Assessment methods for eating disorders and body image disorders

Ferenc Túry*, Hayriye Güleç, Elisabeth Kohls
Semmelweis University, Institute of Behavioural Sciences, Budapest, Hungary

Abstract

The growing interest in the treatment and research of eating disorders has stimulated the development of assessment methods, and there are now many questionnaires for evaluating behavioral and attitudinal characteristics of eating pathology. The present article sets out to review the assessment tools that are widely used in clinical practice and research. In particular, it covers self-report measures with summaries of their psychometric properties. It also presents diagnostic questionnaires based on the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV), diagnostic criteria. The instruments described include screening questionnaires, measurement tools for specific eating disorder symptoms, measurement of quality of life in eating disorders, and some tools for the measurement of body image disorder, a common feature of eating disorders. There is also a discussion of distorting factors that decrease the authenticity of assessment tools. These problems arise from the definition of some constructs and from the phenomena of denial and concealment, which are frequent among eating-disordered individuals. The frequent co-occurrence of other psychopathological features (e.g., multiimpulsive symptoms) shows that other psychological phenomena should also be evaluated in line with the assessment of eating disorders.

© 2010 Elsevier Inc. All rights reserved.

Keywords: Anorexia nervosa; Assessment methods; Binge eating disorder; Body image; Bulimia nervosa; Eating disorders

Introduction

Eating disorders came to the forefront of psychiatric/psychosomatic disorders in the last third of the 20th century. The reasons for this upsurge in interest include the serious outcome of anorexia nervosa (AN) and bulimia nervosa (BN) [1], the appearance of some newer forms [e.g., binge eating disorder (BED) [2,3]], and the increased incidence rate of AN during the past 50 years, particularly in females 10 to 24 years old [4].

The growing interest in eating disorders has resulted in the development of various assessment tools for screening and clinical evaluation. The first assessment of an eating-disordered patient requires a medical examination (possibly including laboratory tests) and an evaluation of the detailed history of the illness. Further assessment often involves questionnaires and interview methods, whose psychometric properties vary substantially. There are some very general problems to be faced in the assessment of eating disorders. The reluctance of patients to cooperate, poor compliance, denial of the illness, manipulative behavior, and hidden signs and symptoms generate obstacles in everyday practice [5,6]. Moreover, eating disorders can take on different characteristics during childhood. Some instruments have been developed specifically for the assessment of childhood eating disorders (for review, see reference [7]).

The major diagnostic categories that are addressed in most of the assessment methods are AN, BN, and BED. Binge eating disorder (i.e., binges without bulimic compensatory behaviors) is included in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)*, as a provisional diagnosis and has become an important category with high prevalence [2]. Some methods of assessment address certain transdiagnostic psychopathological issues, which are common in various types of eating disorder. Low self-esteem, impairment of social functioning (e.g., interpersonal problems and isolation), multiimpulsive features (e.g., alcohol and drug abuse, kleptomania,
suicidality, self-harm behavior), or other cognitive and behavioral components (e.g., preoccupation with weight and food and obsession) may occur in most of the diagnostic categories [8].

Eating and body image disorders show behavioral, attitudinal, cognitive, and perceptual characteristics. Some methods of assessment address only one of these dimensions, while others attempt to assess the disorders across a range of dimensions. In general, the behavioral signs and symptoms of a disorder may appear to be the simplest aspects to measure, but they are not always clearly evident in eating disorders. The definition of binges is problematic, because objective binges sometimes strongly differ from subjective binges; for example, the patient may believe she/he is binging but, actually, it cannot be supported objectively [9]. As for the cognitive factors, many distorted cognitions relating to eating and body shape are widespread in the general population (e.g., “people will like me more if I am thin”), and the boundaries between the healthy and pathological conditions are unclear. Moreover, perceptual, cognitive, and emotional factors play a role in the development of the body image. It is important to stress that the concept of body image cannot be solely limited to visual input. Some authors use the term body experience, reflecting the complexity of this concept [10].

Measurement methods for eating and body image disorders are usually administered by either interview or self-reporting. Structured clinical interviews are widely seen as gold standards in clinical practice and research. They are essential for the detection of physical signs and symptoms, medical complications, somatic and psychiatric comorbidity, and different subtypes and essential to clarify differential diagnostic issues [11]. The major measures with excellent psychometric properties include the Structured Clinical Interview for DSM-IV TR Axis I Disorders (SCID-I [12]), Eating Disorder Examination (EDE [13,14]), and the Structured Interview for Anorexic and Bulimic Disorders (SIAB-EX [15]). The SCID-I is a general psychiatric interview method based on the criteria of DSM-IV TR, including items for assessing eating disorders symptoms. The 12th version of EDE was published in 1993 and is the primary tool in treatment studies. Its 23 variables include diagnostic items and form four subscales: eating concern, shape concern, dietary restraint, and weight concern. Having gone through several revisions, the SIAB-EX provides valid diagnosis of AN and BN (including subtypes) consistent with both the DSM-IV and the International Statistical Classification of Diseases, 10th Revision (ICD-10), criteria [16]. All the structured clinical interviews should be administered after thorough training to ensure the validity of the diagnosis and assessment of symptoms. In clinical settings and in the definition of the complex constructs, such as binge and dieting, they can be useful and more precise than self-report measures. All clinical interviews, however, are time consuming and costly. Self-report methods are widely used in screening, research, and assessment of eating and body image pathology. Although the administration of self-report measures is easy and economic, a number of limitations arise from subjective interpretations of the phenomena particular to eating disorders, such as binges [17] and denial of the illness resulting in the concealment of symptoms, which might threaten the validity of these measures [18]. The effect of social desirability can also distort the interpretation of the results. The real occurrence of amenorrhea, binge, and purging behavior is often kept secret. Distress and shame often lie behind concealment [18]. These influences can cause underreporting of the severity of the symptoms, mostly in anorectics.

This review summarizes some of the best-known self-report assessment tools with good psychometric qualities. It is intended to provide a pool of options for both clinicians and researchers. There are also good manuals and reviews summarizing several useful methods [17,19–22].

General measures of eating disorder symptoms

Eating Disorders Inventory

The Eating Disorders Inventory (EDI [23]) was developed to measure behavioral and psychological traits in AN and BN. The items are based on extensive literature and research experience [24]. The EDI has 64 items comprising eight subscales: drive for thinness, bulimia, body dissatisfaction, ineffectiveness, perfectionism, interpersonal distrust, interoceptive awareness, and maturity fears. Three subscales assess the attitudes toward weight, body shape, and eating (drive for thinness, bulimia, and body dissatisfaction), and the remaining five subscales assess the psychological characteristics of individuals with eating disorders. The EDI was revised to EDI-2 by Garner [25], retaining eight subscales but adding 27 items that constituted three new subscales: asceticism, impulse regulation, and social insecurity. Recently, the EDI was revised to EDI-3 by Garner [26], retaining the same items from EDI-2. Some of the items show different item loadings compared with EDI-2 [27]. The three EDI-2 subscales assessing the core eating pathology symptoms are unchanged in the EDI-3 except for the addition of one item from the EDI-2 interoceptive awareness subscale to the bulimia scale and the body dissatisfaction scale. Low self-esteem, personal alienation, interpersonal insecurity, interpersonal alienation, interoceptive deficits, emotional dysregulation, perfectionism, asceticism, and maturity fears subscale make up the rest of the EDI-3.

The EDI, EDI-2, and EDI-3 are 6-point forced-choice self-report scales transformed into a 4-point scale; the three extreme responses related to disordered behaviors are scored, and the choice opposite to the most anorexic response gets a score of 0. The subscales scores can be used separately or summed to give a total score. For EDI-3, seven additional composite scores may be obtained by summing the T-scores...
of two or more scale scores. Eating Disorders Inventory, EDI-2, and EDI-3 are widely used and are easy to administer and score. The EDI can be used both as a screening instrument and a tool for diagnosis. It is a good measure of treatment outcome and useful at differentiating the levels of severity [28]. As a screening instrument for eating disorders, using untransformed scoring of the items in the nonclinical population has been suggested to prevent low variability and lack of internal consistency [29].

Generally, EDI subscales have good psychometric properties. Internal consistencies of the eight subscales range from .80 to .91. However, Cronbach’s α values are less than .80 for the three new subscales of EDI-2 [30]. The maturity fears subscale showed lower Cronbach’s α values in both studies. Test–retest reliability was higher than .80 among a nonclinical female sample across a 3-week period [31]. Test–retest reliability of EDI-3 showed very good test–retest correlations, with the intervals being 1 week (.79–.95), 3 weeks (.81–.97), or 1 year (.41–.75) [27].

The EDI and all of its subscales have been found to discriminate AN patients from college females and BN from control participants, but the results were contradictory when trying to discriminate between subgroups of AN and BN. An attempt to replicate the factor structure with nonclinical samples led to conflicting results, suggesting differences in how clinical and nonclinical individuals interpret the items [20].

Eating Disorder Examination Questionnaire

The EDE Questionnaire (EDE-Q [32]) is adapted from the semistructured interview EDE [13] and is a widely used measure of eating disorder symptoms. The questionnaire is designed to assess and describe the specific psychopathology of eating disorders, including severity of eating pathology, key behavioral problems, and associated disturbances over the past 28 days. It consists of 33 items and contains four subscales, as does the EDE. Luce and Crowther [33] reported excellent internal consistencies for the EDE-Q subscales, with Cronbach’s α ranging from .78 to .93, and the test–retest reliability ranged from .81 to .94 across a 2-week period. Several studies have made comparisons between EDE and EDE-Q. Fairburn and Beglin [32] reported high similarities in the assessment of behavioral symptoms, for example, dieting and self-induced vomiting (correlations higher than .80). However, several studies suggest greater variability between EDE and EDE-Q when assessing more complex symptoms such as binge eating. Overall, this measure shows excellent psychometric properties in different studies, which supports its use in research and clinical settings. Its version 6.0 was published in 2008 [34]. According to Wade et al. [35], the EDE is likely to perform well as a diagnostic and predictive tool. However, caution should be exercised with respect to the use and interpretation of the individual subscales in nonclinical populations of young adolescents.

Anorexia Nervosa Inventory for Self-Rating

The Anorexia Nervosa Inventory for Self-Rating (ANIS [36]) was developed as an instrument for diagnostic purposes and for longitudinal assessments of anorectic patients. Items were selected from a larger item pool by expert ratings and subjected to factor analysis. The factor analytic studies yielded a 32-item self-rating scale with six factors: figure consciousness, feeling of insufficiency, anacasm, adverse effect of meals, sexual anxieties, and bulimia. The ANIS is designed to discriminate between healthy and anorectic subjects and is intended as an assessment tool during treatment of anorectic patients. Several studies [36,37] showed the ANIS as a psychometrically sound instrument (internal consistencies between .57 and .90).

DSM-IV diagnostic questionnaires

In addition to interviews (SCID and SIAB-EX, mentioned above), there are some self-report instruments based on the DSM-IV criteria [2].

Eating Disorder Diagnostic Scale

The Eating Disorder Diagnostic Scale (EDDS [38]) has recently been developed as a brief self-report scale for diagnosing AN, BN, and BED. Its items were adapted from validated structured interviews for the assessment of eating pathology. The scale consists of 22 items. The authors reported this scale as reliable and valid in different studies [38,39] and recommend it for clinical and research applications.

Questionnaire for Eating Disorder Diagnoses

The Questionnaire for Eating Disorder Diagnoses (Q-EDD [40]) is a 50-item self-report questionnaire. It operationalizes DSM-IV [2] criteria for eating disorders into a self-report format yielding categorical labels: the general categories are eating disordered (matching the DSM-IV diagnosis) and non-eating disordered (not matching). Within the non-eating-disordered category, there are two subcategories: asymptomatic (no eating disorder symptoms) and symptomatic (no diagnosable disorder, but symptoms). Within the eating-disordered category, six specific diagnoses may be distinguished: AN, BN, eating disorder not otherwise specified (EDNOS)—subthreshold BN, EDNOS—menstruating AN, EDNOS—nonbinge BN, and BED. The authors described several studies, which analyzed the psychometric properties of the Q-EDD and yielded sound results. In general, they reported strong support for its use in terms of discrimination between (1) eating disorders vs. no eating disorder; (2) eating disorder, symptomatic vs. asymptomatic; and (3) AN and BN.
The Questionnaire on Eating and Weight Patterns—Revised (QEWP-R [41]) is a self-administered 28-item questionnaire. It has been developed as a screening instrument to assist in the identification and diagnosis of individuals with BED according to the DSM-IV provisional diagnostic criteria [2], BN, and subclinical forms of BED. The psychometric properties are limited, but overall, the QEWP-R shows adequate reliability and validity [42].

Screening questionnaires

Eating Attitudes Test, Eating Attitudes Test-26

The Eating Attitudes Test (EAT [43]) was developed to assess attitudes and behaviors related to eating disorders via self-report. The original EAT consisted of 40 items; the 26-item version was developed later [44]. Each item is rated on a Likert-type scale from 1 (always) to 6 (never). Reliability and validity have been documented for both EAT and EAT-26 [43,44]. The EAT shows internal consistencies between .76 and .94 for different samples. Moreover, the EAT is highly correlated with the EDI. Although the EAT-26 is supported by reliability and validity studies, several doubts have been raised as to whether its sensitivity and specificity are sufficient for screening [40].

Bulimia Test

The original form of the Bulimia Test (BULIT [45]) was published in 1984, and the revised form (BULIT-R [46]) was published in 1991. It is a 36-item self-report questionnaire. Binge eating, compensatory behavior, and weight and shape concerns consistent with core diagnostic criteria for BN are measured on a 5-point Likert scale. Of the 36 items, 26 are used to determine the final score. Total scores are obtained by summing the 28 items and range from 28 to 140, and 104 is considered as cutoff for BN classification. The authors report high internal consistency (between .97 and .98) and test–retest reliability (=.95); validity data are also quite strong [46–48] showing high correlations with other related assessments of eating disorders and behaviors, as well as DSM-III-R [49] and DSM-IV [2] BN diagnosis.

Bulimia Investigatory Test, Edinburgh

The Bulimia Investigatory Test, Edinburgh (BITE [50]), a self-rating scale for BN, is designed for use in epidemiological studies and as a measure for treatment outcome. The questionnaire consists of 33 items (contains a total score, a symptom subscale, and a severity subscale). The test aims to detect and describe binge eating. The BITE showed internal consistencies between .96 for the symptom subscale and .62 for the severity subscale [50,51].

SCOFF

The SCOFF [52] was designed to provide a simple and short screening tool for nonspecialists (comparable with the CAGE questionnaire for alcohol misuse [53]). The authors developed five questions addressing core features of AN and BN; these give the acronym SCOFF. Two or more positive answers are recommended as cutoff, indicating a likely case of AN or BN. Its purpose is to raise suspicion of a likely case rather than to diagnose. A high correlation has been found between the questionnaire and a clinical interview based on DSM-IV criteria [54]. However, the reported sensitivity rate of 100% [52] could not be replicated in other studies [55]. Mond et al. [56] studied the comparative validity of the EDE-Q [13,14] and the SCOFF in screening for cases in a primary care sample of young adult women. The authors conclude that both measures performed well in terms of their ability to detect cases and to exclude noncases of commonly occurring eating disorders in primary care settings. The findings suggest the EDE-Q to perform somewhat better than the SCOFF and being more robust to effects on validity of age and weight. It was concluded that these findings need to be weighed against the brevity advantages of the SCOFF. In sum, the SCOFF may be considered as a sound screening tool with sufficient psychometric properties.

Short Evaluation of Eating Disorders

The Short Evaluation of Eating Disorders (SEED [57]) was developed for the quick assessment of key eating disorder symptoms and repeated assessments over time. The questionnaire consists of six items, measuring three main symptoms for AN (degree of underweight, fear of gaining weight, and distortion of body perception) and BN (amount of binge eating, amount of compensatory behavior, and overconcern with body weight and shape). This allows the calculation of two total severity indices for AN and BN. This total severity indices range from 0 (no symptoms) to 3 (extreme symptoms); see Bauer et al. [57] for detailed description and operationalization. The questionnaire may be used repeatedly, even after short time intervals. Data from two patient samples and one nonpatient sample as well as therapist ratings were used to investigate different validity aspects of the SEED, overall with positive results.

Measures of specific eating disorder symptoms

Mizes Anorectic Cognitions Scale

The Mizes Anorectic Cognitions Scale (MAC [58]) was developed to evaluate the typical cognitions of eating disorders that precede the behavioral components of both AN and BN. It provides a continuous measure of the severity of cognitive distortions in a 5-point Likert scale. There are three domains in the self-report scale: the perception of
weight and eating as the basis of approval from others, the belief that rigid self-control is fundamental for self-worth, and the rigidity of efforts to regulate weight and eating. The MAC has advantages in identifying high-risk individuals, is useful for the detection of therapeutic change, and is sensitive to decreases in eating pathology. There are three versions [59]. The original MAC contains 45 items, and there is a 33-item shortened form. The MAC—Revised (MAC-R) comprises 24 items, with three subscales (each contains eight items): self-control and self-esteem; appearance, weight, and approval; rigid weight regulation and fear of weight gain. The possible range is 24 to 120 scores for MAC-R.

The psychometric qualities of the MAC have been corroborated by several studies [60–64]. Cronbach’s α is reported between .75 and .89 for internal consistency of the subscales, and .91 for the entire scale. Test–retest reliability showed a correlation coefficient of .78 with 2 months between the two data collection. As for MAC-R, the Cronbach’s α was .82 to .85 for the subscales and .90 for the whole test. In the validity studies, the MAC correlated well with other scales of eating pathology and discriminated eating-disordered patients and other psychiatric patients [60–64]. The MAC-R also showed good concurrent validity and criterion-related validity. The instrument is good at discriminating between AN and BN [63].

Forbidden Food Survey

The Forbidden Food Survey (FFS [65]) is designed to assess food phobia. A basic assumption is that eating-disordered subjects experience a high degree of anxiety related to certain foods. The instrument contains 45 food items. These represent foods from five groups and three caloric levels (low, medium, and high). The items are rated in a 5-point Likert scale, from “I would feel very good about myself after eating this food,” to the answer “… very bad.”

A study by Schlundt and Johnson [66] showed Cronbach’s α above .80 for all scales except those for milk, low-calorie foods, and beverages. The test–retest reliability was .63 for the domain milk and .90 for the domain meat. In different diagnostic groups, the binge-purging subjects scored the highest, and non-eating-disordered subjects showed the lowest scores. The intercorrelations with other eating disorder scales were relatively high. The FFS seems to be a valid instrument for differentiating clinical and nonclinical groups [12,66].

Three-Factor Eating Questionnaire or Eating Inventory

The Three-Factor Eating Questionnaire (TFEQ [67]) or Eating Inventory is a self-report instrument to evaluate three dimensions of eating behavior: cognitive restraint in eating, susceptibility to periodic disinhibition of control over eating, and perceived hunger. These factors play a role in the pathomechanism of eating disorders, including obesity. The questionnaire contains 51 items in two parts: 36 items ask for true/false judgment, and 15 items are scored in a 4-point Likert-type scale. The items were derived from the Restraint Scale [68], from an obesity questionnaire, and from the clinical experience of the first author. The cutoff score of the three subscales are 14, 12, and 11. The factor structure is inconsistent in different studies. However, these factors have prognostic value because high cognitive restraint and low disinhibition predict a successful outcome in losing weight, whereas high disinhibition relates to binge eating.

The reliability and validity studies showed good internal consistencies, with Cronbach’s α of .80 to .93 for the three subscales [69]. The test–retest reliability was .80 to .93 for the subscales in a small sample. The TFEQ scores correlated well with the severity of binge eating, permitting easy differentiation between bulimic and control subjects. The cognitive restraint subscale is applicable for monitoring treatment progress in obesity and BED, because positive response to treatment correlates with increases in scores on this subscale, and decreases in the scores in the disinhibition and hunger subscales. The TFEQ is a sensitive indicator of changes in the cognitive and behavioral components of restraint, disinhibition, and hunger.

Dutch Eating Behavior Questionnaire

The Dutch Eating Behavior Questionnaire (DEBQ [70]) is a 33-item self-report questionnaire with a 5-point Likert scale to assess three separate factors of eating behavior: restrained eating, emotional eating, and external eating. The items were derived from 100 items used in the authors’ previous research and from other obesity questionnaires. The Restraint Eating Scale contains items relating to weight control. The Emotional Eating Scale examines the amount of eating related to emotional states. The items of the External Eating Scale deal with external cues stimulating eating.

Van Strien et al. [70] reported Cronbach’s α of .94 to .95. The test–retest reliability was .92 after 2 weeks [71]. The factor structure of the DEBQ proved to be stable, although Ogden [72] described that the Restraint Eating Scale confounds two aspects of caloric restriction, the intentions to diet and the success of dieting. Convergent, discriminative, and concurrent validity studies have corroborated the good psychometric characteristics of the DEBQ [19,72].

Binge Eating Scale

The Binge Eating Scale (BES [73]) was originally developed to assess binge eating experience in obese individuals, but it has been applied to other populations as well. The BES includes 16 items; a global score for eating disturbance is obtained by summing the responses to the 16 items. Each item includes “I” statements, and participants are asked to choose the statement that best fits his or her situation. The BES provides a total score of severity, but this score can neither be used to measure binge frequency nor give specific data for diagnostic purposes. One study found
the BES to be better as an initial screening tool for BED but is less accurate in identifying the non-BED individuals [74]. Internal consistency of the BES was studied by Gormally et al. [73]. They compared the total scores to each item on the scale, and y² tests of significance were reported 9.1 or higher (P<.01). The comparison between the BES scores and interview-rated binge eating severity levels revealed that the BES significantly discriminated between interviewer-rated severity levels.

**Binge Scale Questionnaire**

The Binge Scale Questionnaire [75] was developed to assess the behavioral and attitudinal dimensions of binge eating and vomiting. It is a nine-item multiple choice questionnaire that gives a continuous global score of severity related to bulimic symptomatology. The Binge Scale Questionnaire is basically a screening instrument and cannot be used for diagnostic purposes because it only reveals the occurrence of binges and vomiting and does not assess their frequency [76]. A score of 10 or less on the scale represents occurrence of binges and vomiting and does not assess their frequency. A score of 10 or less on the scale represents the norm for the normal population, and a score of 15 or higher suggests the presence of an eating disorder. Hawkins and Clement [75] reported Cronbach’s α of .68 for the total score. One-month test–retest reliability of the total score is reported as .88. Although the Binge Scale Questionnaire provides a total score, the authors also conducted a principal components factor analysis with varimax rotation, which yielded a two-factor solution: feeling of guilt and concern about binge eating, and behavioral aspects of binge eating. The correlation between the Binge Scale Questionnaire and the BULIT-R is reported to be .93.

**Revised Restraint Scale**

The Revised Restraint Scale has been developed to identify individuals who are chronically concerned about their weight and attempt to control or reduce it by dieting or curtailing food intake. The first version contained five items and was developed by Herman and Mack [77] to measure chronic dieting. This scale has been revised by Herman and Polivy [68] to an 11-item form. The final version contains 10 items with two subscales: weight fluctuation and concern for dieting subscale [78]. As for the scoring format, some items are scored in a 5-point and others in a 4-point Likert-type scale. The higher the score on the overall scale, the higher the probability of unsuccessful dieting with frequent attempts to restrain eating, resulting in loss of control. The norms of the total score for females and males are suggested to be around 16 and 12, respectively. Test–retest reliability of the scale is quite stable and high (ranging from .74 to .95). Internal consistency of the whole scale typically exceeds .75 but varies considerably depending on the sample characteristics (either obese or normal weight subjects). The subscale consistencies are even lower [78]. As for the content validity of the scale, a two-factor solution seems reasonable in general; however, three or more factor solutions have also been reported when the sample included individuals with BN or obesity. Some evidence has been reported of the scale being confounded by socially desirable response sets.

**Assessment of body image**

A special problem in the assessment of eating disorders is the regular coincidence of body image disorder, especially in AN. Body image disorder has become a topic of intensive study in recent decades, and several instruments have been developed to measure it. In this section, three frequently used assessment tools will be reviewed.

The most common elements of assessment are dissatisfaction, self-evaluation of body size and body composition, attitudes about gaining weight, preoccupations, and rituals (e.g., checking and avoidance behaviors [79]).

**Body Shape Questionnaire**

The Body Shape Questionnaire (BSQ [80]) contains 34 items, each with a 6-point Likert-type scale, and measures concerns with body shape and size among adult men and women. The questions relate to the previous 4 weeks. The total score range is 34 to 204. The items were derived by conducting semistructured interviews with various groups of women including patients with AN and BN. There is a shortened form of the BSQ [81], which omits two items and divides the rest into two 16-item scales. These two short scales were found to correlate well with each other and with the original 34-item form (r=.96–.99). There is another short and validated form of the BSQ containing 14 items [82].

The reliability studies showed high consistency; Cronbach’s α was .97 [83]. The test–retest reliability was high with a correlation of .88 in a 3-week span. In the validity studies, the BSQ was compared with other body image tests (e.g., the EDI Body Dissatisfaction Subscale), and the correlation proved to be high (between .61 and .81). In other studies, different clinical and nonclinical groups were compared, and the BSQ scores could distinguish these groups [83]. The BSQ provides a comprehensive measure of body shape and weight concerns and is useful for the detection of individuals having a high risk of eating disorders.

**Physical Appearance State and Trait Anxiety Scale**

The Physical Appearance State and Trait Anxiety Scale (PASTAS [84]) measures the anxiety associated with 16 body sites. Eight of them are weight related, and eight are non-weight related. The instrument has trait and state forms, the latter being designed for repeated use. The frequency and the intensity of the anxiety related to the body part is measured by a five-point Likert scale. Rutt et al. [85] performed a reliability study of the PASTAS. The factor structure was examined with exploratory and confirmatory
factor analyses in a sample of Hispanic women with normal body mass index. A new three-factor solution (weight-related lower torso, weight-related midtorso, and non-weight related) was found, with fit indices ranging from .88 to .93.

**Body Attitude Test**

The Body Attitude Test (BAT [86]) was developed for female eating-disordered patients. It is a practical self-report instrument containing 20 items with a 6-point Likert-type scale. The range of the item scores is 0 to 5 and the maximum score is 100. The authors recommend a cutoff score of 36 for pathological body attitude. The psychometric properties of BAT have been tested in a large number of patients and control subjects, for example, eating-disordered persons, weight watchers, and normal subjects [10,86,87]. The factor analyses yielded a stable four-factor structure: negative appreciation of body size, lack of familiarity with one’s own body, general body dissatisfaction, and a rest factor. Moreover, the BAT was administered together with other self-report measures (Body Attitude Questionnaire [88], Body Shape Questionnaire, EDI, Rosenberg Self-Esteem Scale) in an anorectic and in a bulimic sample, and in a group of female university students. The results indicate that the negative body attitude expressed on the BAT is related to other signs of negative body experience. The BAT differentiates well between clinical and nonclinical subjects and between anorectics and bulimics.

**The Sociocultural Attitudes Towards Appearance Scale-3**

The Sociocultural Attitudes Towards Appearance Scale-3 (SATAQ-3 [89]) measures multiple aspects of a societal influence on body image and eating disturbances. It is a revision and extension of the SATAQ, first reported by Heinberg et al. [90]. In the study of Thompson et al. [89], factor analyses were conducted to determine the underlying structure of the revised scale. The resulting subscales consisted of 30 items (Internalization-General, nine items; Information, nine items; Pressure, seven items; Internalization-Athlete, five items). Cronbach’s α on the subscales ranged between .92 and .96. Moreover, the SATAQ-3 showed excellent convergent validity with measures of body image and eating disturbances [89]. Eating-disturbed and eating-disordered samples had higher scores on SATAQ-3 subscales than a control sample. The authors recommend this measurement for basic risk factor studies as well as for gauging the efficacy of prevention and treatment programs.

**The Body Appreciation Scale**

The Body Appreciation Scale (BAS [91]) is an easily administered 13-item self-report questionnaire, reflecting the body appreciation and the positive aspects of the body image. The generally applied body image measurement methods address mainly psychopathological features and the negative aspects of the body image. One important characteristic of the BAS is that it is relatively gender neutral.

The items of the BAS are rated along a 5-point Likert-type scale (1–5, from “never” to “always”). Higher scores reflect greater body appreciation. The psychometric properties were examined in four stages. The Cronbach’s α was .94, showing a good internal consistency. The temporal stability over a 3-week period was adequate (r=.90). The results of the exploratory factor analysis provided support for the unidimensionality of the BAS. The questionnaire demonstrated evidence of good construct and incremental validity as well [91].

**Measurement of quality of life in eating disorders**

**The Eating Disorders Quality of Life Instrument**

The Eating Disorders Quality of Life Instrument (EDQOL [92]) is a health-related quality of life measure that was developed to assess the impact of the illness on various areas of eating-disordered patients. It is a 25-item instrument consisting of four subscales: psychological, physical/cognitive, financial, and work/school. The items were generated in three primary steps: domain generation, content generation, and item generation. Scale scores are obtained by averaging the items of each scale. Similarly, total score is obtained by averaging all the items of the EDQOL. Higher scores indicate better quality of life. The EDQOL was validated on female participants with a large sample size (N=558). The Cronbach’s α coefficients for single subscales are above .84, showing good internal consistency. One-week test–retest reliabilities were quite high (above .87) except the work/school subscale (.14). The EDQOL differentiates between eating-disordered groups compared with diet–exercise and non-eating-disordered groups; the eating-disordered group showed greater impairment on all subscales, lending support for known groups validity. The EDQOL is more sensitive to subscale variability compared with generic measures of quality of life.

**The Eating Disorders Quality of Life Scale**

The Eating Disorders Quality of Life Scale (EDQOLS [93]) was developed as a condition-specific quality of life measure for adolescents and adults with eating disorders. The EDQOLS has 40 items and 12 domains. Four sources of material were used for domain and item generation: the peer-reviewed literature, treatment provider interviews, patient/client interviews, and internet-based first person narratives. Final domains consist of cognitive functioning, education/vocation, family and close relationships, relationships with others, future/outlook experience, appearance, leisure, psychological health, emotional health, values and beliefs,
physical health, and eating attitudes. The EDQOLS includes response scaling with a 5-point scale, and the total score is obtained summing the item degrees. The mean total score in a multisite eating-disordered sample was 110 (SD=27.6; range, 56–187), and the highest score was 200. Higher scores indicate better quality of life. The scale was found to be internally consistent (Cronbach’s α, .96). Cronbach’s α values for domains and the total score were higher than .50. Patterns of quality of life across all variables were in the expected direction, and the measure was sensitive to changes occurring over time.

The reviewed instruments are summarized in Table 1. The availability of the questionnaires may be subject to copyright, and the user must contact the authors.

### Conclusion

The complex nature of eating disorders calls for the development of more comprehensive measurement tools. The self-report measures of eating disorders used at present do not evaluate the full range of behavioral and attitudinal symptoms [6]. New methods are also required to accommodate newly appearing syndromes and subtypes (e.g., orthorexia nervosa, which means an obsessional dependence on healthy food).

There are several distorting factors that impair the authenticity of the instruments. First, some problems arise from the definition of some constructs (e.g., body image, preoccupation with food, etc.). Second, the very common

<table>
<thead>
<tr>
<th>Name</th>
<th>Reference no.</th>
<th>No. of items</th>
<th>Self-administered/rater administered</th>
<th>Time estimation (min)</th>
<th>Reliability a</th>
<th>Validity a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structured clinical interviews</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCID-I</td>
<td>[12]</td>
<td>Semistructured interview</td>
<td>30–60</td>
<td>++</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>EDE</td>
<td>[13]</td>
<td>Semistructured interview</td>
<td>30–60</td>
<td>++</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>SIAB-EX</td>
<td>[15,16]</td>
<td>Semistructured interview</td>
<td>30–60</td>
<td>++</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>General measures of eating disorder symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDI</td>
<td>[23]</td>
<td>64</td>
<td>Self-administered</td>
<td>15</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>EDI-2</td>
<td>[25]</td>
<td>91</td>
<td>Self-administered</td>
<td>20</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>EDI-3</td>
<td>[26]</td>
<td>91</td>
<td>Self-administered</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDE Questionnaire (EDE-Q)</td>
<td>[32]</td>
<td>38</td>
<td>Self-administered</td>
<td>10–15</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>ANIS</td>
<td>[36]</td>
<td>32</td>
<td>Self-administered</td>
<td>10</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>DSM-IV diagnostic questionnaires</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDDS</td>
<td>[38]</td>
<td>22</td>
<td>Self-administered</td>
<td>5–10</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Q-EDD</td>
<td>[40]</td>
<td>50</td>
<td>Self-administered</td>
<td>15–20</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>QEWR-P</td>
<td>[41]</td>
<td>28</td>
<td>Self-administered</td>
<td>5–15</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Screening questionnaires</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EAT</td>
<td>[43]</td>
<td>40</td>
<td>Self-administered</td>
<td>10</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>EAT-26</td>
<td>[44]</td>
<td>26</td>
<td>Self-administered</td>
<td>5</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>BULIT-R</td>
<td>[45], revised: [46]</td>
<td>36</td>
<td>Self-administered</td>
<td>10</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>BITE</td>
<td>[50]</td>
<td>33</td>
<td>Self-administered</td>
<td>10</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>SCOFF</td>
<td>[52]</td>
<td>5</td>
<td>Self-administered</td>
<td>2</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>SEED</td>
<td>[57]</td>
<td>6</td>
<td>Self-administered</td>
<td>&lt;5</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Measures of specific eating disorder symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAC</td>
<td>[58]</td>
<td>45</td>
<td>Self-administered</td>
<td>10</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>MAC-R</td>
<td>[63]</td>
<td>24</td>
<td>Self-administered</td>
<td>5</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>FFS</td>
<td>[65]</td>
<td>44</td>
<td>Self-administered</td>
<td>10</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>TFEQ</td>
<td>[67]</td>
<td>51 (in two parts: 36+15)</td>
<td>Self-administered</td>
<td>10–15</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>DEBQ</td>
<td>[70]</td>
<td>33</td>
<td>Self-administered</td>
<td>10</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>BES</td>
<td>[73]</td>
<td>16</td>
<td>Self-administered</td>
<td>5</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Binge Scale Questionnaire</td>
<td>[75]</td>
<td>9</td>
<td>Self-administered</td>
<td>2–3</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Revised Restraint Scales</td>
<td>[68]</td>
<td>10</td>
<td>Self-administered</td>
<td>2–3</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Assessment of body image</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body Shape Questionnaire (BSQ)</td>
<td>[80]</td>
<td>34</td>
<td>Self-administered</td>
<td>10</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>BSQ—shortened form</td>
<td>[81]</td>
<td>16</td>
<td>Self-administered</td>
<td>5</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>PASTAS</td>
<td>[84]</td>
<td>16–16 (state, trait)</td>
<td>Self-administered</td>
<td>5–5</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>BAT</td>
<td>[86]</td>
<td>20</td>
<td>Self-administered</td>
<td>5–10</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>SATAQ-3</td>
<td>[89]</td>
<td>30</td>
<td>Self-administered</td>
<td>10</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>BAS</td>
<td>[91]</td>
<td>13</td>
<td>Self-administered</td>
<td>5</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Measurement of quality of life in eating disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDQOL</td>
<td>[92]</td>
<td>25</td>
<td>Self-administered</td>
<td>5</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>EDQOLS</td>
<td>[93]</td>
<td>40</td>
<td>Self-administered</td>
<td>10–15</td>
<td>++</td>
<td>++</td>
</tr>
</tbody>
</table>

* a – indicates incomplete data; +, some information; ++, excellent psychometric properties.
occurrence of denial and concealment among eating-disordered individuals may lead to subjective underestimation of the symptoms. Third, other psychopathological features (e.g., multimpulsive symptoms) frequently co-occur with eating disorders, indicating the need to evaluate additional psychological phenomena. Fourth, there are major concerns surrounding culture-specific individual differences [20] influencing both the course of illness in general and the reactions of individuals to the specific assessment tools (self-report and interview). Fifth, the gender specificity of the instruments has to be taken into consideration. The majority of the measurement tools were developed for females. There are relatively gender-neutral questionnaires (e.g., BAS [91]).

The increase of the prevalence of the male disorders raises concerns surrounding culture-specific individual differences [20] influencing both the course of illness in general and the reactions of individuals to the specific assessment tools (self-report and interview). Fifth, the gender specificity of the instruments has to be taken into consideration. The majority of the measurement tools were developed for females. There are relatively gender-neutral questionnaires (e.g., BAS [91]).

Eating-disordered patients have to be assessed using several methods in parallel, in a process involving several sessions rather than a single cross-sectional evaluation [11,94]. It is difficult, however, to find the right balance between maximizing information and optimizing costs and benefits. One possibility, in line with the technological improvements in mental health applications, is the increased use of online assessment tools. In the future, we can expect a higher number of studies validating online instruments of eating disorders.

Acknowledgments

This work was supported by funding from the Marie Curie Research Training Network INTACT (Individually Tailored Stepped Care for Women with Eating Disorders; reference number: MRTN-CT-2006-035988). Participating centers: Heidelberg, Germany (coordinating centre): H. Kordy, S. Bauer; St. Etienne, France: F. Lang; Prague, Czech Republic: H. Papežová; London, UK: U. Schmidt; Braga, Portugal: P. Machado; Budapest, Hungary: F. Türy; Utrecht, the Netherlands: R. Adan; Geneva, Switzerland: A. Golay; NetUnion sarl, Lausanne, Switzerland: T. Lam.

References


