

Are the Criterion B Binge-Eating Disorder Symptoms Interchangeable in Conveying the
Latent Binge-Eating Trait? An Item Response Theory Analysis

By
© 2020
Brianna Richson
University of Kansas
B.A., University of Iowa, 2015

Submitted to the graduate degree program in Psychology and the Graduate Faculty of the
University of Kansas in partial fulfillment of the requirements for the degree of Master of Arts.

Chair: Kelsie T. Forbush, Ph.D.

Tera Fazzino, Ph.D.

Rick Ingram, Ph.D

Date Defended: 24 February 2020

Date Accepted: 24 February 2020

The thesis committee for Brianne Richson certifies that this is the approved version of the following thesis:

Are the Criterion B Binge-Eating Disorder Symptoms Interchangeable in Conveying the Latent Binge-Eating Trait? An Item Response Theory Analysis

Chair: Kelsie T. Forbush, Ph.D.

Date Approved: ___ 24 February 2020 _____

Abstract

Binge-eating disorder (BED) is one of the most prevalent eating-disorder diagnoses among nationally representative samples of adults. Yet, few studies have evaluated the psychometrics of the BED diagnostic criteria. The BED Criterion B symptoms represent specific symptoms that may accompany binge-eating episodes. A minimum of three Criterion B symptoms must be endorsed to meet full BED diagnostic criteria. However, few studies have examined the Criterion B symptoms individually to clarify how well each symptom captures underlying BED pathology. The purpose of the present study was to use item response theory (IRT) to identify: 1) how much underlying binge-eating pathology must be present to endorse each symptom (i.e., item difficulty) and 2) how well each symptom differentiates among individuals with different levels of binge-eating pathology (i.e., item discrimination). Participants ($N=254$) were adults (85.0% female) with BED who completed eligibility screening for a BED group-treatment study. I hypothesized that ‘eating alone’ and ‘eating large amounts when not hungry’ would be the most ‘difficult’ items. Second, I hypothesized that ‘eating alone’ would demonstrate the greatest ability to discriminate those at the high end of the latent trait spectrum from relatively lower trait individuals. Results indicated that ‘uncomfortably full’ and ‘eating large amounts when not hungry’ had the highest difficulty parameters and were most difficult for high trait level individuals to endorse. ‘Eating rapidly’ and ‘eating alone’ were the most discriminating symptoms. Finally, the only symptom that provided unique item information at any trait level was ‘eating large amounts when not hungry.’ Results have implications for refining the Criterion B symptoms given the lack of unique item information relayed by the majority of the symptoms, the symptoms’ dimensionality, and variable monotonicity.

Table of Contents

Are the Criterion B Binge-Eating Disorder Symptoms Interchangeable in Conveying the Latent Binge-Eating Trait? An Item Response Theory

| | |
|------------------------------------|----|
| Analysis..... | 1 |
| BED Diagnostic Criteria..... | 1 |
| Criterion B Symptom Validity..... | 3 |
| Item Response Theory..... | 5 |
| Purpose and Hypotheses..... | 7 |
| Method..... | 8 |
| Participants and Procedure..... | 8 |
| Measures..... | 8 |
| Statistical Analyses..... | 9 |
| Results..... | 10 |
| Criterion B Symptom Frequency..... | 10 |
| EFA..... | 10 |
| IRT 2PL Model..... | 10 |
| Discussion..... | 11 |
| References..... | 17 |
| Appendix..... | 22 |

Are the Criterion B Binge-Eating Disorder Symptoms Interchangeable in Conveying the Latent Binge-Eating Trait? An Item Response Theory Analysis

Binge-eating disorder (BED) is characterized by the recurrent consumption of an objectively large amount of food in a discrete amount of time, such as two hours (American Psychiatric Association, 2013). Recent research estimates that the three-month prevalence of BED ranges from 5.6 to 6.9% among adults recruited from the community (Hay et al., 2015). BED is associated with increased annual healthcare costs and lower health-related quality-of-life (Agh et al., 2015). Specific medical conditions that drive the increased healthcare costs associated with BED include obesity and related conditions such as type II diabetes, asthma, gastrointestinal dysfunction, and high blood pressure (Olguin et al., 2017). Despite the clear public-health burden associated with BED, few studies have critically evaluated the psychometric properties of the BED diagnostic criteria. Critical evaluation of the BED diagnostic criteria at the symptom level is necessary to inform how specific binge-eating symptoms in BED are representative of severity.

BED Diagnostic Criteria

BED requires the presence of binge-eating episodes in the absence of behaviors to compensate for binge eating. Specifically, Criterion A1 specifies that a person must consume an abnormally large amount of food in a discrete period time given the circumstances (American Psychiatric Association, 2013). Individuals must also endorse experiencing a loss-of-control (LOC) over their eating during periods of over-consumption (Criterion A2; American Psychiatric Association, 2013). LOC refers to the subjective feeling that one could not stop eating once started. Criteria A1 and A2 represent the required components of a 'binge-eating episode.' Binge-eating episodes must occur at a minimum frequency of once per week for a three-month

period (Criterion D; American Psychiatric Association, 2013). To meet full diagnostic criteria for BED, individuals must also endorse at least three of the following Criterion B symptoms that may accompany their episodes: 1) 'eating rapidly,' 2) 'uncomfortably full,' 3) 'eating large amounts when not hungry,' 4) 'eating alone,' and 5) 'feeling disgusted/depressed/guilty' (American Psychiatric Association, 2013).

Initial evidence for the validity of BED as a distinct diagnosis was published by Spitzer et al. (1992). The authors conducted a field trial of preliminary diagnostic criteria for BED, including assessment of the 'associated features' that accompany binge eating (now called the Criterion B symptoms). The criteria were tested in several different samples, including: participants enrolled in weight control programs (i.e., receiving behavioral weight control treatment and nutrition counseling), participants formerly engaged in a weight control program, and a general sample of community adults. The authors examined how inclusion of the Criterion B symptoms as diagnostic criteria affected BED prevalence rates within each sample. Results indicated that including the Criterion B symptoms as diagnostic criteria only slightly lowered BED prevalence rates, which was taken as evidence that few individuals with BED did not endorse at least three of the Criterion B symptoms.

Spitzer et al. (1992) also assessed the internal consistency of the proposed BED diagnostic criteria, as well as how each associated feature correlated with the total number of diagnostic criteria that were fulfilled. In the sample of participants recruited from weight control programs, the correlations of Criterion B symptoms with the total symptom score ranged from .50 to .66. In the sample of community members, correlations ranged from .55 to .71. However, further interpretation of which symptoms were least or most strongly correlated with the total symptom score was limited because the authors only reported correlation ranges rather than

individual correlations. Thus, it was unclear from the original study that evaluated the BED criteria whether the Criterion B symptoms were interchangeable in terms of how strongly they corresponded with overall BED psychopathology.

Criterion B Symptom Validity

Although endorsement of any three Criterion B binge-eating symptoms is necessary for a diagnosis of BED, research demonstrated mixed findings for the validity of individual Criterion B symptoms (Klein et al., 2016; Vannucci et al., 2013; White & Grilo, 2011). One way that researchers have sought to examine the validity of the Criterion B symptoms is by assessing the predictive value of individual symptoms. White and Grilo (2011) tested the total predictive value (an index of the percent agreement derived from sensitivity, specificity, and positive/negative predictive power) for each Criterion B symptom in a sample of individuals with BED. Results demonstrated that ‘eating alone’ (B4) and ‘eating large amounts when not hungry’ (B3) were the best overall identifiers of binge-eating in persons with BED, supporting the criterion-related validity of Criterion B3 and B4. An additional aim of White and Grilo (2011) was to discern an optimal threshold for how many Criterion B symptoms should be endorsed to correctly identify cases of BED. The authors tested both positive and negative predictive power to identify an endorsement threshold for the Criterion B symptoms that maximized identification of ‘true’ BED cases and minimized incorrect BED diagnosis (i.e., false positives). The results from White and Grilo (2011) supported the current threshold of three binge-eating symptoms being endorsed in terms of maximizing sensitivity and specificity for identifying true BED cases.

Klein et al. (2016) assessed the extent to which each individual binge-eating symptom was associated with distress on the Eating Disorder Diagnostic Scale (Stice et al., 2000). In a logistic regression model, ‘feeling disgusted/depressed/guilty’ (B5) was the only Criterion B

symptom that uniquely predicted distress about binge eating or weight gain associated with binge eating. However, ‘eating large amounts when not hungry’ (B3) and ‘eating alone’ (B4) demonstrated significant correlations with distress of $r=.28$ and $r=.32$, respectively (Klein et al., 2016). The authors also utilized receiver-operating-characteristic (ROC) curves to calculate the area under the curve (AUC) to elucidate the optimal number of symptoms for predicting distress without compromising sensitivity and specificity. Results also supported the current three-symptom threshold, consistent with previous findings of White and Grilo (2011).

Finally, Vannucci et al. (2013) examined eating-disorder (ED) specific clinical impairment measured by the Clinical Impairment Assessment (CIA) in two groups of college-aged women who endorsed LOC eating. One group endorsed less than three Criterion B symptoms in addition to LOC, and one group endorsed three or more Criterion B symptoms in addition to LOC. Adjusted odds ratios from binary logistic regression analyses demonstrated no significant difference between groups in the odds of meeting the CIA threshold for clinically significant impairment; thus, meeting the *DSM-5* diagnostic threshold of endorsing three Criterion B symptoms was not related to increased odds of reporting clinically significant impairment due to an ED. Results from Vannucci et al. (2013) contrast the previously mentioned studies that did support the three-symptom threshold’s ability to discern true BED cases.

One additional way to further understand how well the Criterion B symptoms represent BED pathology is to consider how often each of the symptoms are endorsed by individuals who binge eat. For instance, items that correspond with a higher latent binge-eating trait presumably may only be endorsed by a small subset of individuals with more severe psychopathology. Two previously mentioned studies reported ‘eating alone’ (B4) was the least commonly endorsed symptom (White & Grilo, 2011; Klein et al., 2016). Thus, there may be preliminary evidence to

suggest that B4 may represent an uncommon symptom of binge eating that may discern individuals who experience more severe BED from those who have fewer BED problems. One previously mentioned study reported that ‘feeling disgusted/depressed/guilty’ was the most commonly endorsed symptom (White & Grilo, 2011). In contrast, Klein et al. (2016) identified ‘eating large amounts when not hungry’ (B3) as the most commonly endorsed symptom, and ‘feeling disgusted/depressed/guilty’ (B5) had a slightly lower endorsement rate. While results for the most commonly endorsed symptom are inconsistent in the literature, B3 and B5 are commonly endorsed items that may not distinguish persons with low versus high severity of BED, given they are endorsed by virtually all persons who meet criteria for a BED diagnosis.

In summary, research largely suggests that the three-symptom threshold for Criterion B symptoms has predictive validity for identifying clinical cases of BED. However, differences among the Criterion B symptoms are observed in terms of the symptoms’ associations with distress. These differences imply that the Criterion B symptoms may differ in how well they assess BED severity. To my knowledge, research has not yet assessed how well the Criterion B symptoms correspond with latent binge-eating. Therefore, the purpose of the current study was to apply item response theory (IRT) to critically evaluate the BED Criterion B symptoms. To date, no studies have applied IRT to understanding whether Criterion B symptoms are equally reflective of latent binge-eating severity in a sample of people with BED. Few studies within the ED field have utilized IRT, and no studies have applied IRT specifically to the BED diagnostic criteria.

Item Response Theory

Item response theory (IRT) is a psychometric method that originated in the field of education that can also be applied to the assessment of psychiatric disorders (Yang & Kao 2014).

IRT addresses an inherently problematic assumption of items within a scale: that items are equal in terms of their contributions to a respondent's total "trait" level (Thorpe & Favia, 2012). In IRT, the term "trait" refers to any unobservable characteristic that a set of items purportedly measures (e.g., depressive symptoms, well-being, etc.). IRT mathematically relates the latent trait spectrum, measured by a set of items (or symptoms), to participants' responses. The relationship between the latent trait and the probability of an item being endorsed, given latent trait level, is represented graphically by plotting each item on an item characteristic curve, or ICC (Ramsay & Reynolds, 2000). In this study, the ICC would be a graph of the relationship between latent level of binge eating and the probability of endorsing a given symptom (Baker, 2001). The ICC provides information about how an item functions across the latent trait spectrum by plotting how the probability of endorsing an item changes as latent trait level varies. The slope of an item's ICC indicates an item's ability to distinguish individuals with different trait levels based on their response to that item. An item with a gradually positively sloping ICC indicates a gradual increase in the probability that the item will be endorsed as trait level increases. In contrast, an item with a steep positive slope indicates a marked increase in the probability that an item will be endorsed as trait level increases (i.e., that item is better able to distinguish different latent trait levels).

Relatedly, an item information curve (IIC) is a curve that plots an item's information value across the latent trait spectrum. Item information has an inverse relationship with an item's measurement error, and therefore provides an index of an item's measurement precision across the latent trait spectrum. Item information value is positively related to how well an item discriminates at a given latent trait level. If an item information value is large for a given latent trait level, the value indicates that the item has good measurement precision at that trait level.

The width of an item information curve's peak conveys information about the proportion of the latent trait spectrum for which an item has favorable measurement precision for. For example, an item with a narrow information peak may be highly informative for a specific trait level, whereas an item with a broader information peak may be more informative for a larger proportion of the latent trait spectrum.

Purpose and Hypotheses

The purpose of this study was to use IRT to identify whether the Criterion B symptoms are equally reflective of latent binge-eating trait level in a sample of people with BED. The aims of this study were to use IRT to: 1) identify the relationship between latent trait level and Criterion B symptom endorsement (i.e., item difficulty) and 2) examine how the probability of endorsement for each symptom changes across the latent trait spectrum (i.e., item discrimination). Specifically, for a specific Criterion B symptom to have a high item discrimination value, the probability of the symptom being endorsed should sharply increase as the latent trait level increases. In contrast, a symptom with poor discrimination would show minimal change in the probability of endorsing the symptom, even as the latent trait level increases.

Based on findings from Klein et al. (2016) and White and Grilo (2011) that 'eating alone' (B4) and 'eating large amounts when not hungry' (B3) were significantly correlated with distress and demonstrated favorable criterion-related validity, I hypothesize that B4 and B3 will represent the two most 'difficult' Criterion B symptoms (i.e., symptoms with higher trait level being necessary for endorsement). Second, I hypothesize that 'eating alone' (B4) will demonstrate high discriminability at high latent trait levels, given this item's relatively low rate of endorsement

among individuals with BED reported in both Klein et al. (2016) and White and Grilo (2011) in addition to this item's significant association with distress.

Method

Participants and Procedure

Participants were adults ($N=254$) with full-threshold BED recruited from two Midwestern communities as a part of a BED treatment study conducted at two Midwestern sites (Peterson et al., 2009). The sample was predominantly comprised of Caucasian women with overweight and obesity (see Table 1). Treatment conditions included self-help group cognitive-behavioral therapy, therapist-led group cognitive behavioral therapy, therapist-assisted group cognitive-behavioral therapy, and a waitlist control condition. Only baseline data were utilized for this study. Participants were recruited through community advertisements or referrals from ED and general-health providers.

Measures

Eating Disorder Examination

Participants completed the Eating Disorder Examination (EDE), a semi-structured interview that assesses ED-specific psychopathology and behaviors, with a primary focus on the previous 28 days as well as the two months prior (Fairburn & Cooper, 1993; Cooper & Fairburn, 1987). ED diagnoses can be derived from the EDE. The EDE assesses the *DSM-5* Criterion B binge-eating symptoms in relation to binge-eating episodes that the participant considers representative of episodes that occurred over the previous six months. A question about each Criterion B symptom yields a dichotomous response (0 = Feature not present; 1 = Feature present) for each symptom. The EDE also yields a global score and four subscales: Restraint, Eating Concern, Shape Concern, and Weight Concern. Interrater reliability was reported by the

authors of the original study with intraclass correlation coefficients; the authors reported that coefficients for the EDE global score and subscales ranged from .955 to .982 (Peterson et al., 2009). Relative to EDE norms for treatment-seeking individuals with BED, the sample was within one standard deviation above or below average for the four EDE subscales (Wilfley et al., 2000). Thus, the levels of ED psychopathology as measured by the EDE in this sample appear normative for individuals with BED who seek treatment.

Inventory of Depressive Symptomatology-Self Report

Participants completed the Inventory of Depressive Symptomatology-Self Report (IDS-SR; Rush et al., 2000), a 30-item measure of depressive symptoms experienced over the past week. Items measure depressive symptoms on a four-point Likert scale, and the measure generates a total score (with a maximum score of 84). The IDS-SR demonstrates good internal consistency, convergent validity with other established measures of depression, and concurrent validity (Trivedi et al., 2004; Rush et al., 1996; Rush et al., 2000; Corruble et al., 1999). In the current sample, internal consistency was acceptable (Cronbach's $\alpha=.875$). Based on previous research conducted by Corruble et al. (1999) examining IDS-SR scores in a sample of individuals with varying depressive symptom levels, IDS-SR mean scores in this sample at baseline indicate the presence of mild to moderate depressive symptoms (see Table 1).

Statistical Analyses

Frequencies and descriptive statistics for various participant characteristics (e.g., ethnicity, age, BMI, 3-month OBE total) and Criterion B symptom endorsement were calculated in SPSS statistical software (version 24).

Unidimensionality of the Criterion B symptoms was assessed in Mplus statistical software using exploratory factor analysis (EFA) for categorical items. IRT analyses were

conducted using the R package ‘ltm’ (Rizopoulos & Rizopoulos, 2018). Using ‘ltm,’ the Criterion B symptoms were examined in a two-parameter logistic (2PL) IRT model, which estimates parameters for both difficulty and discrimination. Both ICCs and item information curves (IICs) were plotted for the five Criterion B symptoms in R.

Results

Criterion B Symptom Frequency

The frequency of each Criterion B symptom’s endorsement is displayed in Table 2. The most commonly endorsed symptom in the sample was ‘eating large amounts when not hungry’ (B3) whereas the least commonly endorsed symptom was ‘eating rapidly’ (B1).

EFA

Mplus statistical software was utilized to run one-factor and two-factor EFA models to evaluate the unidimensionality of the five Criterion B symptoms (i.e., to assess the extent to which the symptoms represent a shared factor). Results indicated that the one-factor Criterion B symptoms model did not have a good fit to the data ($\chi^2=11.742$ [5], $p<.05$, CFI=.839, TLI=.679, RMSEA=.073), implying that the Criterion B symptoms are not unidimensional. Alternatively, results indicated that the two-factor Criterion B symptoms model did provide good fit to the data ($\chi^2=.054$ [1], $p>.05$, CFI=1.000, TLI=1.225, RMSEA=.000), suggesting that the symptoms may represent related but separate facets of binge eating.

IRT 2PL Model

A 2PL IRT model containing the five Criterion B symptoms estimated difficulty and discrimination parameters for each Criterion B symptom in R statistical software (see Table 3). When plotted on an ICC (see Figure 1), ‘eating rapidly’ (B1) and ‘eating alone’ (B4) demonstrated clear monotonicity (i.e., the probability of endorsement for these symptoms

increased as the binge-eating trait level increased). ‘Feeling disgusted/depressed/guilty’ (B5) was slightly monotonic. ‘Eating large amounts when not hungry’ (B3) and ‘uncomfortably full’ (B2) demonstrated decreased likelihood of endorsement as the binge-eating trait level increased and, thus, were not monotonic; their decreased likelihood of endorsement as the trait increased is reflected in these symptoms’ negative discrimination parameters.

‘Eating large amounts when not hungry’ (B3) had a markedly lower probability of being endorsed at the high end of the latent trait spectrum (see Figure 1). ‘Feeling disgusted/depressed/guilty’ (B5) was overall the least difficult symptom to endorse. The symptoms with the highest overall difficulty parameters were ‘uncomfortably full’ (B2) and ‘eating large amounts when not hungry’ (B3); as noted previously, these items demonstrated decreased likelihood of endorsement as trait level increased, and thus their high difficulty parameters reflect their rarity of endorsement among higher trait levels. In contrast, ‘eating rapidly’ (B1) and ‘eating alone’ (B4) had lower raw difficulty parameters than B3 and B2, and were more difficult for lower trait level individuals to endorse. Finally, B1 and B4 had the highest discrimination parameters, with the two symptoms’ similar ICCs indicating that the symptoms discriminated well between above average and below average trait levels.

IICs for the Criterion B symptoms indicated that only one symptom conveyed differential item information at a particular trait level: ‘eating large amounts when not hungry’ (B3) conveyed maximal information at high trait levels (see Figure 2). In contrast, the remaining Criterion B symptoms demonstrated flat item information curves across trait levels.

Discussion

This was the first study to use IRT to identify whether the Criterion B BED symptoms are equally reflective of latent binge-eating trait level. Although hypotheses were not fully

supported, results revealed critical information about the utility of BED symptoms for assessing binge-eating pathology as reflected by the Criterion B symptoms. As I hypothesized, 'eating large amounts when not hungry' (B3) had a relatively high difficulty parameter; however, B3 was more difficult to endorse as trait level increased, contrary to my hypothesis. This result may reflect the relationship between binge eating and lowered sensitivity to satiety (Boutelle et al., 2017). The lack of endorsement of B3 by high trait individuals may also reflect a tendency to misinterpret hedonic hunger (i.e., wanting to eat as a means of pleasure despite a lack of need for energy intake) as physical hunger (Lowe & Butryn, 2007). I also hypothesized that 'eating alone' (B4) would be one of the most difficult symptoms (require greater trait level for endorsement relative to other symptoms), which results did not fully support. Among the three symptoms for which the probability of endorsement increased with trait level, B4 was more difficult than 'feeling disgusted/depressed/guilty' (B5), but less difficult than 'eating rapidly' (B1). The lack of full support for this hypothesis may reflect the fact that binge eating commonly occurs when people are alone, rather than when people are in social settings (Latner & Clyne, 2008).

I also hypothesized that 'eating alone' (B4) would produce the highest discrimination parameter. Somewhat consistent with this hypothesis, B4 emerged with the second highest discrimination parameter. As indicated by discrimination parameter estimates and ICCs, 'eating rapidly' (B1) showed the steepest increase in probability of endorsement as the binge-eating trait increased. Specifically, B1 was able to discriminate individuals with below-average vs. above-average trait levels of binge eating. Additionally, 'eating alone' (B4) demonstrated a gradual increase in the probability of endorsement as the binge-eating trait level increased. In contrast, 'eating large amounts when not hungry' (B3) demonstrated a negative discrimination parameter

due to individuals at the highest end of the binge-eating trait spectrum having a decreased likelihood of endorsing this symptom. ‘Uncomfortably full’ (B2) also demonstrated a slight negative discrimination parameter.

Four out of the five Criterion B symptoms did not demonstrate variability in amount of item information across binge-eating trait level: ‘eating rapidly’ (B1), ‘uncomfortably full’ (B2), ‘eating alone’ (B4), and ‘feeling disgusted/depressed/guilty’ (B5). The uniformity in item information observed is consistent with research conducted by Klein et al. (2016) that found only ‘feeling disgusted/depressed/guilty’ (B5) uniquely predicted distress about binge eating. The lack of observed variability in item information may be due to the fact that this study consisted of individuals who met full diagnostic criteria for BED and, thus, all participants in my study endorsed at least three Criterion B symptoms; therefore, there may not be as much variability in symptom endorsement patterns relative to samples that also included individuals with subthreshold BED. In contrast, ‘eating large amounts when not hungry’ (B3) provided maximal information for individuals with the highest binge-eating trait level.

One notable limitation of this study was that the sample consisted exclusively of individuals who met full BED diagnostic criteria; thus, all participants endorsed at least three out of five Criterion B symptoms. The range of the binge-eating trait represented in analyses in the current study is, therefore, restricted. Moreover, there were certain symptoms that were endorsed by nearly the entire sample. Had this study been conducted in a sample in that included subthreshold BED or among a community sample that was not recruited specifically for ED concerns, greater variability in patterns of symptom endorsement and binge-eating trait levels may have been represented. Variability in the sample is particularly important for assessing item information; items that are endorsed at equivalent rates across trait levels and/or that are

endorsed by everyone may not convey unique information. However, given that the purpose of this study was to evaluate the information that each Criterion B symptom provides about individuals with BED, results are nonetheless informative.

Another limitation of this study is the violation of certain IRT statistical assumptions. As previously noted, a one-factor model containing the Criterion B symptoms did not provide a good fit to the data, indicating that the Criterion B symptoms are not unidimensional and that the unidimensionality assumption of IRT was not met. However, research on how the existence of multidimensionality among items influences IRT model analyses indicates that parameter estimates (i.e., difficulty and discrimination) are minimally impacted by violation of the unidimensionality assumption (Crişan et al., 2017; Bonifay et al., 2015). Relatedly, the IRT assumption of local independence among items was inherently violated by this research question because the BED diagnostic criteria require the endorsement of multiple Criterion B symptoms. Finally, results indicated that some but not all Criterion B symptoms were monotonic (i.e., that the probability of the symptom being endorsed increased as the binge-eating trait increased). For example, ‘eating large amounts when not hungry’ (B3) was less likely to be endorsed by individuals at the high end of the trait spectrum. However, the finding that B3 may be more representative of less pathological presentations is clinically informative when considering how well this item distinguished individuals with high binge-eating trait levels from individuals with lower levels. Nonetheless, given the violation of the monotonicity assumption in IRT, results should be interpreted with caution.

Finally, a limitation of this study is the ethnic and racial homogeneity of the sample. Pooled data from large scale epidemiological studies suggest that lifetime binge eating is more prevalent among ethnic minority groups relative to non-Latino Caucasians (Marques et al.,

2011). Thus, future studies are needed to replicate findings in a sample with greater ethnic and racial diversity.

Despite limitations, this study has notable strengths. First, this study examined Criterion B symptoms in a sample of individuals diagnosed with BED through semi-structured clinical interviews, whereas most prior research on the Criterion B symptoms relied on self-report questionnaires to assess BED symptoms. Given the nuanced nature of certain BED symptoms such as LOC and the consumption of an objectively large amount of food, assessment of BED through semi-structured clinical interviews rather than self-report questionnaires may be preferable (Fairburn & Beglin, 1994). Second, this study is the first to utilize a latent trait statistical method to assess if the Criterion B symptoms represent a shared latent construct and to evaluate how latent levels of the binge-eating construct were associated with Criterion B symptom endorsement.

There are several directions for future research. First, researchers could further explore the dimensionality of the Criterion B symptoms in individuals with full- and sub-threshold BED and, if further research supports the Criterion B symptoms' multidimensionality, potentially revise the diagnostic criteria to more accurately characterize BED symptoms. For example, perhaps there are two facets of binge eating measured by the Criterion B symptoms that could represent BED subtypes or diagnostic specifiers. However, to my knowledge, no other research thus far has examined the factor structure of the Criterion B symptoms. Second, given the result that certain symptoms (such as B5, 'feeling disgusted/depressed/guilty') provided minimal unique item information when endorsed, future research should consider the incremental validity of each Criterion B symptom to eliminate potential redundancy among the BED symptoms in future revisions of the diagnostic criteria. For example, results of this study indicated that

‘feeling disgusted/depressed/guilty’ (B5) was the least difficult symptom to endorse, which may reflect this symptom’s overlap with BED’s separate distress criterion (Criterion C) that requires the additional presence of notable distress pertaining to the person’s binge eating behavior to meet full diagnostic criteria. Future research on the Criterion B symptoms could support narrowing of the diagnostic criteria to evaluate if the Criterion B symptoms provide meaningful diagnostic information beyond LOC. This study did not include LOC in analyses because all individuals with BED must endorse LOC.

In conclusion, this was the first study to use IRT to identify the item difficulty and item discrimination of the BED Criterion B symptoms. A key implication of this study is that the Criterion B symptoms may not provide incremental information in characterizing binge-eating pathology among individuals who meet full diagnostic criteria for BED. Additionally, the lack of item information provided by the majority of the Criterion B symptoms implies that the Criterion B symptoms may have high measurement error. To the extent that required symptoms do not add diagnostically useful information or have non-negligible measurement error, individuals may be misdiagnosed as having sub-threshold BED despite having a high level of binge-eating pathology that warrants treatment. Additionally, clinicians assessing for BED may erroneously assume that endorsing a given Criterion B symptom reflects an aspect of binge eating that is shared by all the Criterion B symptoms or that endorsement always equates to higher pathology, whereas results of this study suggest that the Criterion B symptoms are not unidimensional or totally monotonic. Thus, continued refinement of the BED diagnostic criteria is warranted to ensure accurate diagnosis and meaningful symptom assessment.

References

- Agh, T., Kovács, G., Pawaskar, M., Supina, D., Inotai, A., & Vokó, Z. (2015). Epidemiology, health-related quality of life and economic burden of binge eating disorder: a systematic literature review. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 20(1), 1-12. doi: <https://doi-org/10.1007/s40519-014-0173-9>
- American Psychiatric Association. (2013). *Feeding and Eating Disorders*. In *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC.
- Baker, F. B. (2001). *The basics of item response theory* (2nd ed.). ERIC.
- Bonifay, W.E., Reise, S.P., Scheines, R., & Meijer, R.R. (2015). When are multidimensional data unidimensional enough for structural equation modeling? An evaluation of the DETECT multidimensionality index. *Structural Equation Modeling: A Multidisciplinary Journal*, 22, 504-516. doi: <https://doi-org/10.1080/10705511.2014.938596>
- Boutelle, K.N., Knatz, S., Carlson, J., Bergmann, K., & Peterson, C.B. (2017). An Open Trial Targeting Food Cue Reactivity and Satiety Sensitivity in Overweight and Obese Binge Eaters. *Cognitive and Behavioral Practice*, 24(3), 363-373. doi: <https://doi-org/10.1016/j.cbpra.2016.08.003>
- Cooper, Z., & Fairburn, C. (1987). The eating disorder examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders*, 6(1), 1-8. doi: [https://doi-org/10.1002/1098-108X\(198701\)6:1<1::AID-EAT2260060102>3.0.CO;2-9](https://doi-org/10.1002/1098-108X(198701)6:1<1::AID-EAT2260060102>3.0.CO;2-9)
- Corruble, E., Legrand, J.M., Duret, C., Charles, G., & Guelfi, J.D. (1999). IDS-C and IDS-SR: Psychometric properties in depressed in-patients. *Journal of Affective Disorders*, 56(2-3), 95-101. doi: [https://doi-org/10.1016/S0165-0327\(99\)00055-5](https://doi-org/10.1016/S0165-0327(99)00055-5)

- Crişan, D.R., Tendeiro, J.N., & Meijer, R.R. (2017). *Applied Psychological Measurement*, 41(6), 439-455. doi: <https://doi-org/10.1177/0146621617695522>
- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16(4), 363-370. doi: [https://doi-org/10.1002/1098-108X\(199412\)16:4<363::AID-EAT2260160405>3.0.CO;2-%23](https://doi-org/10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-%23)
- Fairburn, CG.; Cooper, Z. (1993). The Eating Disorder Examination. In: Fairburn, CG.; Wilson, GT., editors. *Binge eating: Nature, assessment, and treatment*. 12th. New York: Guilford Press; p. 317-360.
- Hay, P., Girosi, F., & Mond, J. (2015). Prevalence and sociodemographic correlates of DSM-5 eating disorders in the Australian population. *J Eat Disord*, 3, 19. <https://doi-org/doi:10.1186/s40337-015-0056-0>
- Klein, K. M., Forney, K. J., & Keel, P. K. (2016). A preliminary evaluation of the validity of binge-eating disorder defining features in a community-based sample. *International Journal of Eating Disorders*, 49(5), 524-528. doi: <https://doi-org/10.1002/eat.22479>
- Latner, J. D., & Clyne, C. (2008). The diagnostic validity of the criteria for binge eating disorder. *International Journal of Eating Disorders*, 41(1), 1-14. doi: <https://doi-org/10.1002/eat.20465>
- Lowe, M.R., & Butryn, M.L. (2007). Hedonic hunger: A new dimension of appetite? *Physiology & Behavior*, 91(4), 432-439. doi: <https://doi-org/10.1016/j.physbeh.2007.04.006>
- Marques, L., Alegria, M., Becker, A. E., Chen, C. n., Fang, A., Chosak, A., & Diniz, J. B. (2011). Comparative prevalence, correlates of impairment, and service utilization for eating disorders across US ethnic groups: Implications for reducing ethnic disparities in

- health care access for eating disorders. *International Journal of Eating Disorders*, 44(5), 412-420. doi: <https://doi-org/10.1002/eat.20787>
- Olguin, P., Fuentes, M., Gabler, G., Guerdjikova, A. I., Keck, P. E., & McElroy, S. L. (2017). Medical comorbidity of binge eating disorder. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 22(1), 13-26. doi: <https://doi-org/10.1007/s40519-016-0313-5>
- Peterson, C. B., Mitchell, J. E., Crow, S. J., Crosby, R. D., & Wonderlich, S. A. (2009). The efficacy of self-help group treatment and therapist-led group treatment for binge eating disorder. *American Journal of Psychiatry*, 166(12), 1347-1354. doi: <https://doi-org/10.1176/appi.ajp.2009.09030345>
- Ramsay, M. C., & Reynolds, C. R. (2000). Development of a scientific test: A practical guide. *Handbook of psychological assessment*, 21-42.
- Rizopoulos, D., & Rizopoulos, M. D. (2018). Package 'ltm'. URL <http://wiki.r-project.org/rwiki/doku.php>.
- Rush, A.J., Carmody, T., & Reimitz, P.E. (2000). The Inventory of Depressive Symptomatology (IDS): clinician (IDS-C) and self-report (IDS-SR) ratings of depressive symptoms. *International journal of methods in psychiatric research*, 9(2), 45-59. doi: <https://doi-org/10.1002/mpr.79>
- Rush, A.J., Gullion, C.M., Basco, M.R., Jarrett, R.B., & Trivedi, M.H. (1996). The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychological Medicine*, 26(3), 477-486. doi: <https://doi-org/10.1017/S0033291700035558>
- Spitzer, R. L., Devlin, M., Walsh, B. T., Hasin, D., Wing, R., Marcus, M., . . . Agras, S. (1992).

- Binge eating disorder: A multisite field trial of the diagnostic criteria. *International Journal of Eating Disorders*, 11(3), 191-203. doi: [https://doi-org/10.1002/1098-108X\(199204\)11:3<191::AID-EAT2260110302>3.0.CO;2-S](https://doi-org/10.1002/1098-108X(199204)11:3<191::AID-EAT2260110302>3.0.CO;2-S)
- Stice, E., Telch, C. F., & Rizvi, S. L. (2000). Development and validation of the Eating Disorder Diagnostic Scale: a brief self-report measure of anorexia, bulimia, and binge-eating disorder. *Psychological assessment*, 12(2), 123. doi: <https://doi-org/10.1037//1040-3590.12.2.123>
- Thorpe, G. L., & Favia, A. (2012). Data Analysis Using Item Response Theory Methodology: An Introduction to Selected Programs and Applications. *Psychology Faculty Scholarship*, 20.
- Trivedi, M.H., Rush, A.J., Ibrahim, H.M., Carmody, T.J., Biggs, M.M., Suppes, T.,...Kashner, T.M. (2004). The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychological Medicine*, 34(1), 73-82. doi: <https://doi-org/10.1017/S0033291703001107>
- Vannucci, A., Theim, K. R., Kass, A. E., Trockel, M., Genkin, B., Rizk, M., . . . Aspen, V. (2013). What constitutes clinically significant binge eating? Association between binge features and clinical validators in college-age women. *International Journal of Eating Disorders*, 46(3), 226-232. doi: <https://doi-org/10.1002/eat.22115>
- White, M. A., & Grilo, C. M. (2011). Diagnostic efficiency of DSM–IV indicators for binge eating episodes. *Journal of Consulting and Clinical Psychology*, 79(1), 75. doi: <https://doi-org/10.1037/a0022210>

- Wilfley, D.E., Schwartz, M.B., Spurrell, E.B., & Fairburn, C.G. (2000). Using the Eating Disorder Examination to Identify the Specific Psychopathology of Binge Eating Disorder. *International Journal of Eating Disorders*, 27(3), 259-269. doi: [https://doi-org/10.1002/\(SICI\)1098-108X\(200004\)27:3<259::AID-EAT2>3.0.CO;2-G](https://doi-org/10.1002/(SICI)1098-108X(200004)27:3<259::AID-EAT2>3.0.CO;2-G)
- Yang, F. M. (2014). Item response theory for measurement validity. *Shanghai Archives of Psychiatry*, 26(3), 171. doi: <https://doi-org/10.3969/j.issn.1002-0829.2014.03.010>

Appendix

| | |
|--|----|
| Table 1. Participant Characteristics..... | 23 |
| Table 2. Criterion B Symptom Endorsement Frequency..... | 24 |
| Table 3. Criterion B Symptom Difficulty and Discrimination Parameters..... | 25 |
| Figure 1. Criterion B Symptom Item Characteristic Curves..... | 26 |
| Figure 2. Criterion B Symptom Item Information Curves..... | 27 |

Table 1***Participant Characteristics***

| <hr/> Characteristic <hr/> | |
|--|---------------|
| Gender <i>n</i> (%) | |
| Female | 216 (85.0) |
| Male | 31 (12.2) |
| Missing | 7 (2.8) |
| Ethnicity <i>n</i> (%) | |
| Caucasian | 238 (93.0) |
| African American | 2 (.8) |
| Native American | 1 (.4) |
| Multiracial | 1 (.4) |
| Other | 5 (2.0) |
| Missing | 7 (2.8) |
| Psychiatric Medication <i>n</i> (%) | |
| Yes | 194 (75.8) |
| No | 49 (19.1) |
| Missing | 11 (4.3) |
| Age <i>M</i> (<i>SD</i>) | 47.1 (10.5) |
| Missing <i>n</i> (%) | 7 (2.8) |
| BMI <i>M</i> (<i>SD</i>) | 39.0 (7.9) |
| Missing <i>n</i> (%) | 2 (.8) |
| OBE 3-Mo. Total | 66.30 (41.83) |
| EDE | |
| Eating Concern | 1.88 (1.23) |
| Restraint | 1.56 (1.32) |
| Weight Concern | 3.38 (1.10) |
| Shape Concern | 3.57 (1.03) |
| IDS-SR | 24.81 (11.42) |

Note. BMI: Body Mass Index. OBE: Objective Binge-Eating Episodes. EDE: Eating Disorder Examination. IDS-SR: Inventory of Depressive Symptomatology-Self Report.

Table 2***Criterion B Symptom Endorsement Frequency***

| Criterion B Symptom | Frequency and Percentage of Sample that Endorsed [<i>n</i> (%)] |
|---|--|
| Eating rapidly (B1) | 173 (68.1) |
| Uncomfortably full (B2) | 234 (92.1) |
| Eating large amounts when not hungry (B3) | 251 (98.8) |
| Eating alone (B4) | 185 (72.8) |
| Feeling disgusted/depressed/guilty (B5) | 249 (98.0) |

Table 3*Criterion B Symptom Difficulty and Discrimination Parameters*

| Symptom | <i>b</i> (Difficulty) | <i>a</i> (Discrimination) |
|---|--------------------------|------------------------------|
| Eating rapidly (B1) | -.721 | 1.488 |
| Uncomfortably full (B2) | 9.692 | -0.257 |
| Eating large amounts when not hungry (B3) | 2.184 | -14.332 |
| Eating alone (B4) | -1.049 | 1.249 |
| Feeling disgusted/depressed/guilty (B5) | -45.814 | .085 |

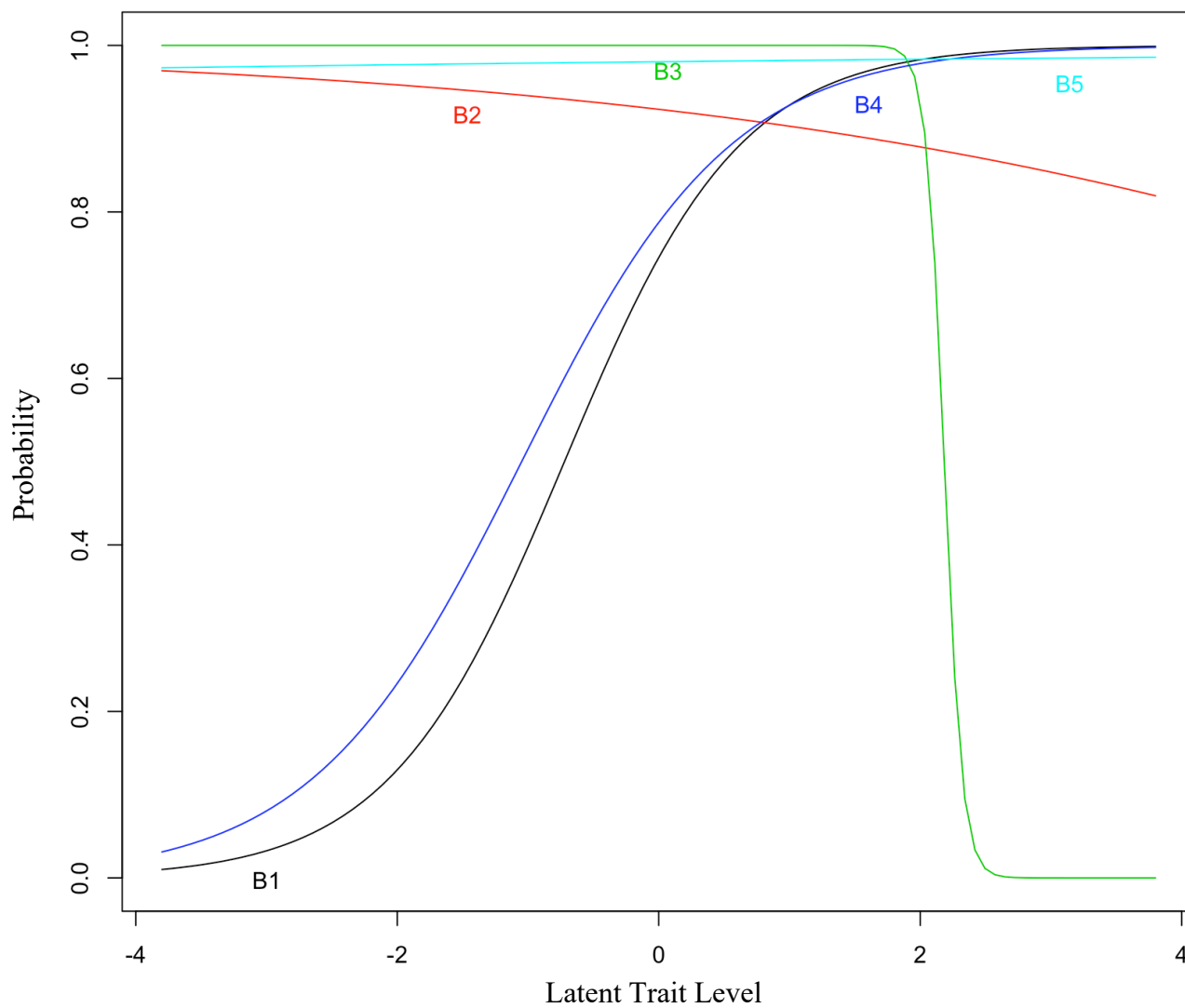


Figure 1. Criterion B Symptom Item Characteristic Curves

Note. B1: Eating rapidly. B2: Uncomfortably full. B3: Eating large amounts when not hungry.

B4: Eating alone. B5: Feeling disgusted/depressed/guilty.

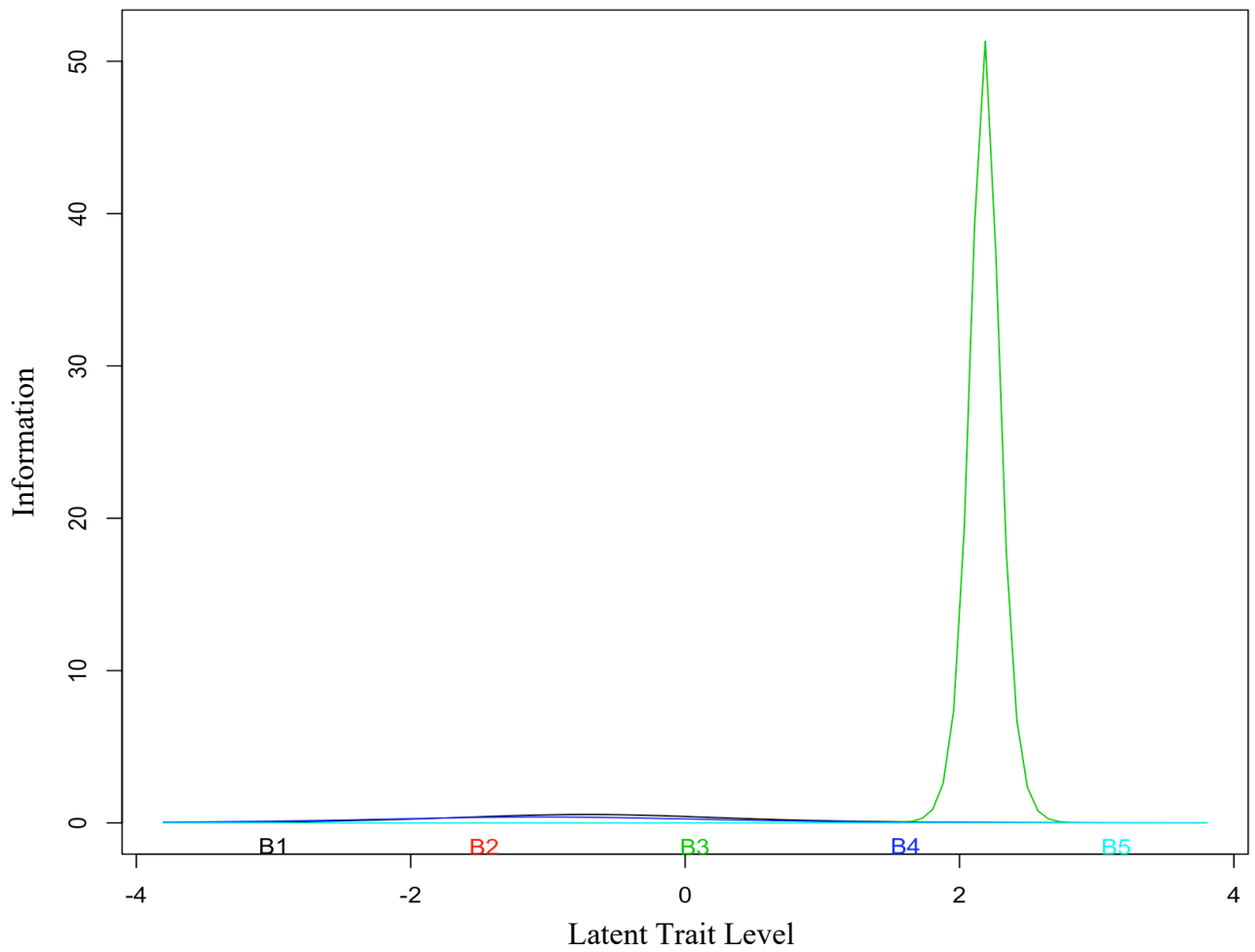


Figure 2. Criterion B Symptom Item Information Curves

Note. B1: Eating rapidly. B2: Uncomfortably full. B3: Eating large amounts when not hungry.

B4: Eating alone. B5: Feeling disgusted/depressed/guilty.