional Eating subscale from the DEBQ (van Strien Frijters, Bergers, & Defares, 1986) asks participants about the frequency of eating in response to negative emotions. Responses ranged from 1 = *never* to 5 = *very often*. Items were averaged ($\alpha = .95$ at baseline). This scale has demonstrated internal consistency ($\alpha = .86$ to $.97$) and shows convergent and predictive validity for binge eating onset (Stice, Presnell, & Spangler, 2002; van Strien et al., 1986).

Functional impairment—Items from the SAS (Weissman & Bothwell, 1976) assessed psychosocial functioning in the family, peer group, school, and work spheres. See Study 1 for evidence of the reliability and validity for this measure ($\alpha = .72$ at baseline).

Mental health and health treatment—Participants answered the following questions: *How often have you seen a psychologist, psychiatrist, counselor, or therapist because of mental health problems (depression, anxiety, etc.) in the last 6 months?* and *How often have you seen a doctor (physician) because of illness, injury, long-term health problems, or for regular check-ups in the last 6 months?* The mental health service and health service items showed 1-year test-retest reliability for assessment-only controls ($r = .89$ and $r = .82$ respectively).

Results and Discussion

Participants with threshold or subthreshold bulimia nervosa reported an average of 8 binge eating episodes and 11 compensatory behavior episodes over the past month. They also reported significantly greater mental health and health treatment and psychosocial impairment than those who were free of an eating disorder (Stice et al., 2008). Those who experienced bulimic pathology received an average of 12 hours of mental health treatment over the study period, compared to an average of 3.7 hours for non-afflicted participants. Those who experienced bulimic pathology made an average of 14 visits to physicians over the study, versus 9 visits for non-afflicted individuals. In addition, these adolescents reported significantly greater mental health treatment and functional impairment during this episode of eating pathology than was the case before they showed onset. Thus, it appears that these adolescents experienced clinically meaningful bulimic pathology.

To assess whether participants with bulimic pathology could be subtyped along dieting and negative affect dimensions, their scores on the DEBQ dietary restraint scale and the sadness, guilt, and anxiety scales from the PANAS were submitted to a *k-means* cluster analysis as in Study 1. Results identified a pure dieting group ($n=33$; 72%) and a dietary-negative affect group ($n=13$; 28%). ANOVA models indicated that the dietary-negative affect group showed significantly greater self-reported dieting, as well as significantly greater sadness, anxiety, and guilt, relative to the dietary group (Table 2), with the latter three effects accounting for more variance than the former effect. The dietary and dietary-negative affect subtypes did not differ on age, ethnicity, maternal education, paternal education, or BMI (based on directly measured weight and height) (all p -values $>.05$).

To investigate the test-retest reliability of the dietary-negative affect subtyping distinction we performed a second cluster analysis on these 46 participants using data that was collected 4-weeks after baseline. We used the same variables and again specified that we expected two clusters. Results identified a pure dieting group ($n = 29$; 63%) and a dietary-negative affect group ($n = 17$; 37%). A cross-tab analysis indicated high agreement between cluster membership on the two occasions ($\kappa = .61$), which is good according to Fleiss (1981). In total, 83% of the participants were classified in the same subtype on the two occasions separated by a period of 4-weeks. For comparison purposes, we also examined the 4-week test-retest reliability of the purging-nonpurging subtyping distinction. As in Study 1, participants who reported at least two episodes of vomiting for weight control or use of laxatives or diuretics for weight control during the past month were placed in the purging subtype, the remainder was placed in the non-purging subtype. This was done at baseline and a 4-week follow-up. Results indicated modest agreement between cluster membership on the two occasions ($\kappa = .29$), which is poor according to Fleiss (1981). In total, 74% of the participants were classified in the same subtype on the two occasions.

To investigate the concurrent validity of this subtyping distinction, we tested whether the groups differed at baseline on bulimic symptoms, body dissatisfaction, distress about body image, emotional eating, mental health treatment, and functional impairment. The dietary-negative affect subtype reported significantly more binge eating episodes in the past month, greater overvaluation of weight and shape, distress about their body image, emotional eating, and functional impairment than the dietary subtype, although the differences for compensatory behavior frequency, body dissatisfaction, and mental health treatment not significant (Table 2).

To investigate the predictive validity of this subtyping scheme, we tested whether the dietary-negative affect subtype would be less likely to show recovery from bulimic pathology period over the 3-year follow-up (i.e., not meet criteria for threshold or subthreshold bulimia nervosa over at least a 1-year period during follow-up) and whether there were more likely to show increases in BMI over follow-up. We controlled for intervention condition in the models to adjust for the fact that participants were randomized to prevention programs or control conditions. A logistic regression model indicated that the dietary-negative affect subtype was significantly less likely to show persistent recovery from bulimic pathology over the 3-year follow-up relative to the dietary subtype ($p = .011$, $OR = 6.19$); 31% of the dietary-negative affect participants recovered from bulimic pathology during follow-up, relative to 67% of the dietary subtype (Table 2).² A growth mixture model tested whether the subtypes showed

²Post hoc analyses tested whether baseline bulimic symptoms severity (reflecting a composite of past month frequency of binge eating and compensatory behaviors, and severity of overvaluation of weight/shape) predicted recovery from bulimic pathology in Study 1 and Study 2 to address the possibility that this accounted for the predictive validity findings. Results indicated that baseline symptom severity did not significantly predict recovery from bulimic pathology in Study 1 ($OR = 0.99$, $p = .169$) or in Study 2 ($OR = 0.98$, $p = .439$) after controlling for treatment condition. These results imply that baseline symptom severity is not responsible for the predictive validity effects for bulimic pathology recovery in the studies.

differential change in the rate of change in BMI over the 3-year follow-up. There were no differences in change in BMI across groups ($t(1, 175) = -.47, p = .64$).

As in Study 1, we conducted exploratory LPA to test whether it produced effects that were similar to those produced by the cluster analysis and to test whether a 2-cluster solution was superior to 1- and 3-cluster solutions. We used the same procedure described in Study 1. In comparing the 2-cluster model (BIC = 456.91, Entropy = .94) with the 1-cluster model (BIC = 502.45), we rejected the null hypothesis that there was no improvement in a 2-cluster solution. In comparing the 2-cluster model (BIC = 456.91, Entropy = .94) with the 3-cluster model (BIC = 463.94, Entropy = .84), there was insufficient evidence that the 3-cluster model increased the parsimony of the model (LMR-LRT = 33.16, $p = .51$). Thus, we rejected the 1- and 3-cluster models in favor of the 2-cluster model. Results identified a pure dieting group ($n=33$; 72%) and a dietary-negative affect group ($n=13$; 28%). There was 100% agreement between the k -means cluster analysis and the LPA cluster membership ($\kappa = 1.0$). Because classification was identical for these two analytic procedures, all effects reported in Table 2 are identical across the cluster analytic and LPA solutions.

To examine the concurrent and predictive validity for the purging-nonpurging subtype we compared participants who were assigned to the purging subtype ($n = 13$) to those assigned to the non-purging subtype ($n = 33$) at baseline on the concurrent and predictive validation variables listed in Table 2. The purging subtype reported significantly more binge eating in the past month, greater overvaluation of weight and shape, body dissatisfaction, distress about body image, and emotional eating than the nonpurging subtype, but the differences for compensatory behavior frequency, mental health treatment, and psychosocial impairment were not significant. In addition, the percentage of the participants in the purging (39%) and nonpurging (64%) subtypes that showed at least a 1-year recovery from bulimic pathology during the 3-year follow-up was not statistically significant ($n.s.$, $OR = 2.78$). The purging and nonpurging subtypes did not show differential change in BMI over the 3-year follow-up ($t(1, 175) = -0.17, p = .86$; Table 2). Thus, whereas the dietary-negative affect subtypes differed on 8 of the 11 validation variables, the purging-nonpurging subtypes differed on only 5 of these variables; a binomial test indicated that this was a statistically significant difference ($p < .05$). In addition, the average variance explained for the continuous variables was smaller for the purging-nonpurging distinction than for the dietary-negative affect distinction (8% vs. 21% respectively).

In sum, the dietary-negative affect subtyping distinction emerged in this community sample of adolescent women with threshold or subthreshold bulimia nervosa, providing further evidence of the reliability of this distinction. The dietary-negative affect subtyping distinction also showed 4-week test-retest reliability in this sample. The fact that two subtyping approaches produced results that showed perfect agreement provided additional evidence for the reliability of this distinction. With regard to concurrent validity, the dietary-negative affect subtype reported greater bulimic symptoms, related disturbances (distress about body image and emotional eating), and functional impairment. In terms of predictive validity, the dietary-negative affect subtype exhibited a lower likelihood of showing persistent recovery from bulimic pathology over the 3-year follow-up than the dietary subtype. Results also suggested that the dietary-negative affect distinction showed stronger test-retest reliability, concurrent validity, and predictive validity than did the purging-nonpurging subtyping distinction. There were significantly fewer differences between the nonpurging and purging subtypes on the validation variables and this subtyping distinction accounted for less variance in the validation variables than was the case for the dietary-negative affect subtyping distinction.

General Discussion

Although the relatively small samples sizes examined in the present report limit the confidence that can be placed in the robustness of this subtyping distinction, the results are consistent to those that have emerged in four previous studies of individuals with bulimia nervosa, two previous studies of individuals with binge eating disorder, and two studies of individuals with subthreshold eating disturbances (e.g., Chen & Le Grange, 2007; Grilo, Masheb, & Wilson, 2001; Stice & Agras, 1999; Stice & Fairburn, 2003). Thus, to date, this subtyping distinction has emerged in 10 separate samples conducted by 3 independent research groups. This subtyping distinction has emerged in samples of treatment seeking and community-recruited individuals with various degrees of bulimic and binge eating pathology, in samples of adults and adolescents, and in samples containing solely females and those containing a mix of both genders. These results imply that this subtyping distinction is reliable for a variety of inpatient, outpatient, and community populations. The use of a variety of measures in these studies also implies that this is a robust distinction. In addition, the present study and Grilo, Masheb, and Wilson (2001) found that this subtyping distinction showed 4-week test-retest reliability (82-83% of participants were classified similarly on two occasions). Similar results also emerged when split half-replication was employed (Stice & Agras, 1999) and when different clustering approaches were used in the present studies and in Grilo, Masheb, and Berman (2001).

The present results and those from past studies (Chen & Le Grange, 2007; Grilo, Masheb, & Berman, 2001; Stice & Agras, 1999; Stice & Fairburn, 2003) also provide substantial evidence for the validity of this subtyping distinction. The dietary-negative affect subtype typically reported greater eating disordered attitudes and eating disordered behaviors, elevations in related disturbances, such as body image concerns, treatment seeking, functional impairment, and elevated psychiatric comorbidity, relative to the dietary subtype. Although the difference in compensatory behavior frequency found in Study 1 did not replicate in Study 2, which appeared to be a product of the lower statistical power in Study 2, the findings replicated for the remaining 11 validation measures that were used in both studies. However, because not all validation measures were used in both studies, it not possible to draw conclusions regarding the replication for depressive symptoms, disinhibition, hunger, impulsivity, and body dissatisfaction. Importantly, three studies found that individuals with the dietary-negative affect subtype showed a poorer response to psychotherapeutic treatment and three studies found that they show a more protracted course of bulimic pathology, though the prospective effect from Study 1 was only marginal when we used LPA rather than cluster analysis to subtype participants. These findings suggest that this subtyping distinction has prognostic significance that may inform treatment development and planning. This is noteworthy because few consistent predictors of response to treatment and natural clinical course for bulimia nervosa have emerged in the literature (Fairburn, Agras, Walsh, Wilson, & Stice, 2004; Fairburn et al., 2003).

Another important finding from this and prior studies is that the dietary-negative affect subtyping distinction had significantly greater concurrent and predictive validity than does the purging-nonpurging subtyping distinction from DSM-IV. Indeed, it appears that the evidence-base is stronger for the dietary-negative affect distinction than for the purging-nonpurging distinction. Although studies have found that the purging subtype has shown greater eating pathology, mood and anxiety disorders, and personality disturbances, most of these studies suggested that there were more similarities than differences between these two subtypes (Garfinkel et al., 1996; McCann, Rossiter, King, & Agras, 1991; Tobin, Griffing, & Griffing, 1997; Walters et al., 1993). Gleaves, Lowe, Green, Cororve, and Williams (2000) conducted a taxometric analysis and found that purging and nonpurging subtypes of bulimia nervosa did not appear to represent distinct latent taxons, but rather differed in degree on the continuous

measures of eating pathology. Ironically, the present report appears to be the first to provide evidence that the purging-nonpurging distinction prospectively predicts clinical course of bulimic pathology, although this finding only emerged in only one of the two prospective studies. If revisions to the DSM will be guided by empirical evidence, it will be important for additional research groups to contrast the reliability and validity of the purging-nonpurging and the dietary-negative affect subtyping distinctions in independent samples. Although it might seem counter-intuitive to subtype eating disorders along a non-eating dimension, each of the eating disorders currently has non-behavioral symptoms (e.g., fear of weight gain and overvaluation of weight and shape). In addition, negative affect is a prominent feature of various psychiatric diagnoses (e.g., major depression). Finally, the concurrent and predictive validity evidence-base appears to be stronger for the dietary-negative affect distinction than for the purging, non-purging distinction.

Results from the present study and previous studies suggest that dietary restraint is a central feature of bulimia nervosa, that negative affect occurs in only a subset of cases, and that the combination of dietary restraint and depressive affect signals a more severe variant of this disorder that is more difficult to treat and shows a more protracted course. The fact that individuals in the dietary-negative affect subtype show elevations on two putative maintenance factors for bulimic pathology theoretically contributes to the greater persistence of this eating disturbance for this subtype. It is important to acknowledge that, with some exceptions (e.g., Herman & Mack, 1975), studies that used objective measures of caloric intake typically show that individuals with elevated scores on dietary restraint scales, relative to those with lower scores, do not consume significantly fewer calories during single eating episodes (Hetherington et al., 2000; Jansen, 1996; Stice, Fisher et al., 2004), multiple eating episodes (Martin et al., 2005; Rolls et al., 1997; Sysko, Walsh, Schebendach, & Wilson, 2005), or 2-12 week observation periods (Bathalon et al., 2000; Stice, Cooper, Schoeller, Tappe, & Lowe, 2007; Tuschl et al., 1990). These findings imply that dieting scales identify individuals who are attempting to curb an overeating tendency, rather than those who are successfully entering a negative energy balance necessary for weight loss. Thus, it may be the confluence of a negative affect and an overeating tendency that increases the risk for a greater severity of bulimic pathology that shows a more chronic course. The overeating tendency may be due to biological factors, such as abnormalities in dopamine-based reward circuitry (Bowirrat & Oscar-Berman, 2005) or weight suppression (Lowe et al., 2006). Irrespective, the elevated negative affect for this subtype relative to the dieting subtype, theoretically causes people to engage in maladaptive behaviors, such as binge eating and compensatory behaviors, to reduce this negative affect, increase positive affect, or simply distract themselves from the emotional distress (Stice & Agras, 1999). These findings imply that it might be beneficial to include a treatment component that addresses affective disturbances for the participants in the dietary-negative affect subtype, such as cognitive-behavioral therapy for mood disturbances.

Although the present studies used diagnostic interviews, a longitudinal follow-ups, and community recruited samples, they had several limitations. First, the relatively small sample sizes limited our ability to detect differences between groups. However, the limited power did not prevent us from detecting the moderate to large effects reported in these studies and this problem applied equally to the analyses comparing the dietary-negative affect subtypes and the purging subtypes. Second, the small samples also limit the confidence that can be placed in the findings. Yet, the fact that similar findings have emerged in 10 separate samples involving a total of 1152 participants mitigates this concern. Third, the confidence that can be placed in the findings would have been increased if we had used multiple reporter data. It is possible that depression-induced reporting biases contributed to the observed differences between subtypes reported in this study. Future studies should collect objective data (electrolyte markers) or confederate reports to remedy this limitation.

Given the accumulating evidence-base for the reliability and validity of the dieting-negative affect distinction, we think it would be worthwhile for additional research groups to investigate this subtyping scheme. It will be important to examine the predictive validity of this distinction, as it appears to have prognostic significance. As noted, it would also be useful for future studies to examine objective biological markers that may differentiate between these subtypes. Finally, it may be advantageous to test whether treatments that are tailored to the different clinical presentation of these two subtypes are more effective than a single treatment that does not take this difference into account. It is possible that this will improve the efficacy of treatments for this pernicious eating disorder.

Acknowledgments

This research was supported by grants MH01708 and MH/DK61957 from the National Institute of Health.

References

- Bathalon G, Tucker K, Hays N, Vinken A, Greenberg A, McCrory M, et al. Psychological measures of eating behavior and the accuracy of 3 common dietary assessment methods in healthy postmenopausal women. *American Journal of Clinical Nutrition* 2000;71:739–745. [PubMed: 10702167]
- Bowirrat A, Oscar-Berman M. Relationship between dopaminergic neurotransmission, alcoholism, and reward deficiency syndrome. *American Journal of Medical Genetics (Neuropsychiatric)* 2005;132B: 29–37.
- Beck AT, Steer RM, Garbin M. Psychometric properties of the Beck Depression Inventory: 25 years of evaluation. *Clinical Psychology Review* 1988;8:77–100.
- Bond MJ, McDowell AJ, Wilkinson JY. The measurement of dietary restraint, disinhibition and hunger: An examination of the factor structure of the Three Factor Eating Questionnaire (TFEQ). *International Journal of Obesity* 2001;25:900–906. [PubMed: 11439306]
- Burton E, Stice E. Evaluation of a healthy-weight treatment program for bulimia nervosa: A preliminary randomized trial. *Behaviour Research & Therapy* 2006;44:1727–1738. [PubMed: 16458252]
- Chen EY, Le Grange D. Subtyping adolescents with bulimia nervosa. *Behaviour Research and Therapy* 2007;45:2813–2820. [PubMed: 17949682]
- Fairburn, CG.; Cooper, Z. The eating disorder examination. In: Fairburn, C.; Wilson, G., editors. *Binge eating: Nature, assessment, and treatment*. NY: Guilford; 1993. p. 317–360.
- Fairburn CG, Harrison PJ. Eating disorders. *Lancet* 2003;361:407–416. [PubMed: 12573387]
- Fairburn CF, Stice E, Cooper Z, Doll HA, Norman PA, O'Connor ME. Understanding persistence of bulimia nervosa: A five-year naturalistic study. *Journal of Consulting and Clinical Psychology* 2003;71:103–109. [PubMed: 12602430]
- Fairburn CF, Wilson GT, Agras WS, Welch T, Stice E. Early change in treatment predicts outcome in bulimia nervosa. *American Journal of Psychiatry* 2004;161:2322–2324. [PubMed: 15569910]
- Fleiss, JL. *Statistical Methods for Rates and Proportions*. 2. New York: Wiley & Sons; 1981.
- French SA, Jeffery RW, Wing RR. Food intake and physical activity: A comparison of three measures of dieting. *Addictive Behaviors* 1994;19:401–409. [PubMed: 7992675]
- Garfinkel PE, Lin B, Goering P, Spegg C, Goldbloom D, Kennedy S, et al. Purging and non-purging forms of bulimia nervosa in a community sample. *International Journal of Eating Disorders* 1996;20:231–238. [PubMed: 8912035]
- Gleaves DH, Lowe MR, Green BA, Cororve MB, Williams TL. Do anorexia and bulimia nervosa occur on a continuum? A taxometric analysis. *Behavior Therapy* 2000;31:195–219.
- Glynn SM, Ruderman AJ. The development and validation of an eating self-efficacy scale. *Cognitive Therapy and Research* 1986;10:403–420.
- Grilo CM. Subtyping female adolescent psychiatric inpatients with features of eating disorders along dietary restraint and negative affect dimensions. *Behaviour Research and Therapy* 2004;42:67–78. [PubMed: 14744524]

- Grilo CM, Masheb RM, Berman RM. Subtyping women with bulimia nervosa along dietary and negative affect dimensions: A replication in a treatment-seeking sample. *Eating and Weight Disorders* 2001;6:53–58. [PubMed: 11300547]
- Grilo CM, Masheb R, Wilson GT. Subtyping binge eating disorder. *Journal of Consulting and Clinical Psychology* 2001;69:1066–1072. [PubMed: 11777111]
- Herman CP, Mack D. Restrained and unrestrained eating. *Journal of Personality* 1975;43:647–660. [PubMed: 1206453]
- Hetherington MM, Bell A, Rolls BJ. Pleasure and monotony: Effects of repeat exposure on pleasantness, preference and intake. *British Food Journal* 2000;102:507–521.
- Hetherington MM, Bell A, Rolls BJ. Pleasure and monotony: Effects of repeat exposure on pleasantness, preference and intake. *British Food Journal* 2000;102:507–521.
- Jansen A. How restrained eaters perceive the amount they eat. *British Journal of Clinical Psychology* 1996;35:381–392. [PubMed: 8889079]
- Killen JD, Taylor CB, Hayward C, Haydel KF, Wilson DM, Hammer L, et al. Weight concerns influence the development of eating disorders: A 4-year prospective study. *Journal of Consulting and Clinical Psychology* 1996;64:936–940. [PubMed: 8916622]
- Kreft, I.; De Leeuw, J. *Introducing Multilevel Modeling*. London: Sage Publications; 1998.
- Lowe MR, Annunziato RA, Markowitz JT, Didie E, Bellace DL, Riddell DL, et al. Multiple types of dieting prospectively predict weight gain during the freshman year of college. *Appetite* 2006;47:83–90. [PubMed: 16650913]
- Martin C, Williamson D, Geiselman P, Walden H, Smeets M, Morales S, et al. Consistency of food intake over 4 eating sessions in the laboratory. *Eating Behaviors* 2005;6:365–72. [PubMed: 16257810]
- McCann JD, Rossiter EM, King RJ, Agras WS. Nonpurging bulimia: A distinct subtype of bulimia nervosa. *International Journal of Eating Disorders* 1991;10:679–687.
- McLachlan, GJ.; Peel, D. *Finite Mixture Models*. New York: Wiley & Sons; 2000.
- McCarthy M. The thin ideal, depression, and eating disorders in women. *Behavioral Research and Therapy* 1990;28:205–218.
- Mendelson, B.; White, D.; Mendelson, M. *Manual for the Body-Esteem Scale for Adolescents and Adults*. 1996. Unpublished manual
- Muthén, LK.; Muthén, BO. *Mplus User's Guide*. 4.1. Los Angeles: Muthén & Muthén; 2006.
- Patton JH, Stanford MS, Barratt ES. Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology* 1995;51:768–774. [PubMed: 8778124]
- Penick E, Nickel E, Powell B, Liskow B, Campell J, Dale T, et al. The comparative validity of eleven alcoholism typologies. *Journal of Studies on Alcohol* 1999;60:188–202. [PubMed: 10091957]
- Pettinati HM. The use of selective serotonin reuptake inhibitors in treating alcoholic subtypes. *Journal of Clinical Psychiatry* 2001;62:26–31. [PubMed: 11584872]
- Polivy J, Herman CP. Dieting and binge eating: A causal analysis. *American Psychologist* 1985;40:193–204. [PubMed: 3857016]
- Raftery AE. Bayesian model selection in social research. *Sociological Methodology* 1995;25:111–164.
- Rolls BJ, Castellanos VH, Shide DJ, Miller DL, Pelkman CL, Thorwart ML, Peters JC. Sensory properties of a nonabsorbable fat substitute did not affect regulation of energy intake. *American Journal of Clinical Nutrition* 1997;65:1375–1383. [PubMed: 9129465]
- Singer, JD.; Willet, JB. *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford: Oxford University Press; 2003.
- Stice E. A prospective test of the dual pathway model of bulimic pathology: Mediating effects of dieting and negative affect. *Journal of Abnormal Psychology* 2001;110:124–135. [PubMed: 11261386]
- Stice E, Agras WS. Predicting onset and cessation of bulimic behaviors during adolescence: A longitudinal grouping analyses. *Behavior Therapy* 1998;29:257–276.
- Stice E, Agras WS. Subtyping bulimics along dietary restraint and negative affect dimensions. *Journal of Consulting and Clinical Psychology* 1999;67:460–469. [PubMed: 10450616]
- Stice E, Agras WS, Telch CF, Halmi C, Mitchel J, Wilson T. Subtyping binge eating disordered women along dietary restraint and negative affect dimensions. *International Journal of Eating Disorders* 2001;30:11–27. [PubMed: 11439405]

- Stice E, Burton EM, Shaw H. Prospective relations between bulimic pathology, depression, and substance abuse: Unpacking comorbidity in adolescent girls. *Journal of Consulting and Clinical Psychology* 2004;72:62–71. [PubMed: 14756615]
- Stice E, Cooper JA, Schoeller DA, Tappe K, Lowe MR. Are dietary restraint scales valid measures of moderate-to long-term dietary restriction? Objective biological and behavioral data suggest not. *Psychological Assessment* 2007;19:449–458. [PubMed: 18085937]
- Stice E, Fairburn CG. Dietary and dietary-depressive subtypes of bulimia nervosa show differential symptom presentation, social impairment, comorbidity, and course of illness. *Journal of Consulting and Clinical Psychology* 2003;71:1090–1094. [PubMed: 14622085]
- Stice E, Fisher M, Lowe MR. Are dietary restraint scales valid measures of acute dietary restriction? Unobtrusive observational data suggest not. *Psychological Assessment* 2004;16:51–59. [PubMed: 15023092]
- Stice E, Killen JD, Hayward C, Taylor CB. Age of onset for binge eating and purging during adolescence: A 4-year survival analysis. *Journal of Abnormal Psychology* 1998;107:671–675. [PubMed: 9830254]
- Stice E, Marti N, Spoor S, Presnell K, Shaw H. Dissonance and healthy weight eating disorder prevention programs: Long-term effects from a randomized efficacy trial. *Journal of Consulting and Clinical Psychology* 2008;76
- Stice E, Presnell K, Spangler D. Risk factors for binge eating onset: A prospective investigation. *Health Psychology* 2002;21:131–138. [PubMed: 11950103]
- Stunkard AJ, Messick S. The Three Factor Eating Questionnaire to measure dietary restraint, disinhibition, and hunger. *Journal of Psychosomatic Research* 1985;29:71–83. [PubMed: 3981480]
- Suris A, Borman PD, Lind L, Kashner TM. Aggression, impulsivity, and health functioning in a veteran population: Equivalency and test-retest reliability of computerized and paper-and-pencil administrations. *Computers in Human Behavior* 2004;23:97–110.
- Sysko R, Walsh TB, Schebendach J, Wilson GT. Eating behaviors among women with anorexia nervosa. *American Journal of Clinical Nutrition* 2005;82:296–301. [PubMed: 16087971]
- Tobin DL, Griffing A, Griffing S. An examination of subtype criteria for bulimia nervosa. *International Journal of Eating Disorders* 1997;22:179–186. [PubMed: 9261657]
- Tuschl RJ, Laessle RG, Platte P, Pirke KM. Differences in food-choice frequencies between restrained and unrestrained eaters. *Appetite* 1990;14:9–13. [PubMed: 2310178]
- van Strien T, Frijters JER, Bergers GPA, Defares PB. The Dutch eating behavior questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. *International Journal of Eating Disorders* 1986;5:295–315.
- Waller, G. Psychology of binge eating. In: Fairburn, CG.; Brownell, KD., editors. *Eating disorders and obesity: A comprehensive handbook*. London: Guilford; 2002. p. 98-102.
- Walsh BT. DSM-V from the perspective of the DSM-IV experience. *International Journal of Eating Disorders* 2007;40:S3–S7. [PubMed: 17573695]
- Walters EE, Neale MC, Eaves LJ, Heath AC, Kessler RC, Kendler KS. Bulimia nervosa: A population-based study of purgers versus nonpurgers. *International Journal of Eating Disorders* 1993;13:265–272. [PubMed: 8477298]
- Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Archives of General Psychiatry* 1976;33:1111–1115. [PubMed: 962494]
- Williamson DA, Martin CK, York-Crowe E, Anton SD, Redman LM, Han H, Ravussin E. Measurement of dietary restraint: Validity tests of four questionnaires. *Appetite* 2007;48:183–192. [PubMed: 17101191]
- Wonderlich SA, Crosby RD, Mitchell JE, Engel SG. Testing the validity of eating disorder diagnoses. *International Journal of Eating Disorders* 2007;40:S40–S45. [PubMed: 17683095]

Differences Between the Dietary and Dietary-Negative Affect Subtypes and Between the Nonpurging and Purging Subtypes at Baseline on Clustering Variables, Concurrent Validation Variables, and Predictive Validation Variables in Study 1

Clustering variables		Dietary subtype			Dietary-negative affect subtype			Nonpurging subtype			Purging subtype					
		M	(SD)		M	(SD)		M	(SD)		M	(SD)		M	(SD)	
				Percent variance explained												Percent variance explained
TFEQ cognitive restrained scale		11.4	(3.4)	9.4	(3.7)	7.6**		10.4	(3.6)	10.5	(3.7)	0.0				0.0
Beck Depression Inventory		13.5	(5.8)	25.5	(9.2)	40.5***		17.5	(9.4)	20.3	(9.7)	2.1				2.1
PANAS sadness scale		2.2	(0.6)	3.5	(0.8)	47.0***		2.7	(0.9)	2.9	(1.0)	0.6				0.6
PANAS anxiety scale		1.9	(0.6)	3.0	(0.8)	43.3***		2.2	(0.8)	2.6	(1.0)	4.8*				4.8*
PANAS guilt scale		2.6	(0.8)	3.9	(0.7)	44.4***		3.0	(0.9)	3.4	(1.0)	3.8*				3.8*
<u>Concurrent validation variables</u>																
EDDI binge eating frequency (past month)		7.0	(4.8)	15.4	(13.1)	16.5***		7.8	(6.0)	13.6	(12.5)	8.1**				8.1**
EDDI compensatory frequency (past month)		18.0	(19.3)	29.2	(23.6)	6.6**		17.4	(20.9)	28.3	(21.5)	6.2*				6.2*
EDDI weight and shape overvaluation		4.6	(1.2)	5.0	(1.2)	3.7*		4.7	(1.2)	4.9	(1.2)	0.4				0.4
TFEQ disinhibition		9.0	(2.4)	10.4	(2.3)	8.6**		9.1	(2.6)	10.1	(2.2)	4.8*				4.8*
TFEQ hunger		8.6	(3.6)	10.4	(3.0)	7.1**		9.3	(3.6)	9.6	(3.3)	0.3				0.3
ESES emotional eating		2.9	(1.1)	3.9	(1.1)	14.8***		3.3	(1.3)	3.4	(1.4)	0.1				0.1
Barratt Impulsivity Scale		67.9	(11.9)	71.9	(13.3)	1.4		69.4	(12.8)	68.9	(12.3)	0.0				0.0
SAS functional impairment		2.1	(0.3)	2.6	(0.6)	24.9***		2.2	(0.4)	2.4	(0.6)	1.6				1.6
		<u>Percentage</u>		<u>Percentage</u>		<u>Odds ratio</u>		<u>Percentage</u>		<u>Percentage</u>		<u>Odds ratio</u>				
Past year mental health treatment		17%		34%		2.7*		17%		33%		2.4*				
<u>Predictive validation variables</u>																
Recovery from bulimic pathology		<u>Percentage</u>		<u>Percentage</u>		<u>Odds ratio</u>		<u>Percentage</u>		<u>Percentage</u>		<u>Odds ratio</u>				
		46%		29%		2.3*		43%		30%		2.4*				
		<u>Change</u>		<u>Change</u>		<u>Pseudo R²</u>		<u>Change</u>		<u>Change</u>		<u>Pseudo R²</u>				
Change in BMI 6-month follow-up		-0.1 (1.5)		0.0 (0.8)		.00		-0.0 (1.4)		0.1 (0.9)		.00				

* $p < .05$,

 $p < .01$,

 $p < .001$.

TFEQ = Three Factor Eating Questionnaire. PANAS = Positive and Negative Affect Scale. EDDI = Eating Disorder Diagnostic Interview. ESES = Eating Self-Efficacy Scale. SAS = Social Adjustment Scale.

Table 2

Differences Between the Dietary and Dietary-Negative Affect Subtypes and Between the Nonpurging and Purging Subtypes at Baseline on Clustering Variables, Concurrent Validation Variables, and Predictive Validation Variables in Study 2

Clustering variables	Dietary subtype			Dietary-negative affect subtype			Nonpurging subtype			Purging subtype		
	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)
DEBQ restrained eating scale	3.4	(0.6)	3.8	(0.5)	12.5**		3.4	(0.6)	3.8	(0.4)	10.8*	
PANAS sadness scale	2.3	(0.6)	4.5	(0.5)	76.7***		2.7	(1.1)	3.3	(1.2)	6.4*	
PANAS anxiety scale	2.0	(0.7)	3.4	(0.8)	41.6***		2.3	(0.8)	2.8	(1.2)	6.5*	
PANAS guilt scale	2.9	(0.9)	4.2	(0.6)	35.9***		3.0	(0.9)	3.7	(1.1)	8.4*	
<u>Concurrent validation variables</u>												
EDDI binge eating frequency (past month)	5.6	(4.5)	9.3	(6.0)	10.2*		5.0	(3.9)	10.8	(5.7)	26.4***	
EDDI compensatory frequency (past month)	8.7	(10.5)	13.8	(8.3)	5.3		9.2	(10.5)	12.5	(8.9)	2.2	
EDDI weight and shape overevaluation	4.6	(0.8)	5.1	(0.8)	13.1**		4.5	(0.7)	4.9	(1.1)	6.2*	
Body dissatisfaction	3.9	(0.6)	4.3	(0.6)	6.0*		3.9	(0.6)	4.2	(0.6)	6.0*	
Distress about body image	3.7	(0.6)	4.2	(0.5)	14.4**		3.8	(0.6)	4.2	(0.4)	9.3*	
DEBQ emotional eating	3.0	(1.1)	3.9	(1.1)	10.6*		3.1	(1.3)	3.8	(1.1)	6.1*	
Mental health treatment	2.2	(3.9)	4.2	(4.9)	4.5*		2.8	(4.3)	2.9	(4.4)	0.0	
SAS functional impairment	2.7	(0.4)	3.3	(0.5)	24.2***		2.8	(0.4)	3.0	(0.6)	3.0	
<u>Predictive validation variables</u>												
Recovery from bulimic pathology	Percentage		Percentage		Percentage		Percentage		Percentage		Odds ratio	
	67%		31%		61%		52%		1.3			
	Change		Change		Change		Change		Pseudo R ²			
	1.0 (2.4)		0.3 (1.9)		-0.0 (1.4)		0.1 (0.9)					
Change in BMI over 3-year follow-up												

Note:

* p < .05,

** p < .01,

*** p < .001.

DEBQ = Dutch Eating Behavior Questionnaire. PANAS = Positive and Negative Affect Scale. EDDI = Eating Disorder Diagnostic Interview. SAS = Social Adjustment Scale.