



Published in final edited form as:

Eat Behav. 2008 December ; 9(4): 398–407. doi:10.1016/j.eatbeh.2008.04.004.

Evaluation of Diagnostic Criteria for Night Eating Syndrome Using Item Response Theory Analysis

Kelly C. Allison, Ph.D^a, Scott G. Engel, Ph.D^b, Ross D. Crosby, Ph.D^{b,c}, Martina de Zwaan, M.D^{b,1}, John P. O'Reardon, M.D^a, Stephen A. Wonderlich, Ph.D^{b,c}, James E. Mitchell, M.D^{b,c}, Delia Smith West, Ph.D^d, Thomas A. Wadden, Ph.D^a, and Albert J. Stunkard, M.D.^a

^a University of Pennsylvania School of Medicine, Department of Psychiatry, Center for Weight and Eating Disorders, 3535 Market St., 3rd Floor, Philadelphia, PA 19104-3309, USA

^b Neuropsychiatric Research Institute, 120 South 8th St., Box 1415, Fargo, ND 58107-1415, USA

^c Department of Clinical Neuroscience, University of North Dakota School of Medicine, Fargo, North Dakota, USA

^d University of Arkansas Medical Sciences, 4301 West Markham, #820, Little Rock, AR 72205, USA

Abstract

Uniform diagnostic criteria for the night eating syndrome (NES), a disorder characterized by a delay in the circadian pattern of eating, have not been established. Proposed criteria for NES were evaluated using item response theory (IRT) analysis. Six studies yielded 1,481 Night Eating Questionnaires which were coded to reflect the presence/absence of five night eating symptoms. Symptoms were evaluated based on the clinical usefulness of their diagnostic information and on the assumptions of IRT analysis (unidimensionality, monotonicity, local item independence, correct model specification), using a two parameter logistic (2PL) IRT model. Reports of (1) nocturnal eating and/or evening hyperphagia, (2) initial insomnia, and (3) night awakenings showed high precision in discriminating those with night eating problems, while morning anorexia and delayed morning meal provided little additional information. IRT is a useful tool for evaluating the diagnostic criteria of psychiatric disorders and can be used to evaluate potential diagnostic criteria of NES empirically. Behavioral factors were identified as useful discriminators of NES. Future work should also examine psychological factors in conjunction with those identified here.

Keywords

IRT; diagnosis; Night Eating Questionnaire; nocturnal ingestions; evening hyperphagia; eating disorder

Original criteria for the night eating syndrome (NES) were based on a patient in whom the disorder was first noted and the subsequent treatment of 25 obese persons referred to a Special Study Clinic because of difficulties managing their obesity (Stunkard, Grace, & Wolff,

Address correspondence to Kelly C. Allison, Ph.D., University of Pennsylvania, Department of Psychiatry, Center for Weight and Eating Disorders, 3535 Market St., Suite 3021, Philadelphia, PA 19104-3309, USA, kca@mail.med.upenn.edu, phone: 215-898-7314; fax: 215-898-2878.

¹Present address: Department of Psychosomatic Medicine and Psychotherapy, University of Erlangen-Nuremberg, Erlangen, Germany

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1955). These criteria were: (1) consumption of at least 25% of daily caloric intake after the evening meal, (2) initial insomnia at least half of the time, and (3) morning anorexia. Revision of the criteria in 1999 added: (1) nighttime awakenings (at least one per night), (2) frequently accompanied by the ingestion of snacks; (3) an increase in the requirement for evening hyperphagia to greater than 50% of daily caloric intake; (4) a duration of three months; and (5) absence of bulimia nervosa and binge eating disorder (BED) (Birketvedt et al., 1999). The last criterion has not been included in other studies, while initial insomnia was not considered in this report.

Studies of NES have employed variations of these criteria (see de Zwaan, Burgard, Schenck, & Mitchell, 2003 for review), including the use of different cut-times for evening hyperphagia: 7 pm (Cerú-Björk, Andersson, & Rössner, 2001; Striegel-Moore, Franko, Thompson, Affenito, & Kraemer, 2006; Stunkard et al., 1996), *after the evening meal* (Birketvedt et al., 1999, Marshall, Allison, O'Reardon, Birketvedt, & Stunkard, 2004; O'Reardon et al., 2004), after 11pm (Striegel-Moore et al., 2006), and simply, *evening eating* (Kuldau & Rand, 1986; Rand et al., 1997). The proportion of calories used to define evening hyperphagia has also varied by study, from 25% (Stunkard et al., 1955, Allison, Grilo, Masheb, & Stunkard, 2005; Allison et al., 2006) to 50% (Birketvedt et al., 1999; Stunkard et al., 1996), to simply *overeating* or *excessive* eating during the evening (Kuldau & Rand, 1986; Rand, MacGreggor, & Stunkard, 1997). Some NES studies have examined putative NES samples which have had nocturnal ingestions *or* evening hyperphagia (Cerú-Björk et al., 2001, Grilo & Masheb, 2004; Napolitano, Head, Babyak, & Blumenthal, 2001) as well as samples that included persons who had either of these or both criteria (Allison et al., 2007; Allison et al., 2006; Allison et al., 2005; O'Reardon et al., 2004; Pawlow, O'Neil, & Malcolm, 2003). Striegel-Moore and colleagues (2004, 2006) tested several of these definitions with a sample of adolescent girls and a representative national sample ages 13 and older, demonstrating decreases in the proportion of those reporting nocturnal eating with increasingly stringent criteria.

Two prevalence studies of NES in the general population for adults revealed rates of 1.5% (Rand et al., 1997) and 5.2% (Lamerz et al., 2005). Prevalence estimates in obesity clinics are higher, ranging from 6% (Cerú-Björk et al., 2001) to 14% (Gluck, Geliebter, & Satov, 2001), and in pre-operative bariatric surgery patients from 8% (Adami, Meneghelli, & Scopinaro, 1999; Allison et al., 2006) to 42% (Hsu, Bentancourt, & Sullivan, 1996). Among psychiatric outpatients 12% have NES, with obese patients 5.2 times more than normal weight patients to be diagnosed (Lundgren et al., 2006). Differences in diagnostic criteria across studies, including the timeframe used for evening hyperphagia and assessment of nocturnal ingestions, may contribute to the inconsistencies in this relationship.

Depressed mood has also been linked to NES in several studies (Birketvedt et al., 1999, O'Reardon et al., 2004; Gluck et al., 2001; de Zwaan, Roerig, Crosby, Karaz, & Mitchell, 2006), with mood falling as the day progresses for many (Birketvedt et al., 1999). Lifetime occurrence of major depressive disorder based on DSM-IV diagnostic criteria is high at 56% (de Zwaan et al., 2006).

Finally, NES has a low to moderate overlap with BED, ranging from 0% to 26.5% (see de Zwaan et al. 2003 for review). Comparisons between individuals with these disorders suggest that they are different constructs (Allison et al., 2005). Probably most specifically, those with NES have more disturbed sleep, and their nocturnal ingestions are not objectively large (Birketvedt et al., 1999). More work is needed regarding the relationship between the timing of binges of those with BED and the concept of evening hyperphagia. However, there is evidence suggests that NES and BED are distinct; thus, NES needs its own, valid diagnostic criteria.

To assess the core features and psychopathology associated with NES, we used the Night Eating Questionnaire (Allison et al., 2008) which contains items descriptive of NES (see Table 1) and has an alpha for the total score of .70. Vander Wal, Waller, Klurfeld, McBurney, and Dhurandhar (2005) verified that the NEQ was positively correlated with increasingly complex definitions of NES. We sought to understand how each of the features described clinically and contained in the NEQ is related to the overall construct of night eating syndrome by implementing Item Response Theory techniques.

Item Response Theory (IRT)

IRT is a theoretical framework for psychological measurement. It provides an attractive alternative to Classical Test Theory (CTT) because of its potential for solving many practical testing problems (Lord, 1980; Weiss & Yoes, 1991). CTT and IRT share the assumption that measurement involves the location of individuals along some latent (i.e., not directly measurable) dimension, typically referred to as the *trait level*. The primary distinction between the methods is that IRT, unlike CTT, is a model-based theory of measurement. IRT assumes that a person's trait level can be estimated from responses to individual items. An IRT model specifies precisely how a person's response to an item is related both to the trait level for that individual as well as several properties of that item (described below). Additionally, IRT estimates an individual's location on the latent trait by using both the pattern of item responses and the estimated item parameters. In contrast, trait level estimates are obtained by summing response across items in CTT.

The IRT model involves four assumptions. (1) The latent construct of interest must be *unidimensional*. All items evaluated are assumed to measure a single common latent trait. (2) *Monotonicity* specifies that as the probability of endorsing one item increases, so does the probability of endorsing other items. (3) The IRT model must fit the observed data by using the correct number of parameters, i.e., *correct model specification*. (4) *Local item independence* dictates that the probability of response to any item is independent of the response to any other item after taking into account person (θ) and item parameters (difficulty, discrimination) (Embretson & Reise, 2000). A violation of this assumption is called local item dependence which is typically the result of multidimensionality and/or redundant content in items.

The objective of this paper was to evaluate the proposed diagnostic criteria for night eating syndrome using IRT methods. Specifically, a series of analyses were conducted to determine the extent to which these proposed criteria for night eating syndrome meet the key assumptions of IRT analyses and provide useful diagnostic information in identifying individuals with NES.

Method

Participants and Procedures

Participants included 1481 individuals (919 women, 562 men) involved in one of six separate studies (Table 2). Diagnosis of NES was not required for inclusion in the IRT analysis.

Study 1 included 929 overweight individuals with type 2 diabetes from five clinics of the Look AHEAD study, a controlled randomized trial of the long-term health consequences of intentional weight loss (Ryan et al., 2003). During a pre-randomization screening visit participants completed the NEQ as part of a prevalence study of NES and BED in this diabetic population (Allison et al., 2007).

Study 2 included 147 bariatric surgery candidates at the University of Pennsylvania. They completed the NEQ as part of the Weight and Lifestyle Inventory (Wadden & Foster, 2006) included in their preoperative psychological assessment.

Study 3 included 177 outpatient psychiatric patients who participated in a prevalence study of NES at two university clinics (Lundgren et al., 2006). Their treating psychiatrists referred them, and they completed the NEQ in the presence of study staff.

Study 4 included 77 individuals who identified themselves as having NES. They completed the NEQ when they presented for a comprehensive outpatient assessment study of NES (O'Reardon et al., 2004).

Study 5 included 45 overweight participants who served as comparison subjects for the NES sample in Study 4. The NEQ was administered to confirm the absence of NES, in conjunction with interview and diary data.

Study 6 included 106 participants who responded to newspaper advertisements seeking individuals who “eat large amounts of food in the evening” or “get up during the night to eat”. The study was conducted at the Neuropsychiatric Research Institute (Fargo, ND) to collect descriptive data on individuals with broadly defined night eating problems (de Zwaan et al., 2006). Those who met inclusion criteria completed an NEQ during a phone interview.

The six samples were combined to provide a large sample size and maximum heterogeneity across respondents, both of which are advantageous for estimating IRT model parameters (Embretson & Reise, 2000). All participants gave their informed consent for their respective studies and each study was approved by an Institutional Review Board.

Measures

The NEQ was developed as a screening instrument for NES (Allison et al., 2008). The items and responses to the items were revised several times from 2000 to 2006. However, six core items based on a combination of the 1955 (Stunkard et al., 1955) and 1999 NES criteria (Birketvedt et al., 1999) have been present throughout the revisions and include: level of morning hunger, time of first meal, percentage of caloric intake after the evening meal, initial insomnia, frequency of awakenings after going to sleep, and frequency of nocturnal eating episodes. Depressed mood was also included initially, but it did not load on the factor generated from the six core items. Therefore, it was excluded because it violated the unidimensionality assumption of IRT. Other items that assess the psychological features of NES, such as degree of control over evening eating and nocturnal eating (awaking from sleep to eat), were not presented in early versions of the NEQ and were not included in the analysis.

The six core items from the NEQ were converted from five point Likert-type scales (0–4) to dichotomous items (i.e., symptom present vs. symptom absent) to permit IRT analysis of the presence or absence of NES (Table 1). We considered the underlying feature of NES to be a marked circadian delay in food intake (O'Reardon et al., 2004) and that two items may independently, or in combination, define this single construct: evening hyperphagia and/or frequent nocturnal ingestions. Consequently, the six core items from the NEQ represented five criteria to be evaluated in the IRT analysis (Table 1).

Statistical Analyses

Unidimensionality—Most IRT models assume that a single latent trait variable is sufficient to explain the common variance among item responses (Stout, 1987). Although no analytical standard exists for establishing unidimensionality among dichotomous items, nonlinear factor analysis of residual covariance terms is the standard practice (Hattie, 1984; Hattie, 1985;

Hambleton & Rovinelli, 1986). The factor structure of night eating symptom criteria was evaluated with a nonlinear factor analysis based upon tetrachoric correlations using Mplus software (Muthén & Muthén, 1998). The unidimensionality assumption was evaluated in terms of the ratio of first to second factor eigenvalues, root mean square residual values, and residual inter-item correlations.

Monotonicity—The monotonicity of these night eating symptoms was evaluated graphically by plotting an “empirical” item response function for each item, with the x-axis representing the number of criteria met on all other items except the item in question (i.e., *rest-score*), and the y-axis representing the conditional proportion meeting the criterion in question. A monotonic item should demonstrate a higher proportion of endorsement at each successive level of the rest-score (Thissen & Orlando, 2001).

Model specification and IRT analysis—A two-parameter logistic (2PL) IRT model (Model 1) was then estimated for the remaining dichotomized items that included separate difficulty parameters but a common slope for all items. This model was then compared to a 2PL IRT model (Model 2) estimating separate difficulty and slope parameters, and a 3 parameter logistic (3PL) IRT model (Model 3) including separate difficulty, slope and guessing parameters for each item. Model fit was compared between models using the difference in -2 log likelihood values distributed as a chi-square. IRT analysis was performed using PARSCALE software (Muraki & Bock, 1997).

Local item independence—The assumption of local item independence was evaluated by comparing the expected frequencies based on estimated IRT item parameters to the observed frequencies across levels of theta for individual items, pairs of items, and item triplets using the method described by Drasgow, Levine, Tsien, Williams, and Mead (1995). To facilitate comparison of chi-square values based on different degrees of freedom, we adjusted the chi-square values for sample size by dividing by the degrees of freedom (Drasgow et al., 1995). Values less than 3.0 indicate good fit, while large ratio statistics, particularly for pairs and triplets, may indicate violations of local independence and/or unidimensionality.

Figure 1 presents item characteristic curves (ICC's) and item information curves (IIC's) for four sample items which illustrate the three common IRT item parameters. ICC's display the probability of endorsing a particular item as a function of the trait level, while IIC's display the amount of psychometric information provided about the latent trait as a function of the trait level. The probability of response to any item is assumed to be a function of three item parameters. (a) The *item threshold* parameter represents the point along the latent trait at which 50% of the respondents endorse the item, and therefore is related to item difficulty. In Figure 1, Items 1 and 3 are of moderate (and equal) difficulty; Item 4 is the least difficult item, while Item 2 is the most difficult. (b) The *item discrimination* parameter indicates the “steepness” of the slope and represents the degree to which an item differentiates participants of different ability. Although Items 1, 2, and 4 differ markedly in terms of threshold, they have identical discrimination parameters (i.e. the slope of the lines is the same), indicating that they are comparable in terms of their ability to discriminate across the latent construct. In contrast, Item 3 has a markedly lower discrimination parameter. The effects of this lower discrimination parameter can be seen in the IIC's in Figure 1, where the curve for Item 3 is much lower than the remaining items. Thus, Item 3 provides less information about the latent construct. (c) Finally, the *lower asymptote* parameter indicates the probability of item endorsement at the lowest levels of the latent trait, and is typically used to represent guessing or base rates of occurrence. The value of the lower asymptote for Item 4 is .50, indicating that 50% of those at the lowest level of the latent construct will endorse this symptom.

Results

Sample Characteristics

The majority of participants in all studies were female (Table 2). Age and BMI varied considerably across trials, with the oldest participants in the Look AHEAD trial (Study 1) and the heaviest participants among bariatric surgery candidates (Study 2). Race and ethnicity of the participants across all studies included approximately 70.4% Caucasian, 23.5% African American, 3.4% Hispanic/Latino, 0.6% Asian American, 0.6% Native American, 1.5% other/unknown.

Unidimensionality

Factor analysis of the dichotomous items revealed a first main factor accounting for 56.3% of the variance, with a first-to-second factor eigenvalue ratio of 3.43:1. Root mean square residual for this model was 0.051, suggesting that the five items entered into the factor analysis are represented adequately by a single unidimensional construct.

Monotonicity

Empirical item response functions for each of the five items were calculated to show the proportion of the sample endorsing each item as a function of the number of the remaining four items that were endorsed (i.e., the *rest-score*). All five items showed clear evidence of monotonicity with increasing item response functions (Figure 2).

Symptom Frequency

Table 3 presents the percent of participants endorsing each of the five items by study. As would be expected, rates of endorsement for all items were lower for the obesity and psychiatric clinic studies (Studies 1–3) and control study (Study 5) than for Studies 4 and 6 that focused specifically on night eating problems.

Model Specification and IRT Analysis

PARSCALE software was used to evaluate three separate 5-item models. The -2 log likelihood value for Model 1 (2PL common slope model) was 8306.87, compared to 8273.04 for Model 2 (2PL fully unconstrained model) and 8273.01 for Model 3 (3PL fully unconstrained model). Thus, Model 2 provided a significantly better fit than Model 1 ($\chi^2 = 33.83$, $df = 5$, $p < .001$). However, the fit for Models 2 and 3 was not significantly different ($\chi^2 = 0.03$, $df = 6$, $p > .05$), suggesting that Model 2, the two parameter fully unconstrained model, provided the most parsimonious fit.

Table 1 shows the parameter estimates and standard errors for this model. Figure 3 shows the corresponding ICC's and Figure 4 the IIC's based upon these parameter estimates. Of note, the difficulty parameters (from Figure 3) for three of the items (evening hyperphagia/nocturnal eating, initial insomnia, and night awakenings) are relatively comparable, ranging from 0.53 to 0.82. This suggests that they provide diagnostic information across a comparable range of the night eating severity (Figure 4). Second, the symptom which provides the most amount of diagnostic information (Figure 4) is evening hyperphagia/nocturnal eating, which are the core features of NES. Finally, two of the items, morning anorexia and delayed morning eating, have considerably lower item response curve slopes (Figure 3) than the remaining items, and consequently, provide considerably less psychometric information (Figure 4).

Local Item Independence

The lower item response curves for morning anorexia and delayed morning eating led to the evaluation of their local item independence. Analysis of the 2PL fully unconstrained 5-item

model characterized across the six samples in Table 3 produced chi-square ratios greater than 3.0 for 3 of 10 item pairs and 5 of 10 item triplets. All of the pair ratios and all but 2 of the triplet ratios above 3.0 included morning anorexia and/or delayed morning eating. The chi-square ratio for the morning anorexia/delayed morning eating pair was 34.4, and all ratios for triplets containing both items were 17.9 or higher. This pattern of results suggests that one or both of these two items violated the assumption of local independence. Consequently, these items were dropped and a new 3-item (evening hyperphagia/nocturnal eating, initial insomnia, and night awakenings) 2PL fully unconstrained IRT model was reevaluated for local item independence. All of the chi-square ratios for single items, item pairs, and triplets were less than 3.0. These results suggest that the 3-item fully unconstrained model is consistent with the IRT assumption of local item independence.

Studies 4 and 6 had the largest number of participants who met all three criteria, at 49% and 39% of the samples, respectively, followed by Studies 3 (22%), 2 (7.5%), 1 (4.5%), and 5 (2%). An additional 47% and 43% of Studies 4 and 6, respectively, met two of the criteria, followed by Studies 2 (26%), 3 (22%), 1 (11%), and 5 (2%).

Discussion

Implications for Night Eating Syndrome

IRT analyses revealed three core features: evening hyperphagia and/or nocturnal eating, initial insomnia, and awakenings from sleep. This analysis combined the two central features of NES, evening hyperphagia and nocturnal ingestions, because consistent occurrence of either of these behaviors represents a delay in the circadian pattern of eating. Further clarification of the parameters for the concept of evening hyperphagia will aid in translating this statistical analysis into clinically meaningful diagnostic criteria. Since evening meal time varies across both individuals and cultures, it seems reasonable to anchor its measurement to ingestions occurring after the evening meal rather than to a specific time. A controlled study revealed a post-evening meal caloric intake of 34.6% (SD = 10.1%) in an NES sample (n = 46) compared to 10.0% (SD = 6.9%) in the control group (n = 43) when based on a diary record of food intake (O'Reardon et al., 2004). This suggests that a standard of 25% (more than two SDs greater than the mean in control participants and one SD less than that of night eaters) may be reasonable when based on the more stringent diary observation, and 50% may be appropriate when retrospective self-report is used.

The IRT analysis revealed initial insomnia as a core feature. This is a surprising finding given that only 57% of the Study 4 NES sample and 66% of the Study 6 NES sample endorsed initial insomnia, the lowest endorsement rate of the three core features. Objectively, both outpatient actigraphy and diary reports (O'Reardon et al., 2004) and inpatient polysomnography (Rogers et al., 2006) showed that the time of sleep onset of night eaters did not differ from that of comparison subjects. Sleep latency (the time spent in bed before falling asleep) is more difficult to measure reliably through actigraphy and polysomnography than sleep onset. In a polysomnography report (Rogers et al., 2006) sleep latency did not differ statistically between the two groups (25.6 ± 29.3 minutes for NES group versus 16.2 ± 22.8 minutes for control group). Thus, subjective self-reports of initial insomnia appear to be at odds with objective sleep data, as has been noted in the sleep literature (Kloss, 2003). Clinically, two useful ways of assessing initial insomnia in NES include reports of difficulty falling asleep or using food late in the evening to induce sleep.

Additionally, this analysis suggests that initial insomnia and night awakenings must both be present for a diagnosis of NES. Less than half of the participants in Studies 4 and 6 identified as having NES would meet all three criteria based on their initial survey responses. This has implications for both the future research and nosology of this syndrome. It raises the question

of whether the severity of some symptoms, namely the hallmark symptoms of evening hyperphagia and nocturnal ingestions, could outweigh the absence of others in determining a diagnosis. These findings also raise the question of whether initial insomnia and awakenings could be considered modifiers of NES, but not mandatory criteria for diagnosis of NES. This idea would be similar to the DSM diagnosis of major depression, where depressed mood and/or loss of interest or pleasure in activities must be endorsed, and only a total of five of nine criteria must be met.

Morning anorexia did not contribute additionally to the model. This salient aspect of night eaters attracted early attention (Stunkard et al., 1955), but subsequent experience has revealed that its frequency, particularly among obese persons, means that it cannot discriminate night eaters.

When evaluating a disorder, it is useful to know what the distribution is of people meeting all or some of the criteria. The majority of participants in non-NES studies, such as the Look AHEAD and Control Group studies, met none or one of the criteria, while a minority met all three. Conversely, the majority of participants in the night eating studies met two or three of the criteria. Thus, IRT seems useful in providing evidence about the relative strength of various items, but ultimately, the clinical interpretation of the disorder may include a broader spectrum of symptoms and behaviors.

Limitations of this study include the examination of only six items of the NEQ. Other items that assess psychological aspects of NES, including cravings and control over eating in the evening and at night, were not present on earlier versions of the NEQ, and, therefore, did not have large enough sample sizes to include in analyses. Future work should examine whether these psychological aspects of NES fall on the same dimension as the behavioral aspects. Additionally, the degree of distress and/or impairment in functioning due to NES, essential aspects of psychiatric diagnoses, were not assessed by the NEQ but should be added in the future. The self-report nature of the data is also limiting. There were no attempts to confirm participant responses on the majority of NEQs. Finally, responses to items on the NEQ were given on a Likert-type scale and later dichotomized based upon cut points derived from clinical experience with NES patients. Interview and behavioral data should be used to confirm such results in future studies.

The current study, in conjunction with a limited literature (Langenbucher et al., 2004), provides evidence that the assumptions of IRT analyses are consistent with evaluation of criteria for NES and other psychiatric disorders. We identified three criteria that offer considerable information about NES (evening hyperphagia and/or nocturnal eating; initial insomnia, and night awakenings) and two criteria that contain less information (morning anorexia and time of first meal). With these core features of NES identified, future work should focus on describing the frequency and severity of these indicators required to comprise a “psychiatric disorder”, along with their relationship to distress and impairment in functioning that we were not able to assess here. These findings should aid in guiding future research on NES by clarifying the items that best identify those who have difficulties with night eating.

Acknowledgements

Studies described in the report are supported by the National Institutes of Health (RO1 DK 056735, RO1 DK60432, and K23 DK60023). The authors declare no conflicts of interest.

We would like to acknowledge the work of the Eating Disorders Look AHEAD Study Group: Kelly C. Allison, Ph.D., Albert J. Stunkard, M.D., and Thomas Wadden, Ph.D. for University of Pennsylvania; Scott Crow, M.D., Joy Johnson-Lind, M.S.W., and Robert Jeffries, Ph.D. for University of Minnesota; John Foreyt, Ph.D., Rebecca Reeves, DrPH, and Brian Hunter, M.A. for Baylor College of Medicine; Delia Smith-West, Ph.D., Vicki DiLillo, Ph.D., and Stacy

Gore, Ph.D. for University of Alabama-Birmingham; and Brent Van Dorsten, Ph.D., James Hill, Ph.D., and Terra Worley, for University of Colorado.

We would also like to thank Nicole Martino, Heidi Marshall, M.S.S., and Heidi Toth for their assistance on Studies 1 and 3–5 and to Lauren Gibbons for her work on Study 2, all from the University of Pennsylvania School of Medicine's Center for Weight and Eating Disorders. Finally, we would like to thank David Dinges, Ph.D. of University of Pennsylvania School of Medicine's Division of Sleep and Chronobiology for consultation on the circadian aspects of eating and sleeping.

References

- Adami GF, Meneghelli A, Scopinaro N. Night eating and binge eating disorder in obese patients. *International Journal of Eating Disorders* 1999;25:335–338. [PubMed: 10191999]
- Allison KC, Crow SJ, Reeves RR, West DS, Foreyt JP, DiLillo VG, Wadden TA, Jeffery RW, Van Dorsten B, Stunkard AJ. the Eating Disorders Subgroup of the Look AHEAD Research Group. The prevalence of binge eating disorder and night eating syndrome in adults with type 2 diabetes mellitus. *Obesity* 2007;15:1285–1291.
- Allison KC, Grilo CM, Masheb RM, Stunkard AJ. Binge eating disorder and night eating syndrome: A comparative study of disordered eating. *Journal of Consulting and Clinical Psychology* 2005;73:1107–1115. [PubMed: 16392984]
- Allison KC, Lundgren JD, O'Reardon JP, Martino NS, Sarwer DB, Wadden TA, Crosby RD, Engel SG, Stunkard AJ. The Night Eating Questionnaire (NEQ): psychometric properties of a measure of severity of the night eating syndrome. *Eating Behaviors* 2008;9:62–72. [PubMed: 18167324]
- Allison KC, Wadden TA, Sarwer DB, Fabricatore AN, Crerand C, Gibbons L, Stack R, Stunkard AJ, Williams N. Night eating syndrome and binge eating disorder among persons seeking bariatric surgery: prevalence and related features. *Obesity* 2006;14(Suppl 2):77S–82S. [PubMed: 16648598]
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*. Washington, DC: American Psychiatric Association; 2000.
- Birketvedt GS, Florholmen J, Sundsfjord J, Osterud G, Dinges D, Bilker W, Stunkard A. Behavioral and neuroendocrine characteristics of the night-eating syndrome. *Journal of the American Medical Association* 1999;282:657–663. [PubMed: 10517719]
- Ceru-Björk C, Andersson I, Rössner S. Night eating and nocturnal eating – two different or similar syndromes among obese patients? *International Journal of Obesity* 2001;25:365–372. [PubMed: 11319634]
- de Zwaan M, Burgard MA, Schenck CH, Mitchell JE. Night time eating: A review of the literature. *European Eating Disorders Review* 2003;11:7–24.
- de Zwaan M, Roerig DB, Crosby RD, Karaz S, Mitchell JE. Nighttime eating: a descriptive study. *International Journal of Eating Disorders* 2006;39:224–232. [PubMed: 16511835]
- Drasgow F, Levine MV, Tsien S, Williams BA, Mead AD. Fitting polytomous item response theory models to multiple-choice tests. *Applied Psychological Measurement* 1995;19:143–165.
- Embretson, SE.; Reise, SP. *Item response theory for psychologists*. Mahwah, New Jersey: Lawrence Erlbaum; 2000.
- Gluck ME, Geliebter A, Satov T. Night eating syndrome is associated with depression, low self-esteem, reduced daytime hunger, and less weight loss in obese outpatients. *Obesity Research* 2001;9:264–267. [PubMed: 11331430]
- Grilo CM, Masheb RM. Night-time eating in men and women with binge eating disorder. *Behaviour Research and Therapy* 2004;42:397–407. [PubMed: 14998734]
- Hambleton RK, Rovinelli RJ. Assessing the dimensionality of a set of test items. *Applied Psychological Measurement* 1986;10:287–302.
- Hattie J. An empirical study of various indices for determining unidimensionality. *Multivariate Behavioral Research* 1984;19:49–78.
- Hattie J. Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement* 1985;9:139–164.
- Hsu LKG, Bentancourt S, Sullivan SP. Eating disturbances before and after vertical banded gastroplasty: a pilot study. *International Journal of Eating Disorders* 1996;19:23–34. [PubMed: 8640199]

- Kloss, J. Daytime sequelae of insomnia. In: Szuba, MP.; Kloss, JD.; Dinges, DF., editors. *Insomnia: principals and management*. Cambridge: Cambridge University Press; 2003. p. 23-42.
- Kuldau JM, Rand CSW. The night eating syndrome and bulimia nervosa in the morbidly obese. *International Journal of Eating Disorders* 1986;5:143–148.
- Lamerz A, Kuepper-Nybelen J, Bruning N, Wehle C, Trost-Brinkhues G, Brenner H, Hebebrand J, Herpertz-Dahlmann B. Prevalence of obesity, binge eating and night eating in a cross-sectional field survey of 6-year-old children and their parents in a German urban population. *Journal of Child Psychology & Psychiatry* 2005;46:385–393. [PubMed: 15819647]
- Langenbucher JW, Labouvie E, Martin CS, Sanjuan PM, Bavly L, Kirisci L, Chung T. An application of item response theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. *Journal of Abnormal Psychology* 2004;113:72–80. [PubMed: 14992659]
- Lord, F. *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1980.
- Lundgren JD, Allison KC, Crow S, O'Reardon JP, Berg KC, Galbraith J, Martino NS, Stunkard AJ. Prevalence of the night eating syndrome in a psychiatric population. *American Journal of Psychiatry* 2006;163:156–158. [PubMed: 16390906]
- Marshall HM, Allison KC, O'Reardon JP, Birketvedt G, Stunkard AJ. The night eating syndrome among nonobese persons. *International Journal of Eating Disorders* 2004;35:217–222. [PubMed: 14994360]
- Muthén, B.; Muthén, L. *Mplus*. Muthén & Muthén: Los Angeles, CA; 1998.
- Muraki, E.; Bock, RD. *PARSCALE: IRT item analysis and test scoring for rating scale data*. Chicago, Illinois: Scientific Software International; 1997.
- Napolitano MA, Head S, Babyak MA, Blumenthal JA. Binge eating disorder and night eating syndrome: Psychological and behavioral characteristics. *International Journal of Eating Disorders* 2001;30:193–203. [PubMed: 11449453]
- O'Reardon JP, Ringel BL, Dinges DF, Allison KC, Rogers NL, Martino NS, Stunkard AJ. Circadian eating and sleeping patterns in the night eating syndrome. *Obesity Research* 2004;12:1789–1796. [PubMed: 15601974]
- Pawlow LA, O'Neil PM, Malcolm RJ. Night eating syndrome: effects of brief relaxation training on stress, mood, hunger, and eating patterns. *International Journal of Obesity* 2003;27:970–978. [PubMed: 12861239]
- Rand CS, MacGreggor AMC, Stunkard AJ. The night eating syndrome in the general population and among postoperative obesity surgery patients. *International Journal of Eating Disorders* 1997;22:65–69. [PubMed: 9140737]
- Rogers NL, Dinges DF, Allison KC, Maislin G, Martino N, O'Reardon JP, Stunkard AJ. Assessment of sleep in women with night eating syndrome. *Sleep* 2006;29:814–819. [PubMed: 16796220]
- Ryan DH, Espeland MA, Foster GD, Haffner SM, Hubbard VS, Johnson KC, Kahn SE, Knowler WC, Yanovski SZ. Look AHEAD Research Group. Look AHEAD (Action for Health in Diabetes): design and methods for a clinical trial of weight loss for the prevention of cardiovascular disease in type 2 diabetes. *Controlled Clinical Trials* 2003;24:610–628. [PubMed: 14500058]
- Stout W. A new item response theory modeling approach with application to unidimensional assessment and ability estimation. *Psychometrika* 1987;55:293–326.
- Striegel-Moore RH, Franko DL, Thompson D, Affenito S, Kraemer HC. Night Eating: Prevalence and demographic correlates. *Obesity Research* 2006;14:139–147.
- Striegel-Moore RH, Thompson D, Franko DL, Barton B, Affenito S, Schreiber GB, Daniels SR. Definitions of night eating in adolescent girls. *Obesity Research* 2004;12:1311–1321. [PubMed: 15340114]
- Stunkard AJ, Berkowitz RI, Wadden TA, Tanrikut C, Reiss E, Young L. Binge eating disorder and the night eating syndrome. *International Journal of Obesity* 1996;2:1–6. [PubMed: 8788315]
- Stunkard AJ, Grace WJ, Wolff HG. The night-eating syndrome: A pattern of food intake among certain obese patients. *American Journal of Medicine* 1955;19:78–86. [PubMed: 14388031]
- Thissen, D.; Orlando, M. Item response theory for items scored in two categories. In: Thissen, D.; Wainer, H., editors. *Test Scoring*. Mahwah, New Jersey: Lawrence Erlbaum; 2001. p. 73-140.
- Vander Wal JS, Waller SM, Klurfeld DM, McBurney MI, Dhurandhar NV. Night eating syndrome: evaluation of two screening instruments. *Eating Behaviors* 2005;6:63–73. [PubMed: 15567112]

- Wadden TA, Foster GD. Weight and Lifestyle Inventory. *Obesity* 2006;14(Suppl 2):99S–118S. [PubMed: 16648601]
- Weiss, DJ.; Yoes, ME. Item response theory. In: Habmleto, RK.; Zaal, J., editors. *Advances in educational and psychological testing: theory and applications*. Norwell, MA: Kluwer Academic Publishers; 1991. p. 94-109.
- World Health Organization International. *Statistical Classification of Diseases and Related Health Problems*, 10th revision. 2. Los Angeles, CA: Practice Management Information Corporation; 2007.

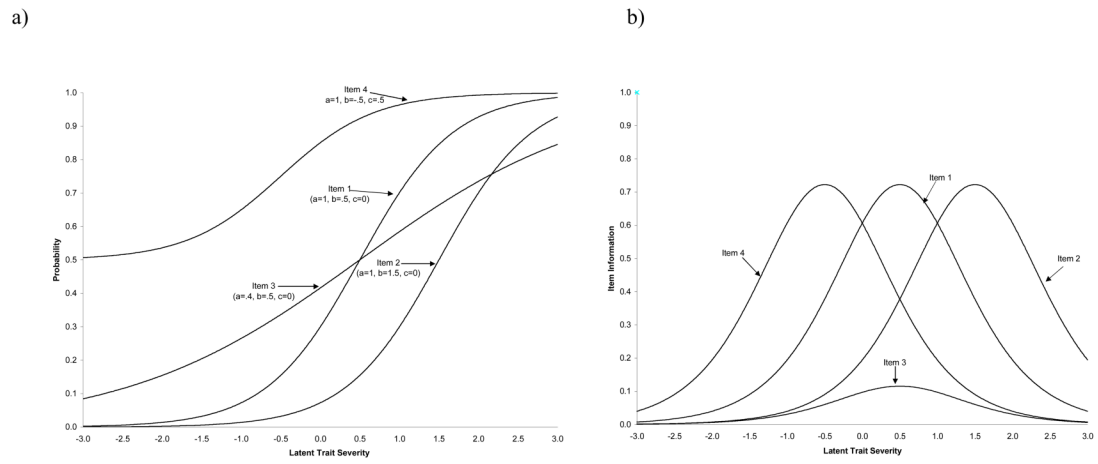


Figure 1.
a) Item characteristic curves (ICC) and (b) item information curves (IIC) for sample test items.

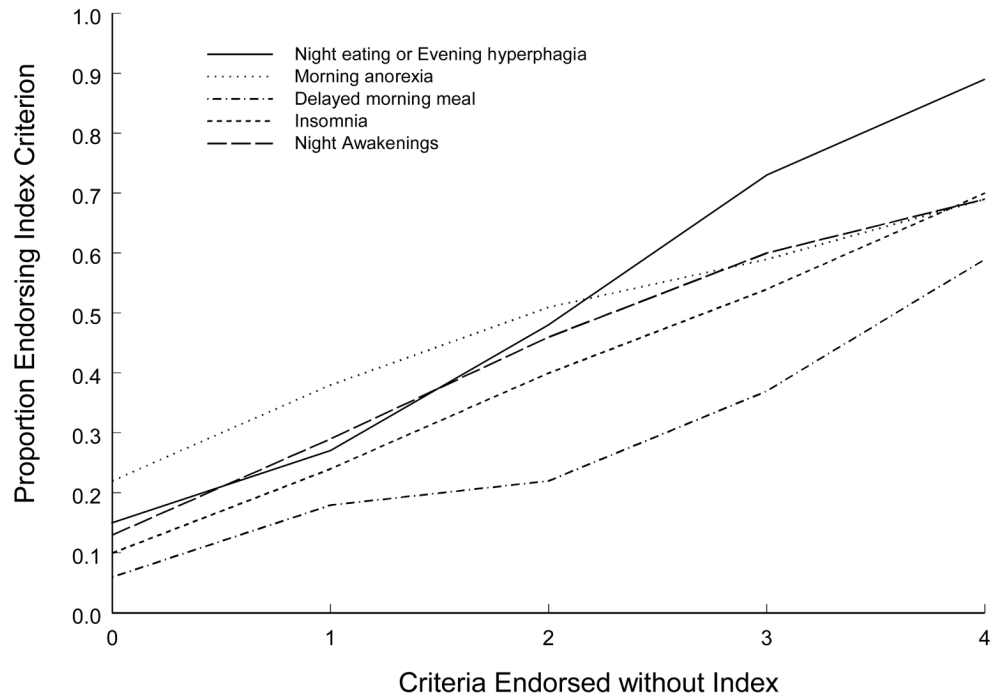


Figure 2.
Empirical item response functions.

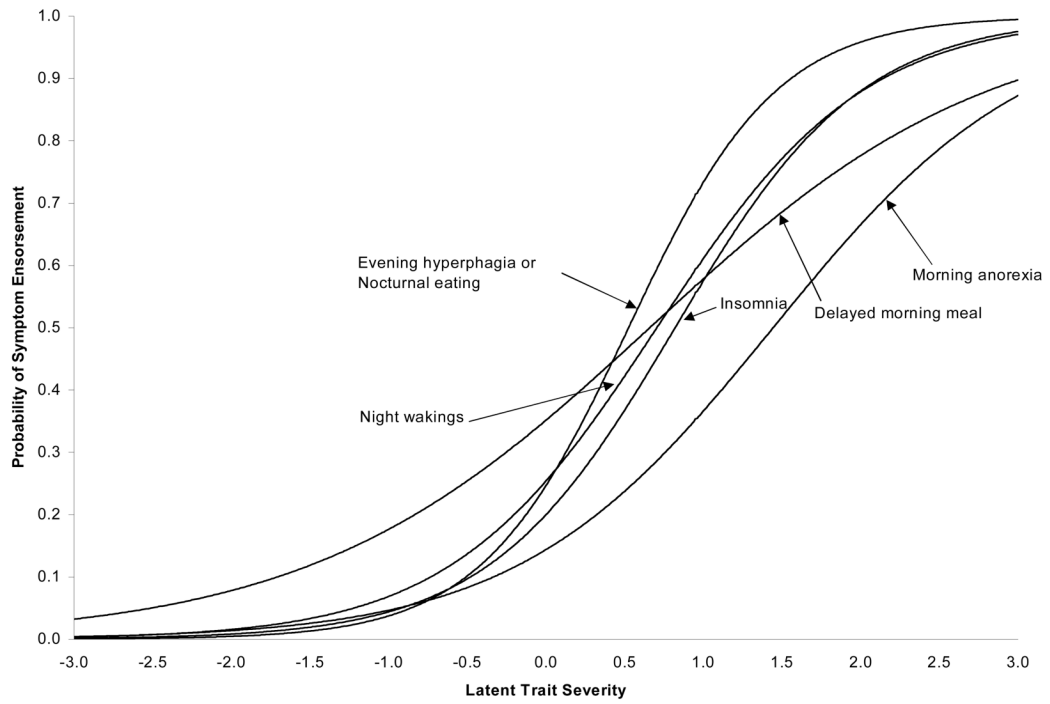


Figure 3.
Item response curves for the 2PL logistic model.

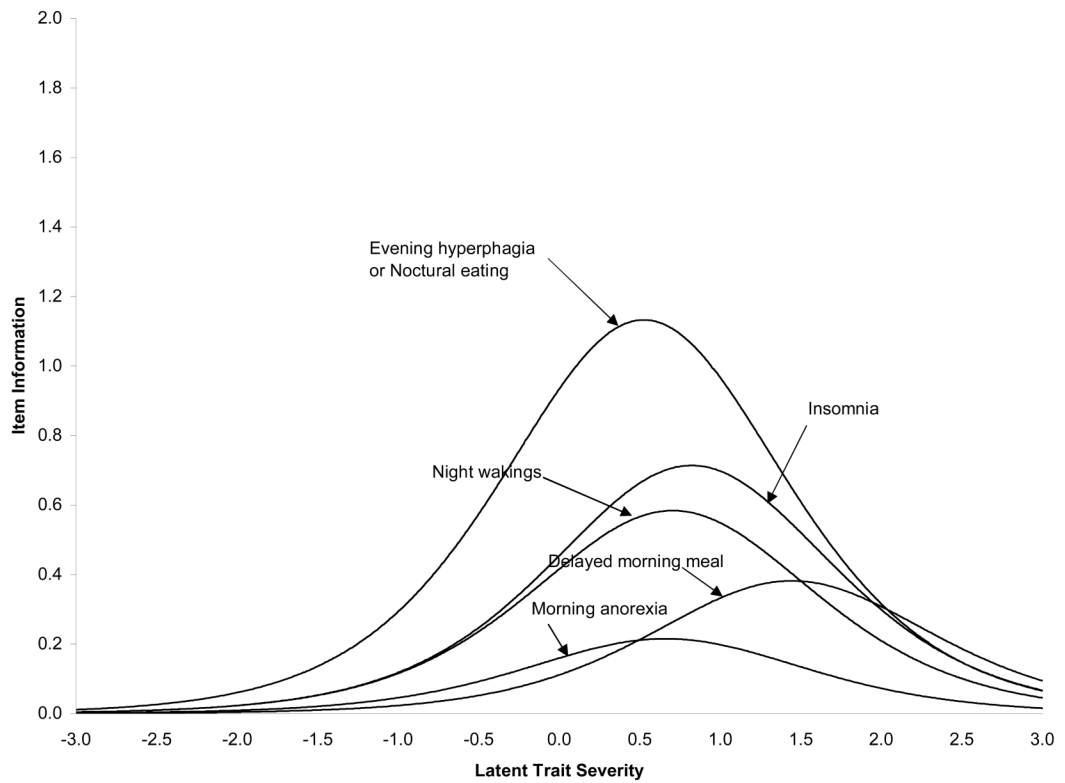


Figure 4.
Item information curves for the 2PL logistic model

Table 1

Cut-off descriptors for Night Eating Questionnaire items retained in the IRT analysis and their parameter estimates (standard errors) for the two parameter logistic fully unconstrained IRT model.

Symptom	Symptom is "present"	Discrimination	Difficulty
Morning anorexia	≤ "Somewhat hungry"	.55 (.04)	.66 (.07)
Delayed morning meal	12:01 PM or later	.73 (.05)	1.44 (.09)
Evening hyperphagia or Nocturnal eating	≥ 26% of calories after supper and/or nocturnal eating "about half the time"	1.25 (.07)	.53 (.04)
Initial insomnia	"About half the time"	.99 (.06)	.82 (.05)
Night awakenings	"More than once a week"	.90 (.05)	.71 (.05)

Table 2

Sample characteristics by study.

Variable	Study 1 Look AHEAD (n=929)	Study 2 Pre- Gastric Bypass (n=147)	Study 3 Psychiatric Clinic (n=177)	Study 4 Stunkard NES Visit (n=77)	Study 5 Stunkard Control (n=45)	Study 6 Phone Interview (n=106)	Combined Sample (n=1479)
Females [N (%)]	546 (58.8%)	103 (70.1%)	124 (70.4%)	48 (62.0%)	30 (66.7%)	68 (64.2%)	919 (62.1%)
Age (yrs) [mean (SD)]	60.1 (6.5)	43.6 (11.5)	40.2 (12.4)	43.4 (10.7)	39.4 (11.3)	49.0 (13.1)	53.8 (12.3)
BMI (kg.m ⁻²) [mean (SD)]	36.2 (6.4)	50.5 (9.4)	29.2 (8.2)	35.4 (7.5)	36.3 (6.1)	30.8 (7.2)	36.3 (8.8)

Table 3

Frequency and percent of participants endorsing symptoms by study.

NES Symptom [N (%)]	Study 1 Look AHEAD (n=929)	Study 2 Pre- Gastric Bypass (n=147)	Study 3 Psychiatric Clinic (n=177)	Study 4 Stunkard NES Visit (n=77)	Study 5 Stunkard Control (n=45)	Study 6 Phone Interview (n=106)	Combined Sample (n=1479)
Morning anorexia	218 (23.5%)	87 (59.2%)	104 (58.8%)	55 (71.4%)	9 (20%)	80 (75.5%)	553 (37.3%)
Delayed morning Meal	119 (12.8%)	25 (17.0%)	55 (31.1%)	40 (51.9%)	8 (17.8%)	46 (43.4%)	293 (19.8%)
Evening hyperphagia and/or nocturnal eating	222 (23.9%)	42 (28.6%)	68 (38.4%)	76 (98.7%)	3 (6.7%)	96 (90.6%)	507 (34.3%)
Initial insomnia	174 (18.7%)	50 (34.0%)	78 (44.1%)	44 (57.2%)	2 (4.4%)	70 (66.0%)	418 (28.2%)
Night awakenings	188 (20.2%)	62 (42.2%)	83 (46.9%)	68 (88.3%)	4 (8.9%)	67 (63.2%)	384 (25.9%)
Total Symptoms [Mean (SD)]	1.0 (1.2)	2.2 (1.3)	2.4 (1.7)	4.5 (1.2)	0.6 (0.8)	4.3 (1.5)	1.6 (1.6)