Readability and comprehension of self-report binge eating measures

Lauren K. Richards [a,*], R. Kathryn McHugh [b,c], Elizabeth M. Pratt [a], Heather Thompson-Brenner [a]

[a] Department of Psychology, Boston University, 648 Beacon Street, Boston, MA 02215, USA
[b] Division of Alcohol and Drug Abuse, McLean Hospital, 115 Mill Street, Belmont, MA 02478, USA
[c] Department of Psychiatry, Harvard Medical School, 401 Park Drive, Boston, MA 02215, USA

ARTICLE INFO

ABSTRACT

The validity of self-report binge eating instruments among individuals with limited literacy is uncertain. This study aims to evaluate reading grade level and multiple domains of comprehension of 13 commonly used self-report assessments of binge eating for use in low-literacy populations. We evaluated self-report binge eating measures with respect to reading grade levels, measure length, formatting and linguistic problems.
Results: All measures were written at a reading grade level higher than is recommended for patient materials (above the 5th to 6th grade level), and contained several challenging elements related to comprehension. Correlational analyses suggested that readability and comprehension elements were distinct contributors to measure difficulty. Individuals with binge eating who have low levels of educational attainment or limited literacy are often underrepresented in measure validation studies. Validity of measures and accurate assessment of symptoms depend on an individual's ability to read and comprehend instructions and items, and these may be compromised in populations with lower levels of literacy.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Limited literacy is highly prevalent in the US (Kutner, Greenber, & Baer, 2005) and is particularly elevated in psychiatric populations (Christensen & Grace, 1999; Currier, Sitzman, & Trenton, 2001). Consideration of the reading grade level of healthcare materials is crucial to the ability of patients to comprehend this information. However, studies of self-report symptom measures for depression (Berndt, Schwartz, & Kaiser, 1983; McHugh & Behar, 2009; Shumway, Sentell, Unick, & Bamberg, 2004), anxiety (McHugh & Behar, 2009; McHugh, Rasmussen, & Otto, 2011), and other psychiatric symptoms (Andrasik, Heimber, Edlund, & Blankenberg, 1981; Beckman & Lueger, 1997) suggest that many of these measures are difficult to comprehend with reading grade levels for most exceeding recommendations (e.g., 5th–6th grade) (Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, AMA, 1999).

Binge eating disorder (BED) is characterized by recurrent out-of-control eating episodes (APA, 2000). A large proportion of individuals with BED have less than a high school education (e.g., 14–23%) (Pike, Dohm, Striegel-Moore, Wilfley, & Fairburn, 2001; Striegel-Moore et al., 2000, 2005). These rates are particularly elevated in ethnic and racial minority groups (Franko et al., 2012), with one study estimating that more than 70% of Hispanic and more than 20% of African American women

with BED had less than high school education (Marcus, Bromberger, Wei, Brown, & Kravitz, 2007). These data may indicate particularly lower reading ability, given that actual reading grade level is often lower than reported education level (Davis et al., 1993; Manly, Jacobs, Touradji, Small, & Stern, 2002).

Moreover, given the under-representation of racial and ethnic minorities in clinical trials for BED (Franko et al., 2012), the effectiveness of psychosocial treatment for these individuals is unknown. To meet the important public health goal of enhancing access and quality of treatment for this population, measures that accurately assess symptoms are needed. Valid self-report assessment requires the ability of patients to comprehend measures; however, no published studies have examined the body of published binge eating symptom measures to characterize their applicability to individuals of differing education and literacy levels. A prior study of measures of bulimia nervosa and anorexia nervosa suggests that many measures approximated recommended grade levels, with 68% of measures falling in the 5th–7th grade range (Petty, Rosen, & Michaels, 2000). However, this study did not provide a comprehensive review of binge eating measures, and it only examined reading grade level, despite the importance of other factors to the comprehension of written materials.

The aim of the current study was to examine the readability of self-report measures of binge eating symptoms. We evaluated reading grade level, and several other factors that may influence comprehension including measure length, format and linguistic problems. We hypothesized that the reading grade level of these measures, on average, would exceed American Medical Association (AMA) recommendations for patient materials (5th–6th grade).

* Corresponding author at: 648 Beacon Street, 6th Floor, Boston MA, 02215, USA.
Tel.: +1 617 353 9610; fax: +1 617 353 9609.
E-mail address: laurenkr@bu.edu (L.K. Richards).

2. Methods

2.1. Selection of measures

A thorough search for self-report measures was conducted to identify all published measures related to the assessment of binge eating. We selected measures that (1) assessed binge eating symptoms, (2) were written in English, and (3) had been validated in at least 2 samples. First, we included self-report measures from the following sources: *The Handbook of Assessment and Treatment Planning for Psychological Disorders* (Craighead & Smith, 2011); *The Assessment of Addictive Behaviors* (Collins & Ricciardelli, 2008); and *The Assessment of Eating Disorders* (Mitchell & Peterson, 2005). Second, we performed a literature search of the *PsycINFO* and *PubMed* databases, using the search words *self-report, self-report measure*, and *self-report instrument*, combined with *binge eating, BED*, and *bulimia* to identify measures not included in the chapters. Finally, we included a measure developed by the authors (Richards, Pratt & Thompson-Brenner, unpublished manuscript) designed to assess the frequency and associated distress of loss of control eating (i.e., one of the central features of binge eating).

These methods resulted in the identification of 22 measures, 9 of which were later excluded because they did not meet the above inclusion criteria or were not obtainable. Of these 13 measures, two did not include standard instructions and thus were excluded from analyses involving instructions.

2.2. Readability assessment

Reading grade level was calculated separately for measure instructions and items using three readability formulas: Flesch Reading Ease (Flesch, 1948); SMOG (McLaughlin, 1969); and FORCAST (Kern, Sticht, Welty, & Hauke, 1976). These indices were averaged to provide a composite reading grade level for each scale's instructions and items, separately and in combination. We excluded response scales from our analyses because the majority were Likert type scales or repeated response choices, and two of the readability formulas require full sentences in order to calculate a grade level.

2.3. Comprehension assessment

Consistent with previous work (McHugh et al., 2011), our comprehension analyses included the following three domains: a) measure length (word count); b) format score, based on a dichotomous rating of the presence or absence of 4 formatting elements that may negatively influence comprehension (reverse scoring, presence of instructions to skip items based on a particular response, shifting of response sets, and the use of double negatives or negative qualifiers prior to a verb); and c) linguistic problems identified using the Question Understanding Aid (QUAID), a validated computer-based program that identifies 5 core linguistic problems related to wording, syntax and semantics (Graesser, Wiemer-Hastings, Kreuz, Wiemer-Hastings, & Marquis, 2000). For QUAID scores, we divided the total number of challenging linguistic elements by the word count to adjust for measure length. For this adjusted index, higher scores indicate a greater number of challenging elements.

2.4. Composite comprehension assessment

Correlations between readability and comprehension indices were calculated to evaluate their degree of overlap. We also created a composite comprehension score by creating and summing standardized values (z scores) for each comprehension domain (word length, format, and QUAID score) in order to rank measures by comprehension difficulty, and to facilitate comparison with rankings by reading grade level. Finally, we combined the two validated measures (readability and QUAID) to provide an overall index of comprehension that was used to rank measures based on level of difficulty.

3. Results

A summary of reading grade levels and comprehension results is presented in Table 1. The mean reading grade level of instructions was 10.7 (SD = 2.3, range = 7.4–14.7). The mean reading grade level for the items was 8.2 (SD = 0.7, range = 7.2–9.4). The mean length of the measures was 576 words (SD = 304.7, range = 174–1218). All the measures examined had at least one challenging formatting element, with the majority of the measures (77%) having two or more challenging elements. Of the eleven scales for which instructions were included, linguistic problems were common in all but two measures; the QUAID tool identified at least one linguistic problem in 80% of measure instructions. For measure items, the QUAID tool identified several linguistically challenging elements in 100% of the measures.

The correlations between composite comprehension scores and the mean readability for items and instructions were nonsignificant (r = −.06, ns), suggesting that these indices measure independent factors. When we compared measure rankings based on comprehension indices versus average reading grade level, a very different pattern emerged. For example, the top two easiest measures relative to readability were the two most difficult measures relative to comprehension. We computed a total composite score by summing the standardized QUAID scores and standardized readability scores (because they are both validated indices of text comprehension), in order to evaluate the relative difficulty of each measure. Measure rankings by readability score, comprehension score and total composite score are displayed in Table 2.

4. Discussion

These results confirm our hypothesis that, on average, binge eating measures are written at a reading grade level higher than the 5th–6th grade reading level recommended by the AMA for patient materials. Our comprehension analyses similarly revealed many challenging elements that may compromise patient comprehension. All of the measures contained at least one difficult formatting component and linguistic problems were common.

Our finding that the additional evaluation of comprehension yielded different results than readability alone is consistent with previous multidimensional assessments of measure complexity (McHugh et al., 2011; Shumway et al., 2004), and suggests that each index measures distinct aspects of measure difficulty. In general, our data support the idea that indices of comprehension should be considered in addition to readability when developing and evaluating self-report measures. Further research is needed to confirm the relative contributions of these variables to measure validity.

To contextualize these findings, we examined the education levels of the samples used in validation studies of the measures included in our analysis. Of the 42 validation studies identified for which education levels were reported, half (n = 21) were conducted exclusively among undergraduate or post-graduate students. The remaining studies included samples recruited from treatment studies, specialty eating disorder or weight loss clinics, community and medical settings. Across these settings, the proportion of subjects with less than a high school level of education ranged from 6%–15.5% (Barnes, Masheb, White, & Grilo, 2011; Celio, Wilfley, Crow, Mitchell, & Walsh, 2004; Dymek-Valentine, Rienecke-Hoste, & Alverdy, 2004; Grilo, Masheb, & Wilson, 2001; Hrabosky et al., 2008; Mond, Hay, Rodgers, Owen, & Beumont, 2004; Mond et al., 2008; Peterson et al., 2007). Notably, two measures (the Eating Disorders Diagnostic Scale and the Eating Attitudes Test) were validated in middle school samples (Stice, Fisher, & Martinez, 2004; Wells, Coope, Gabb, & Pears, 1985). Despite these outliers, overall, the proportions of individuals with low educational attainment included in validation studies are much smaller than those reported in community and epidemiological studies of binge eating (Marcus et al., 2007; Striegel-Moore et al., 2005). Although education level can only serve