Examining the Stability of *DSM–IV* and Empirically Derived Eating Disorder Classification: Implications for *DSM–5*

Carol B. Peterson,
Department of Psychiatry, University of Minnesota Medical School

Scott J. Crow,
Department of Psychiatry, University of Minnesota Medical School

Ross D. Crosby,
Department of Clinical Research, Neuropsychiatric Research Institute and Department of Clinical Neuroscience, University of North Dakota School of Medicine and Health Sciences

Stephen A. Wonderlich,
Department of Clinical Research, Neuropsychiatric Research Institute and Department of Clinical Neuroscience, University of North Dakota School of Medicine and Health Sciences

James E. Mitchell,
Department of Clinical Research, Neuropsychiatric Research Institute and Department of Clinical Neuroscience, University of North Dakota School of Medicine and Health Sciences

Sonja A. Swanson,
Department of Epidemiology, Harvard School of Public Health

W. Stewart Agras, and
Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine

Katherine A. Halmi
Department of Psychiatry, Weill Medical College, Cornell University

Abstract

Objective—The purpose of this investigation was to derive an empirical classification of eating disorder symptoms in a heterogeneous eating disorder sample using latent class analysis (LCA) and to examine the longitudinal stability of these latent classes (LCs) and the stability of *DSM–IV* eating disorder (ED) diagnoses.

Method—A total of 429 females with ED symptoms were assessed using the Eating Disorder Examination every 6 months for 2 years. LCA was used to derive empirical classification at baseline. Latent transition analysis (LTA) was used to examine the longitudinal stability of LCs, and Markov modeling procedures were used to examine *DSM–IV* ED diagnoses over all the time points.

Correspondence concerning this article should be addressed to Carol B. Peterson, Department of Psychiatry, University of Minnesota Medical School, F282/2A West, 2450 Riverside Avenue, Minneapolis, MN 55454-1495. peter161@umn.edu.

Results—LCA yielded a 3-class solution: binge eating and purging, binge eating only, and low body mass index. LTA indicated that these LCs showed greater stability over 2 years than *DSM–IV* diagnoses with the probability of remaining in the same class ranging from 0.69 to 0.91 for LCs and from 0.40 to 0.75 for *DSM–IV* diagnoses. Transition patterns also revealed more stability for LCs with only 21% changing classes compared with 63% of the *DSM–IV* diagnostic categories.

Conclusion—Empirically derived classes of ED symptoms showed greater longitudinal stability than *DSM–IV* diagnoses over a 2-year time period, suggesting that modifying the criteria to be consistent with empirically based classification (e.g., reducing frequency requirements of binge eating and purging) may reduce ED diagnostic crossover in *DSM–5*.

A valid eating disorder (ED) classification system is essential for developing accurate etiological models, effective treatment, and successful prevention programs (Keel et al., 2004; Wonderlich, Joiner, Keel, Williamson, & Crosby, 2007). Currently under revision, the diagnostic criteria for eating disorders as defined by the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV*; American Psychiatric Association, 1994) have inconsistent empirical support for the validity of the currently defined constructs (Wonderlich, Joiner, Keel, Williamson, & Crosby, 2007). In addition, the majority of patients with ED meet criteria for the residual category of not otherwise specified (ED NOS) rather than anorexia nervosa (AN) or bulimia nervosa (BN), and many do not remain in the same ED category or subcategory longitudinally (Eddy et al., 2008; Fairburn & Bohn, 2005; Wade, Crosby, & Martin, 2006; Wilfley, Bishop, Wilson, & Agras, 2007). Empirically derived classification procedures in which psychopathology subgroups are determined statistically may improve the validity of ED diagnoses and, potentially, their longitudinal stability.

Empirically derived classification, including latent class analysis (LCA) and latent profile analysis (LPA), have yielded inconsistent support of the *DSM–IV* ED criteria. As highlighted in a recent review (Crow et al., 2011), close to 20 studies have been conducted using LCA/LPA with mixed findings the likely result of inconsistent samples (e.g., epidemiological vs. clinical), a broad range of indicators and validators (e.g., ED symptoms, co-occurring psychopathology, demographics, and personality), measurement variability, and methodological problems (e.g., local independence violations, sparseness). These studies have yielded as many as six and as few as two profiles/classes. Although several LCA/LPA studies have found latent classes (LCs) resembling *DSM–IV* BN (Bulik, Sullivan, & Kendler, 2000; Keel et al., 2004; Striegel-Moore et al., 2005), AN (Bulik et al., 2000; Keel et al., 2004; Wade et al., 2006), and binge eating disorder (BED; Bulik et al., 2000; Mitchell et al., 2007; Striegel-Moore et al., 2005), most studies have not found AN, BN, and BED categories simultaneously (Crow et al., 2011). In addition, and, in many cases, depending on selected indicators, categories often emerge that are not included in the *DSM–IV* (e.g., perfectionism, impulsivity) as well as asymptomatic classes in epidemiological samples (Crow et al., 2011; Wonderlich et al., 2007).

Diagnostic crossover among *DSM–IV* ED diagnoses has been observed in a number of studies, suggesting that individuals do not typically remain within one ED diagnostic category longitudinally (Keel et al., 2004; Milos, Spindler, Schnyder, & Fairburn, 2005; Tozzi et al., 2005). Particularly common patterns include movement from AN to BN (Eddy et al., 2008) as well as from AN and BN to ED NOS (Agras, Crow, Mitchell, Halmi, & Bryson, 2009). Although the fluctuating course of ED symptoms may be characteristic of these conditions, an alternative classification system might improve the long-term stability of diagnostic categories. Notably, although empirical classification of ED symptoms has been conducted repeatedly, the longitudinal stability of these empirically derived classes has not been investigated. An empirically derived classification system with longitudinal stability could potentially inform the revision of the *DSM* criteria, improving both its validity and clinical utility. The purpose of the present investigation was to determine the LC structure of a heterogeneous sample of individuals with ED symptoms and examine the longitudinal stability of these LCs in comparison to *DSM–IV* diagnoses using latent transition analysis (LTA; Collins et al., 1994).

Method

Participants

Participants were 429 females with heterogeneous ED symptoms who were enrolled in a multicenter longitudinal study in one of three sites in California, New York, or Minnesota (Agras et al., 2009; Crow, Agras, Halmi, Mitchell, & Kraemer, 2002). Participants, recruited from clinics, other ED research studies, and community advertisements, were required to be female, between the ages of 14 and 55, and meet current *DSM–IV* criteria as determined by the Eating Disorder Examination (EDE; Fairburn & Cooper, 1993) for AN ($n = 51$), BN ($n = 99$), BED ($n = 116$), or ED NOS ($n = 163$), which included Partial AN (P-AN; $n = 51$), Partial BN (P-BN; $n = 66$), and Partial BED (P-BED; $n = 46$) based on a priori definitions of these subthreshold conditions (Crow et al., 2002).[1] Of the 1,005 individuals screened by phone, 429 were enrolled in the study. As shown in Table 1, participants were generally well educated and predominantly Caucasian. Average ages ranged between 26.9 and 38.0 for subgroups and 32.2 overall.

This study was approved by the Institutional Review Board at each of the three sites. Participants were screened by phone prior to their baseline assessment at one of the three research centers, at which time written informed consent was obtained.

Measures

The EDE (Fairburn & Cooper, 1993) is an investigator-based interview that was used to determine ED diagnosis and associated symptoms. The EDE assesses psychological aspects

---

[1]Partial anorexia nervosa (AN) (a) meets all *DSM–IV* criteria for AN in the past 12 months but not in the past month or (b) meets all *DSM–IV* criteria for AN in the past 6 months except that body weight is 90% of ideal/expected body weight and is accompanied by amenorrhea or *DSM–IV* Criterion B (i.e., intense fear of weight gain) and Criterion C (i.e., body image disturbance, overvaluation of shape or weight, denial of seriousness of emaciation); partial bulimia nervosa (BN) (a) meets all *DSM–IV* criteria for BN in the past 3 months except for overvaluation of shape or weight or (b) meets all *DSM–IV* criteria for BN except that binge eating and compensatory behaviors occur less than twice per week but occur on average at least once per month for the past 6 months; partial binge eating disorder (BED) meets all *DSM–IV* criteria for BED except that binge eating occurs less than twice per week but occurs on average at least once per month for the past 6 months.

of ED symptoms (with items and subscales scored on a 0–6 scale, with 6 being most pathological) as well as the frequency of behavioral symptoms, including binge eating and purging. The EDE has extensive psychometric data supporting its use and has been found to be a useful measure of assessing change in ED symptoms (Fairburn, Cooper, & O'Connor, 2008). Prior to the commencement of data collection, master's- and doctoral-level assessors received instruction in the EDE using didactic and audiotape training. After a series of role plays and observations, EDE audiotapes were rated for quality by the senior assessor before interviewers were allowed to conduct EDEs for the study. Regular in-person meetings, teleconferences, and e-mail discussions were maintained to minimize site drift. Test-retest and interrater reliability procedures on a subset of interviews yielded coefficients greater than 0.7 on EDE subscales and measures of symptom frequencies (Agras et al., 2009; Rizvi, Peterson, Crow, & Agras, 2000), with the exception of test–retest reliabilities of subjective bulimic episodes, which were lower (0.33). Body mass index (BMI) was also measured at each assessment visit. After baseline assessment, participants were reassessed in person every 6 months for 2 years. In addition, they completed a brief phone interview (Agras et al., 2009) between visits that included an abridged version of the EDE providing data that could be used as needed to determine diagnosis at 6-month visits. Participants were paid $100 for baseline visit and $50 for each 6-month visit.

Diagnostic Categories and Indicator Variables

Indicator variables for LCA and LTA were selected a priori based on previous empirical classification studies (see Crow et al., 2011) and included the following: the presence or absence of objective binge eating (OBE), as defined by any objective bulimic episodes (Fairburn & Cooper, 1993) in which the amount eaten was objectively large, and the participant experienced a sense of loss of control in the past 28 days on the EDE; subjective binge eating (SBE), as defined by the presence or absence of any subjective bulimic episodes (Fairburn & Cooper, 1993) in which the amount eaten was not objectively large but the participant experienced a sense of loss of control in the past 28 days on the EDE; purging, as defined by the presence or absence of any self-induced vomiting or laxative misuse in the past 28 days on the EDE; fear of weight gain, as defined by a score of 4, 5, or 6 on the Fear of Weight Gain item on the EDE; and low BMI, as defined by a current BMI of less than 18.5. The number of indicator variables was limited to five to prevent an artificial inflation in the number of empirical classes (Swanson, Lindenberg, Bauer, & Crosby, in press).

Because of the longitudinal nature of these analyses, *remission status* was defined a priori in order to provide categories for participants who were no longer symptomatic. For the empirical classification, *remission* was defined as the absence of OBE, SBE, vomiting, and laxative use behaviors; the EDE importance of shape, importance of weight, and fear of weight gain scores all less than 4, and a BMI above 17.5 all in the past 28 days. *DSM–IV remission status* was defined as the absence of *DSM–TV* behavioral and cognitive ED symptoms as determined by the assessor at each visit based on the EDE and phone interview.

Statistical Method

For empirically derived diagnoses, exploratory LCA was conducted at each 6-month time point to ensure consistency in structure and content over time. Model fit and number of classes were then determined primarily using Akaike information criterion and sample size-adjusted Bayesian information criterion (Lin & Dayton, 1997; Yang, 2006). For all the 6-month time points beyond baseline, a fixed class was created using the definition of remission described above; otherwise, the model structure and class membership were derived empirically. Once it was established that the empirical class structure was similar across time points, a full latent transition model was examined. LTA models were examined with and without stationary probabilities (transition probabilities that do not change over time). Data supported the invariance model (i.e., stationary probability), assuming equivalence for time intervals. Transition patterns were examined to evaluate migration between empirical classes over time. *DSM–IV* diagnostic stability was examined using first-order Markov modeling.

Results

Of the 429 participants, 386 (90.0%) were assessed at Time 2 (6 months), 350 (81.6%) were assessed at Time 3 (12 months), 328 (76.5%) were assessed at Time 4 (18 months), and 313 (73.0%) were assessed at the final (24-month) time point. Attrition differed slightly by baseline diagnosis and LC, with 84.3%, 73.7%, 87.1%, and 81.6% of participants with AN, BN, BED, and ED NOS, respectively, having assessments during at least three time points; 74.2%, 85.8%, and 84.4% of participants in LC1, LC2, and LC3 at baseline were assessed during at least three time points.

LC Structure

Three LCs (and a remission class for nonbaseline time points) were found at each of the time points with consistency in shape and structure (see Figure 1). As shown in Figure 2, model fit was best for a three-class solution: LC1 ("binge eating and purging") was characterized by highest probability estimates for OBEs (0.75), SBEs (0.65), and vomiting/laxative abuse (0.90), with relatively lower estimates for fear of weight gain (0.49) and low BMI (0.49); LC2 ("binge eating") was characterized by highest probability estimates for OBEs (0.71), SBEs (0.54), and lower estimates for fear of weight gain (0.37), vomiting/laxative abuse (0.00), and low BMI (0.37); LC3 ("low BMI") had higher probability estimates for low BMI (0.56) and fear of weight gain (.0.56) than SBEs (0.25), purging (0.11), and OBEs (0.00).

Table 2 displays the overlap between the LCs and *DSM–IV* diagnoses, revealing that LC1 (binge eating and purging) overlapped with BN, P-AN, and P-BN; LC2 (binge eating) overlapped with BED and P-BED; and LC3 (low BMI) overlapped with AN and P-AN.

Longitudinal Stability

As shown in Table 3, LCs showed relatively high stability over the 2-year period based on LTA, with probability estimates of stability within class over time of 0.69 for LC1, 0.91 for LC2, and 0.83 for LC3. The majority of crossover that did occur took place between all of

the LCs and the remission category. Stability was lower for the *DSM–IV* diagnoses (see Table 4), with stability within diagnosis probability estimates of 0.53 for AN, 0.61 for BN, 0.40 for BED, and 0.75 for ED NOS. As noted previously, movement was particularly notable between the primary diagnoses (AN, BN, BED) and ED NOS (Agras et al., 2009).

Stability was also examined on the basis of the numerical frequency of transitions from one *DSM–IV* diagnosis or from LC to another (excluding movement in and out of the remission category). Over the course of 2 years in the empirical classification, 21% moved to a different LC (not including the asymptomatic class). The most common transition patterns involved individuals remaining in the same LC across all assessments (LC1: $n = 67$; LC2: $n = 124$; LC3: $n = 32$) or migrating from LC1 to LC2 ($n = 34$). For *DSM–IV* diagnoses, 63% moved categories over 2 years. The most common transition patterns for *DSM–IV* involved individuals receiving the same diagnosis across all assessments (ED NOS: $n = 47$; BN: $n = 13$; AN: $n = 6$) or migrating from BED to ED NOS ($n = 33$), from BN to ED NOS ($n = 17$), or from ED NOS to asymptomatic ($n = 6$).

Discussion

We observed in this investigation greater longitudinal stability in empirical classification of ED symptoms over a 2-year period in contrast to *DSM–IV ED* diagnoses. Specifically, three LCs characterized as binge eating and purging, binge eating only, and low BMI showed greater longitudinal consistency and less crossover at five time points than the *DSM–IV* ED diagnostic categories. Consistent with some previous findings (e.g., Bulik et al., 2000), this investigation observed three LCs that roughly resembled the *DSM–IV* categories of AN, BN, and BED. Although the empirically derived classes showed greater stability than the *DSM–IV* groups, the fact that the classes generally map onto *DSM–IV* categories is noteworthy. However, several factors may have improved the stability of the LCs compared with the *DSM–IV* diagnostic categories. First, LC1 included individuals with binge eating and purging from all BMI levels. This finding is consistent with Williamson and colleagues' three-dimensional model of EDs based on taxometric analyses (Williamson, Gleaves, & Stewart, 2005) and suggests that grouping individuals who binge eat and purge regardless of body weight may reduce the common diagnostic migration between AN and BN (Crow et al., 2011; Eddy et al., 2008). Second, the indicator variables for binge eating and purging in this study were defined as the occurrence of these symptoms at least monthly, a frequency threshold that is significantly lower than the requirement of twice per week for the *DSM–IV* criteria. Finally, different "types" of binge eating (OBE/SBE) were included in the model as distinct entities. This distinction has been debated, and the validity of differentiating binge eating episodes based on size remains unclear (Wilfley et al., 2007; Wolfe, Wood Baker, Smith, & Kelly-Weeder, 2009). This investigation did not find that binge eating episode size was a distinctly defining feature of the empirically derived classes, as both OBEs and SBEs were observed in LC1 and LC2.

Several implications for the revision of the *DSM–IV* are supported by the findings of this investigation. First, lowering the frequency requirements from twice per week for binge eating and compensatory behaviors for BN criteria and binge eating for BED criteria may improve longitudinal consistency and reduce crossover between these diagnoses and ED

NOS. Second, the results of this study suggest that the distinction between binge eating episodes that are considered objectively large by clinical raters and those that are regarded as binge eating episodes with the subjective experience of loss of control by patients but are not characterized by the consumption of a large amount of food may not be critical in differentiating diagnostic categories that are longitudinally stable. Finally, this empirically derived classification indicated that grouping individuals who binge eat and purge into the same class, regardless of weight status, improved stability. Although combining individuals who binge eat and purge who are underweight with those who are normal weight or even overweight into the same *DSM* category may reduce longitudinal crossover between AN and BN, the clinical utility of placing underweight individuals into their own category may be more important than diagnostic longitudinal stability (Eddy et al., 2008; Thomas et al., 2010). The fact that the findings of the present study support the conceptualization of at least some ED symptoms (e.g., frequency and size of binge eating episodes) dimensionally implies that whatever cutoff values are set in the diagnostic criteria (e.g., bulimic behaviors occurring once or twice a week) should be determined on the basis of priorities of clinical utility (e.g., diagnostic stability, identifying mortality risk, etc.).

This study is unique in its investigation of the temporal consistency of empirically defined classes of EDs using LTA, as well as the use of a remission class in this context. Strengths of this study include the heterogeneity of ED diagnoses within the sample, the use of the remission class in examining longitudinal stability, the fact that the study was conducted at three different sites, which potentially increases its generalizability, and the use of the EDE as the main assessment instrument. However, several limitations are notable, including longitudinal attrition and the specific designations of ED NOS that may limit the generalizability of these findings to broader ED NOS populations (i.e., absence of "purging" disorder). In addition, the 2-year time frame may have been too short to detect more enduring transition patterns. Findings of this study are contingent on the selected indicators, which, in this case were limited to ED symptoms. Future research is needed to identify the longitudinal stability of alternative indicators (e.g., personality dimensions, biomedical markers, co-occurring psychopathology) in the empirical classification of EDs. Some participants in this naturalistic sample were receiving treatment in the community that may have impacted their diagnostic status longitudinally. Although the examination of the mechanisms of diagnostic migration is beyond the scope of the present study, future research is also indicated to identify the role of specific psychotherapeutic and psychopharmacological interventions in facilitating change in symptom status and remission. An additional limitation of this study is due to the fact that direct statistical comparisons are not feasible using LTA, and data comparisons in this study are descriptive and not inferential. Finally, because the sample was limited to adult females who were primarily Caucasian, these findings may not generalize to broader, ethnically diverse populations as well as to males and adolescents with ED symptoms.

In summary, empirical classification of ED symptoms yielded three classes: binge eating only, binge eating and purging, and low BMI. These classes showed greater stability over 2 years than *DSM–IV* diagnoses, which were characterized by more diagnostic migration. These findings suggest that empirical classification is associated with greater longitudinal stability than the current diagnostic system. In addition, these results suggest that modifying

the diagnostic criteria with lower thresholds of symptom frequency (e.g., binge eating and purging monthly instead of twice per week) and the inclusion of low-BMI individuals in the binge eating and purging category may reduce diagnostic crossover. Future investigations are necessary to replicate these findings in broader and more diverse samples over longer time durations as well as to demonstrate their clinical utility with these as well as with additional indicators including mortality, co-occurring psychiatric symptoms, and medical severity.

Acknowledgments

References

Agras WS, Crow S, Mitchell JE, Halmi KA, Bryson S. A 4-year prospective study of eating disorder NOS compared with full eating disorder syndromes. International Journal of Eating Disorders. 2009; 42:565–570. DOI: 10.1002/eat.20708 [PubMed: 19544557]

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th. Washington, DC: Author; 1994.

Bulik CM, Sullivan PF, Kendler KS. An empirical study of the classification of eating disorders. American Journal of Psychiatry. 2000; 157:886–895. DOI: 10.1176/appi.ajp.157.6.886 [PubMed: 10831467]

Collins LM, Graham JW, Rousculp SS, Fidler PL, Pan J, Hansen WB. Latent transition analysis and how it can address prevention research questions. National Institute of Drug Abuse Research Monograph. 1994; 142:81–111.

Crow, S., Swanson, SA., Peterson, CB., Crosby, RD., Wonderlich, SA., Mitchell, JE. Latent structure analyses of eating disorder diagnoses: Critical review of results and methodological issues. In: Striegel-Moore, R.Wonderlich, SA.Walsh, BT., Mitchell, JE., editors. Developing an evidence-based classification of eating disorders: Scientific findings for DSM-5. Arlington, VA: American Psychiatric Association; 2011. p. 103-120.

Crow SJ, Agras WS, Halmi K, Mitchell JE, Kraemer HC. Full syndromal versus subthreshold anorexia nervosa, bulimia nervosa, and binge eating disorder: A multicenter study. International Journal of Eating Disorders. 2002; 32:309–318. DOI: 10.1002/eat.10088 [PubMed: 12210645]

Eddy KT, Dorer DJ, Franko DL, Tahilani K, Thompson-Brenner H, Herzog DB. Diagnostic crossover in anorexia nervosa and bulimia nervosa: Implications for DSM-V. American Journal of Psychiatry. 2008; 165:245–250. DOI: 10.1176/appi.ajp.2007.07060951 [PubMed: 18198267]

Fairburn CG, Bohn K. Eating disorders NOS (ED NOS): an example of the troublesome "not otherwise specified" (NOS) category in DSM-IV. Behaviour Research and Therapy. 2005; 43:691–701. DOI: 10.1016/j.brat.2004.06.011 [PubMed: 15890163]

Fairburn, CG., Cooper, Z. The Eating Disorder Examination. In: Fairburn, CG., Wilson, GT., editors. Binge eating: Nature, assessment and treatment. 12th. New York: NY: Guilford Press; 1993. p. 317-360.

Fairburn, CG., Cooper, Z., O'Connor, M. The Eating Disorder Examination. In: Fairburn, CG., editor. Cognitive behavior therapy and eating disorders. New York, NY: Guilford Press; 2008. p. 265-308.

Keel PK, Fichter M, Quadflieg N, Bulik CM, Baxter MG, Thornton L, Kaye WH. Application of a latent class analysis to empirically define eating disorder phenotypes. Archives of General Psychiatry. 2004; 61:192–200. DOI: 10.1001/archpsyc.61.2.192 [PubMed: 14757596]

Lin T, Dayton C. Model selection information criteria for non-nested latent class models. Journal of Educational and Behavioral Statistics. 1997; 22:249–264.

Milos G, Spindler A, Schnyder Y, Fairburn CG. Instability of eating disorder diagnoses: Prospective study. British Journal of Psychiatry. 2005; 187:573–578. DOI: 10.1192/bjp.187.6.573 [PubMed: 16319411]

Mitchell JE, Crosby RD, Wonderlich SA, Hill L, le Grange D, Powers P, Eddy K. Latent profile analysis of a cohort of patients with eating disorders not otherwise specified. International Journal of Eating Disorders. 2007; 40(Suppl):S95–S98. DOI: 10.1002/eat.20459 [PubMed: 17886266]

Rizvi SL, Peterson CB, Crow SJ, Agras WS. Test-retest reliability of the Eating Disorder Examination. International Journal of Eating Disorders. 2000; 28:311–316. DOI: 10.1002/1098-108X(200011)28:3<311::AID-EAT8>3.0.CO;2-K [PubMed: 10942917]

Striegel-Moore RH, Franko DL, Thompson D, Barton B, Schreiber GB, Daniels SR. An empirical study of the typology of bulimia nervosa and its spectrum variants. Psychological Medicine. 2005; 35:1563–1572. DOI: 10.1017/S0033291705006057 [PubMed: 16219114]

Swanson SA, Lindenberg K, Brauer S, Crosby RD. A Monte Carlo investigation of factors influencing latent class analysis: An application to eating disorders research. Journal of Eating Disorders. (in press).

Thomas JJ, Delinsky SS, Germain SA St, Weigel TJ, Tangren CM, Levendusky PG, Becker AE. How do eating disorder specialist clinicians apply DSM-IV diagnostic criteria in routine clinical practice? Implications for enhancing clinical utility of DSM-5. Psychiatry Research. 2010; 178:511–517. DOI: 10.1016/j.psychres.2010.05.021 [PubMed: 20591498]

Tozzi F, Thornton LM, Klump KL, Fichter MM, Halmi KA, Kaplan AS, Kaye WH. Symptom fluctuation in eating disorders: Correlates of diagnostic crossover. American Journal of Psychiatry. 2005; 162:732–740. DOI: 10.1176/appi.ajp.162.4.732 [PubMed: 15800146]

Wade TD, Crosby RD, Martin NG. Use of latent profile analysis to identify eating disorder phenotypes in an adult Australian twin cohort. Archives of General Psychiatry. 2006; 63:1377–1384. DOI: 10.1001/archpsyc.63.12.1377 [PubMed: 17146012]

Wilfley DE, Bishop M, Wilson GT, Agras WS. Classification of eating disorders: Toward DSM-V. International Journal of Eating Disorders. 2007; 40:S123–S129. DOI: 10.1002/eat.20436 [PubMed: 17685383]

Williamson DA, Gleaves DH, Stewart TM. Categorical versus dimensional models of eating disorders: An examination of the evidence. International Journal of Eating Disorders. 2005; 37:1–10. DOI: 10.1002/eat.20074 [PubMed: 15690459]

Wolfe BE, Wood Baker C, Smith AT, Kelly-Weeder S. Validity and utility of the current definition of binge eating. International Journal of Eating Disorders. 2009; 42:674–686. DOI: 10.1002/eat. 20728 [PubMed: 19610126]

Wonderlich SA, Joiner TE, Keel PK, Williamson DA, Crosby RD. Eating disorder diagnoses: Empirical approaches to classification. American Psychologist. 2007; 62:167–180. DOI: 10.1037/0003-066X.62.3.167 [PubMed: 17469896]

Yang C. Evaluating latent class analysis models in qualitative phenotype identification. Computational Statistics and Data Analysis. 2006; 50:1090–1104. DOI: 10.1016/j.csda.2004.11.004
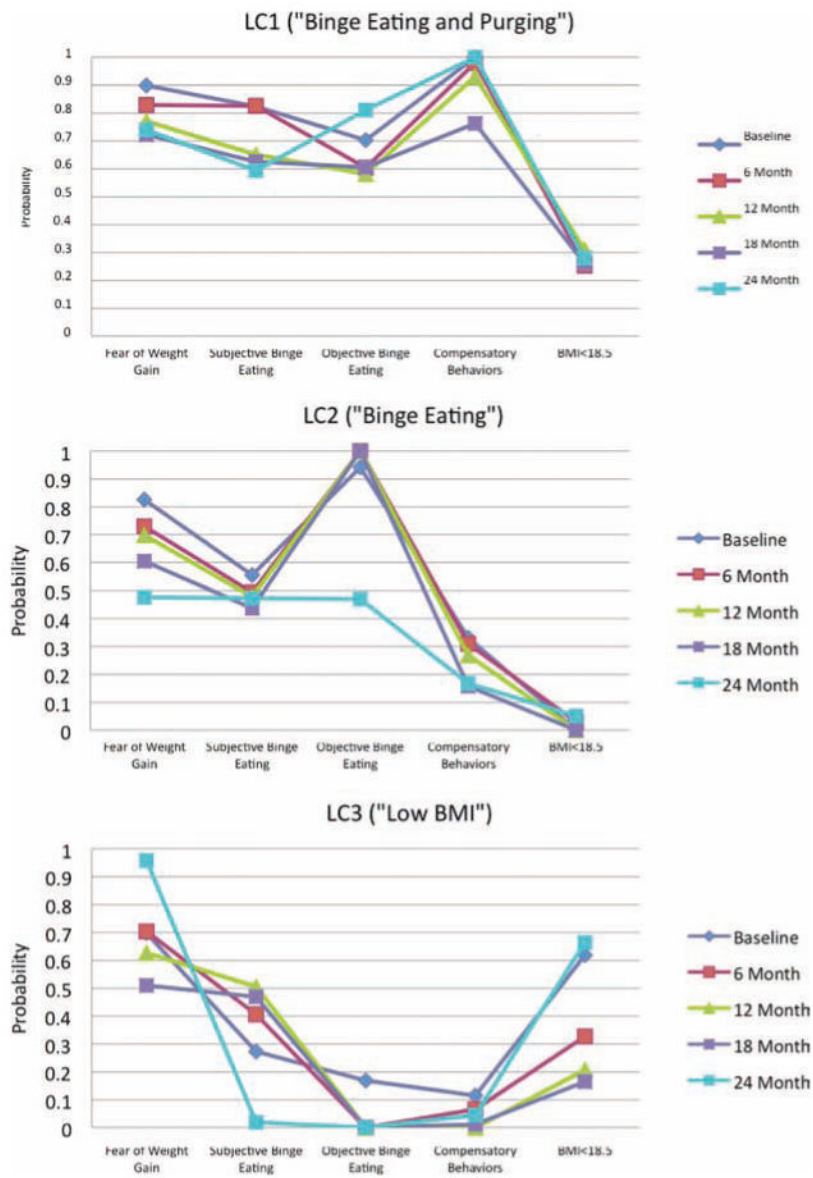
Figure 1.
Longitudinal latent class structures. Latent class (LC) structures (LC1, LC2, LC3) for each time assessed (baseline, 6 month, 12 month, 18 month, 24 month) with probability estimates for each indicator: Fear of Weight Gain, Subjective Binge Eating, Objective Binge Eating, Compensatory Behaviors, Body Mass Index (BMI) less than 18.5.
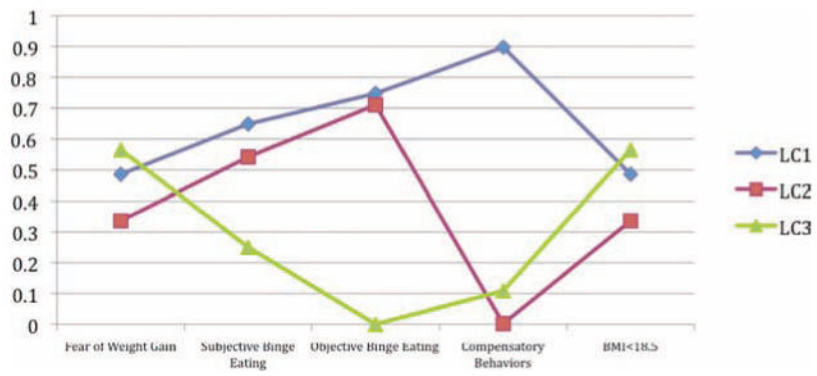
Figure 2.
Latent class structures. Probability estimates of indicators (Fear of Weight Gain, Subjective Binge Eating, Objective Binge Eating, Compensatory Behaviors, Body Mass Index [BMI] less than 18.5) for each latent class (LC): LC1 (Binge Eating/Purging), LC2 (Binge Eating), LC3 (Low BMI).

Table 1

Baseline Demographics (by Diagnosis, Latent Class, and Overall)

Variable	Anorexia nervosa *n* = 51	Bulimia nervosa *n* = 99	Binge eating disorder *n* = 116	Partial anorexia nervosa *n* = 51	Partial bulimia nervosa *n* = 66	Partial binge eating disorder *n* = 46	Latent Class 1 (binge/ purge) *n*= 147	Latent Class 2 (binge eating) *n* = 218	Latent Class 3 (low body mass index) *n* = 64	Overall *n* = 429
Mean age (*SD*)	26.9 (9.6)	29.9 (8.2)	38.0 (7.5)	27.8 (7.4)	30.3 (8.1)	36.6 (8.5)	28.9 (7.9)	35.7 (8.6)	28.3 (9.2)	32.2 (9.1)
Range	15–50	15–49	15–44	15–44	19–48	19–49	15–50	15–50	17–50	15–50
Mean body mass index (*SD*)	16.7 (2.1)	24.7 (8.9)	37.2 (8.6)	18.9 (2.1)	26.8 (7.4)	36.0 (10.5)	23.2 (6.9)	34.2 (10.6)	18.0 (3.3)	28.1 (10.8)
% Caucasian	92.2	92.9	88.8	88.2	87.9	95.6	91.8	89.9	90.6	90.7
% College graduates	41.2	44	52.2	41.2	40.9	56.5	41.0	53.0	51.6	48.5

Table 2

Overlap (N) Between Latent Classes and DSM–IV Diagnosis at Baseline

Variable	Anorexia nervosa $n = 51$	Bulimia nervosa $n = 99$	Binge eating disorder $n = 116$	Partial anorexia nervosa $n = 51$	Partial bulimia nervosa $n = 66$	Partial binge eating disorder $n = 46$
Latent Class 1 (binge/purge) $n = 147$	$n = 16$	$n = 61$	$n = 5$	$n = 21$	$n = 41$	$n = 3$
	Row: 10.9%	Row: 41.5%	Row: 3.4%	Row: 14.3%	Row: 27.9%	Row: 2.0%
	Col.: 31.4%	Col.: 61.6%	Col.: 4.3%	Col.: 41.2%	Col.: 62.1%	Col.: 6.5%
Latent Class 2 (binge eating) $n = 218$	$n = 2$	$n = 37$	$n = 111$	$n = 7$	$n = 19$	$n = 42$
	Row: 0.9%	Row: 17.0%	Row: 50.9%	Row: 3.2%	Row: 8.7%	Row: 19.3%
	Col.: 3.9%	Col.: 37.4%	Col.: 95.7%	Col.: 13.7%	Col.: 28.8%	Col.: 91.3%
Latent Class 3 (low body mass index) $n = 64$	$n = 33$	$n = 1$	$n = 0$	$n = 23$	$n = 6$	$n = 1$
	Row: 51.6%	Row: 1.6%	Row: 0.0%	Row: 35.9%	Row: 9.4%	Row: 1.6%
	Col.: 64.7%	Col.: 1.0%	Col.: 0.0%	Col.: 45.1%	Col.: 9.1%	Col.: 2.2%

Note. Row and Col (Columns) refer to percent within Row (latent class) and percent within Column (eating disorder diagnosis). *DSM–IV = Diagnostic and Statistical Manual of Mental Disorders*, 4th edition.

Table 3

Stability Estimates for Empirical Classification

Classification	Latent Class 1 (binge/purge)	Latent Class 2 (binge eating)	Latent Class 3 (low body mass index)	Latent Class 4 (asymptomatic)
Latent Class 1 (binge/purge)	.69	.19	.09	.03
Latent Class 2 (binge eating)	.00	.91	.02	.07
Latent Class 3 (low body mass index)	.06	.00	.83	.11
Latent Class 4 (asymptomatic)	.04	.19	.25	.53

Table 4

Stability Estimates for DSM–IV Diagnoses

Diagnosis	Anorexia nervosa	Bulimia nervosa	Binge eating disorder	Eating disorder not otherwise specified	Asymptomatic
Anorexia nervosa	.53	.04	.01	.39	.03
Bulimia nervosa	.01	.61	.01	.38	.01
Binge eating disorder	.00	.01	.40	.52	.07
Eating disorder not otherwise specified	.02	.04	.07	.75	.12
Asymptomatic	.01	.00	.02	.40	.57

Note. DSM–IV = Diagnostic and Statistical Manual of Mental Disorders, 4th edition.