Rodgers, R. F., & Franko, D. L. (2015). Screening for eating disorders: An updated guide. In U. Chichester, L. Smolak, & M. P. Levine (Eds.), The Wiley handbook of eating disorders (Vol. 2, pp. 509–523). Chichester, UK: John Wiley & Sons.

# 38

# Screening for Eating Disorders: An Updated Guide

Rachel F. Rodgers and Debra L. Franko
*Department of Counseling and Applied Educational Psychology, Northeastern University, USA*

While base rates of eating disorders (EDs) in the general population are relatively low (see Chapters 5 & 37), early identification is strongly associated with better prognosis (Becker, Franko, Nussbaum, & Herzog, 2004). Furthermore, subthreshold disorders are associated with high levels of distress, and identifying individuals suffering from partial forms of EDs may provide a target for early intervention (Becker et al., 2004). While the literature is not fully convergent, it has been suggested that subthreshold forms of EDs may constitute precursors of full-syndrome EDs, at least among certain groups (Stice, Ng, & Shaw, 2010). Early referral may therefore assist with both improved prognosis and prevention of progression. In addition to early referral, screening for EDs may be conducted with the intention of better targeting an intervention, or establishing prevalence data. Screening tools therefore need to provide accurate information regarding ED status, but also be suitable for use with large groups such as school or university settings, to provide national prevalence data (e.g., Youth Behavior Risk Survey). Developing accurate screens may therefore contribute to better targeting of existing resources and improving the mental health of individuals suffering from or at risk for EDs.

## Properties of Screening Instruments: How to Know When a Screening Tool is Useful

An important requirement for a screening instrument is validity, and in particular criterion validity, defined as how well the screen identifies real cases and non-cases. The criterion validity of a screening instrument can be determined by comparing the results obtained with the screen to the results obtained using a "gold standard" test in the same population, that is, the best available test at the time.

*The Wiley Handbook of Eating Disorders*, First Edition. Edited by Linda Smolak and Michael P. Levine.

## Sensitivity and Specificity

Sensitivity and specificity are two indices of a scale's diagnostic performance, with high values indicating that the scale is highly accurate at classifying cases and non-cases in the correct category. Thus, an instrument with low sensitivity and specificity would both "miss" some of the individuals who actually should have the diagnosis, and misclassify healthy individuals (false positives). The sensitivity of a screening instrument refers to the proportion of cases accurately detected, that is, the proportion of true positives (i.e., cases) correctly identified by the screen. In other words, sensitivity tells us the proportion of people who should have been identified as having an ED who are identified as having an ED.

The specificity refers to the proportion of true negatives identified by the screen. Negatives are people who do not suffer from the disorder, in this case EDs. True negatives are those identified as not having the disorder who don't, while false negatives are those who actually have an ED but are identified as healthy. Thus it assesses the proportion of healthy people who do not have an ED who are classified as not having an ED. The specificity is equal to the number of non-cases correctly identified by the screen over the number of "true" non-cases as determined by the gold standard.

Importantly, there are no fixed guidelines for what constitutes acceptable sensitivity and specificity. Values around 80% are generally considered good; however, it has been suggested that among low-prevalence disorders such as EDs the sensitivity and specificity should be close to 90% (Williams, Hand, & Tarnopolsky, 1982).

## Positive and Negative Predictive Value

The specificity and sensitivity of a test are useful parameters in that they respectively indicate how many individuals will be correctly classified as healthy or presenting the diagnosis. However, sometimes it may be more useful to know how many of the cases identified by a screening test are true cases. In other words, is the screening measure overinclusive in the cases it identifies, or on the contrary, too narrow? For example, are people who are healthy being mistakenly identified as having EDs? Positive and negative predictive values address such questions.

The positive predictive value refers to the proportion of individuals who are identified as cases by the screen who actually are cases. Thus it evaluates the capacity of the test to include only real cases as cases and not be overinclusive. The positive predictive value is highly dependent on the prevalence of the disorder. A correlate of the importance of prevalence in determining positive and negative predictive values is that for low-prevalence disorders, such as anorexia nervosa (AN), the positive predictive value can be very low even when sensitivity and specificity are high (Jacobi, Abascal, & Taylor, 2004). The negative predictive value correspondingly refers to the proportion of non-cases who are correctly identified as healthy by the screen. It evaluates the capacity of the measure to capture all of the cases.

Depending on the objective of the screening, it may be desirable to maximize different properties. In the case of screening for cases for treatment, for example, it is desirable to minimize the number of falsely identified negatives, thus increasing sensitivity, so as to be sure not to miss cases. In the context of wide screening for risk factors for EDs, however, such as might be conducted as an initial step on a multiphase screening program, a screening measure should minimize the proportion of false positives among those scoring above the cut-offs (i.e., the positive predictive value should be maximized) to reduce the number of interviews to be carried out in the second phase.

### Caution in the Use of Screening Instruments for ED

In their review, Jacobi et al. (2004) called for caution in the use of screening instruments, identifying a number of limitations of the available instruments, including an overwhelming focus on diagnosis rather than on identifying high-risk individuals. Jacobi et al. also questioned the external validity of existing screens, in particular their usefulness among diverse populations. They suggested the use of a two-stage procedure that would begin with an initial broad screen with an instrument designed to assess risk status, such as the Weight Concerns Scale (Killen et al., 1993), followed by the administration of a more specific diagnostic instrument, such as the Eating Disorder Examination Questionnaire (Mond, Hay, Rodgers, Owen, & Beumont, 2004; see also Chapters 9 & 40), to those screening positive on the initial measure. Ideally, individuals identified as cases would then be referred for further evaluation and potential treatment. Although Jacobi et al.'s (2004) in-depth review was conducted over a decade ago, in many cases their criticisms are still valid.

## Screening for Cases

### Widely Used Screening Tools

*Eating Attitudes Test (EAT)* Both the original 40-item version of the EAT (Garner & Garfinkel, 1979), as well as the shorter 26-item version (Garner, Olmsted, Bohr, & Garfinkel, 1982), have been extensively used as screening instruments for EDs. Cut-off scores—defined as the score that identifies individuals warranting further evaluation for EDs or subclinical forms—of 30 and 20 have been established for the EAT-40 and EAT-26, respectively (Garner et al., 1982). Early studies of the EAT highlighted its low positive predictive power, particularly for classifying cases of AN (Mann et al., 1983). It was, however, suggested that this failing might stem from changes in diagnostic criteria, as the EAT was developed at a time when eating disorders were not included the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*), and the clinical population included in the initial validation met the Research and Diagnostic Criteria established by Feighner et al. (1972; cited in Spitzer, Endicott, & Robins, 1978), which were very different from those used in *DSM-III* (3rd ed.; American Psychiatric Association, 1980; Mintz & O'Halloran, 2000). In a later validation study of 136 college women using the *DSM-IV* (American Psychiatric Association, 1994) ED criteria, the EAT-40 and EAT-26 demonstrated a sensitivity of 77%, specificity of 95% and 94% respectively, positive predictive power of 82% and 79% respectively, and negative predictive power of 93% and 94% respectively (Mintz & O'Halloran, 2000). These findings suggested that the EAT may in fact be a useful instrument for differentiating between individuals with EDs, regardless of diagnosis, and those without. The EAT has been widely translated and is available in Arabic, Brazilian, Chinese, French, Greek, Italian, Japanese, Portuguese, Spanish, and Turkish.

*Eating Disorder Examination Questionnaire (EDE-Q)* The EDE-Q (Mond et al., 2004) is a 36-item self-report questionnaire derived from the *DSM-IV* (American Psychiatric Association, 1994) criteria for EDs, assessing eating behaviors over the last 28 days. As noted previously, the interview version of the EDE is one of the gold standard instruments often used in the validation of screening instruments (Cooper & Fairburn, 1987). A global score can be calculated as well as four subscales: Restraint, Weight Concerns, Shape Concerns, and Eating

Concerns. An initial analysis identified a cut-off score of 56 on the summed total score (2.3 on the global score). Using this cut-off among a community sample of 208 young women resulted in a sensitivity of 92%, specificity of 86%, and positive predictive value of 30% (Mond et al., 2004). Including two extra criterion items, namely the frequency of objective binge episodes and exercising for weight/shape reasons at least once a week, in addition to the cut-off score resulted in an increase in positive predictive value to 56%. In a later examination of the properties of the EDE-Q among 25 ED cases, analysis suggested 2.8 on the global scale as the optimum cut-off point, resulting in a sensitivity of 96%, specificity of 56%, and positive predictive value of 30% (Mond et al., 2008). Lowering the cut-off score to 1.61, so as to increase sensitivity, resulted in a value of 92% but did not affect the specificity (56%) and positive predictive value (30%).

The usefulness of the EDE-Q in relation to the proposed *DSM-5* (American Psychiatric Association, 2013) criteria (see Chapters 8–11) has also been explored. In a recent study validating the EDE-Q against the EDE-interview version among 217 treatment-seeking individuals, Berg et al. (2012) reported that when using the proposed *DSM-5* criteria, the EDE-Q demonstrated sensitivity ranging from 36.8% to 80.8% according to the diagnosis, specificity ranging from 77.3% to 98%, positive predictive values ranging from 41.2% to 81.5%, and negative predictive values ranging from 76.5% to 98.5%. The EDE-Q performed poorest for binge eating disorder (BED), but reasonably well for AN, bulimia nervosa (BN), and eating disorders not otherwise specified (ED-NOS).

Fairburn, Cooper, Doll, and Davies (2005) explored the longitudinal predictive validity of the EDE-Q among 2,992 young women who screened positive for dieting but did not have a present or past ED. The women were contacted every 6 months over a 2-year period to assess present ED symptomatology. Based on their findings, the authors developed an eight-item screening instrument, which included seven items from the EDE-Q, as well as body mass index (BMI). Using this screening instrument, sensitivity ranged from 69.5% to 79% and specificity from 52% to 76%, depending on the decision tree used, with the final survey therefore displaying good sensitivity and specificity of 79% and 76% respectively. This brief screening instrument therefore appears useful in detecting dieting young women at risk of developing a future ED. The EDE-Q has been translated into many languages including Spanish, Fijian, and Turkish.

*Bulimia Test-Revised (BULIT-R)* The BULIT-R (Thelen, Farmer, Wonderlich, & Smith, 1991) is a 28-item self-report measure designed to assess the presence and degree of bulimic symptomatology according to the *DSM-III-R* (American Psychiatric Association, 1987). Using a cut-off score of 104, initial examinations of the BULIT-R in a sample of 23 women with BN versus 153 controls, as well as a second sample of 53 subjects with scores above and below the cut-off, revealed rates of sensitivity of 83% and 62% respectively, and a specificity of 96% (Thelen et al., 1991). A second study explored the properties of the BULIT-R among 243 female college students (Welch, Thompson, & Hall, 1993). Using the original cut-off of 104, the sensitivity of the scale was 80%, the specificity was 99.5%, the positive predictive value was 80%, and the negative predictive value was 99.5%. In order to increase the positive predictive value, the authors proposed a new cut-off score of 98. Using this threshold, the measure's sensitivity was 100%, the positive predictive value was 71.3%, and the negative predictive value was 100%. A later validation study using *DSM-IV* diagnostic criteria for BN and the initial cut-off score of 104 in a sample of 147 women

revealed a sensitivity of 91%, a specificity of 96%, a positive predictive value of 81%, and a negative predictive value of 98% (Thelen, Mintz, & Vander Wal, 1996). Taken together, these data indicate that the BULIT is a useful and valid screen for identifying bulimic symptomatology. The BULIT has been translated into Korean and Spanish.

*SCOFF* The SCOFF (Morgan, Reid, & Lacey, 1999) is a five-item (Yes/No) screening instrument focusing on the core features of AN and BN:

- **S**ick (self-induced vomiting);
- **C**ontrol (loss of control);
- **O**ne stone (weight loss of 14 lb [6.35 kg]);
- **F**at (feelings of fatness);
- **F**ood (preoccupation with food).

In an initial examination among 116 patients with EDs and 96 controls, the threshold of two positive answers to SCOFF questions had 100% sensitivity for detecting AN or BN and an 87.5% specificity (Morgan et al., 1999). Confirmation of the two positives cut-off came from a study comparing the SCOFF and the EDE-Q in a sample of 257 primary care patients, which reported a sensitivity of 72% and specificity of 81% (Mond et al., 2008).

Additional studies have confirmed the high sensitivity and specificity of the SCOFF. An examination of the properties of the written and orally administered SCOFF in a sample of 178 college students revealed a high level of agreement between these two forms in the prediction of possible EDs using the two positives cut-off (Perry et al., 2002). Among 341 female primary care patients, the questionnaire demonstrated a sensitivity of 86.4% and a specificity of 89.6% (Luck et al., 2002). Using the cut-off of two positives, all women confirmed to be suffering from an ED were detected, in addition to 34 false positives.

Other authors, however, have reported lower levels of sensitivity for the SCOFF. In a sample of 305 graduate students, the SCOFF demonstrated a sensitivity of only 53.3% with a specificity of 90.3% (Parker, Lyons, & Bonner, 2005). The resulting positive predictive value was 66.7%, while the proportion of those who were correctly identified as not having an ED by the screen (NPV) was 88.7%. These findings suggested that the SCOFF may not be as sensitive as previously suggested among this population. Similarly, among a sample of 1,891 Finnish adolescents, the agreement between SCOFF ratings and nurse evaluations was low, with only a 20% agreement for students scoring above the two-positives cut-off (Hautala et al., 2009). The authors suggested that among adolescents, when it is preferable to be overinclusive rather than fail to detect an at-risk individual, a cut-off of one positive might be useful. Chinese, Finnish, French, Italian, Japanese, and Spanish versions of the SCOFF are available.

## Less Widely Used Screening Measures

*Bulimic Investigatory Test, Edinburgh (BITE)* The BITE (Henderson & Freeman, 1987) is a 36-item self-report questionnaire that yields symptom (cut-off score = 20) and severity subscale scores (cut-off = 5), as well as a total score (cut-off = 25). The validity of the BITE was assessed in a sample of 55 ED patients and 27 control women (Waller, 1992). Findings revealed that the BITE performed well in identifying restricting AN cases as well as individuals suffering from BN with and without history of AN; however, the BITE performed poorly (false negative

rate = 60%) in the identification of AN/bulimic subtype. The author cautioned against the use of the BITE for screening, as it appeared to lack sensitivity, in particular regarding cases of normal-weight BN. A number of translations of the BITE exist, including Italian and a Portuguese version developed among a Brazilian sample.

*Eating Disorder Diagnostic Scale (EDDS)* The EDDS (Stice, Telch, & Rizvi, 2000) is a 22-item self-report measure designed to assess the presence of the *DSM-IV* (American Psychiatric Association, 1994) criteria for AN, BN, and BED. The EDDS was validated against the EDE in a sample of 367 women aged 13 to 61. For the three disorders, the EDDS demonstrated a sensitivity ranging from 77% to 93%, a specificity ranging from 96% to 100%, positive predictive values ranging from 80% to 93%, and negative predictive values ranging from 95% to 100%. Similarly, strong validity was found in a second study in which the EDDS was again examined against the EDE in a sample of 443 female adolescents and young adults, yielding a sensitivity of 88%, specificity of 98%, positive predictive value of 74%, and negative predictive value of 98% (Stice, Fisher, & Martinez, 2004). The validity of the Dutch version of the EDDS was also evaluated in a sample of 40 ED patients and 45 control undergraduate students (Krabbenborg et al., 2012). The EDE was again used as the gold standard. The EDDS demonstrated high sensitivity (>95%) and positive predictive value among patients suffering from AN and BN, but lower sensitivity (57%) and positive predictive value (80%) among patients suffering from BED. The authors further identified a cut-off point of 16.5 to distinguish patients from healthy controls. Together these findings suggest that the EDDS is a very useful screening tool. It is unclear why it is not widely used.

*Questionnaire for Eating Disorders Diagnosis (Q-EDD)* The Q-EDD (Mintz, O'Halloran, Mulholland, & Schneider, 1997) is a 50-item self-report questionnaire, based on the *DSM-IV* (American Psychiatric Association, 1994) diagnostic criteria, that aims to evaluate the presence of AN, BN, and four types of ED-NOS (subthreshold bulimia, menstruating anorexia, nonbingeing bulimia, and binge eating disorder). The Q-EDD was validated against a clinical interview and demonstrated a sensitivity of 97%, specificity of 98%, positive predictive value of 94%, and negative predictive value of 99% in distinguishing between eating-disordered individuals all diagnoses combined, and non-eating-disordered individuals. The Q-EDD has been used in large screening studies among young women (Franko et al., 2005).

## Binge Eating Disorder: The New Addition to *DSM-5*

### BULIT

The BULIT has been suggested to be well suited to the evaluation of binge eating as it contains items assessing compensatory behaviors (Vander Wal, Stein, & Blashill, 2011). In a sample of 15 individuals suffering from BED and 26 assessed as having neither full-syndrome nor subclinical BED, using a cut-off score of 80, the BULIT demonstrated a sensitivity of 100%, specificity of 96%, positive predictive value of 94%, and negative predictive value of 100%. Although provided by a single study, these findings confirm that the BULIT can successfully be used to classify individuals suffering from BED.

### The Patient Health Questionnaire Module for Eating Disorders (PHQ-ED)

The PHQ-ED (Spitzer, Kroenke, & Williams, 1999) is composed of six binary (yes/no) items assessing binge eating and compensatory behaviors, and up to two optional follow-up questions that are asked only when binge eating or purging has been endorsed. A recent study examined the properties of the PHQ-ED as a screening instrument in a sample of 348 adults that included 259 cases of BED, BN, or Recurrent Binge Eating (Striegel-Moore et al., 2010). The PHQ-ED showed excellent sensitivity of 100% across all three diagnoses, and good specificity of 91.7% for BED and BN, and 92.4% for Recurrent Binge Eating. However, the positive predictive value was fairly low, with 10.4% for BED and BN, and 18.5% for Recurrent Binge Eating. The authors suggested that the low positive predictive value may have been due to the study sample, which was drawn from a community population with a low prevalence of BED. Clinicians should be cautious in inferring the likely presence of a disorder from a positive screen, particularly when using the PHQ-ED in community samples.

### Bulimic Investigation Test, Edinburgh (BITE)

The BITE, as described above, has also been explored in the context of screening for BED. The BITE's psychometric properties were examined in a sample of 360 consecutive patients attending an outpatient clinic, using both a cut-off of 10 and a cut-off of 20 (Ricca et al., 2000). With the lower cut-off, the BITE displayed a sensitivity of 91%, specificity of 51.4%, positive predictive value of 71.8%, and negative predictive value of 98.2%. Using the higher cut-off of 20, the sensitivity decreased to 33.3%, while the specificity increased to 92.0%, with a positive predictive value of 30.5% and a negative predictive value of 92.9%.

### Binge Eating Scale (BES)

The BES (Gormally, Black, Daston, & Rardin, 1982) is a 16-item self-report questionnaire, scored on a 0–3 scale. A cut-off score of 17 has been established as indicating the presence of binge eating. In a sample of 125 women attending a weight-loss program, only 51.8% of those identified as binge eaters using the BES cut-off score of 17 had a concordant diagnosis using the EDE (Greeno, Markus, & Wing, 1995). In a recent study among 473 consecutive patients presenting for psychological evaluation prior to gastric bypass surgery, using the conventional cut-off score of 17, the BES displayed a sensitivity of 94%, specificity of 76%, positive predictive value of 37%, and negative predictive value of 99% (Grupski et al., 2013). To increase the positive predictive value, the authors conducted the analysis using a higher cut-off score of 27. With this new threshold, the BES displayed a much lower sensitivity of 37%, a specificity of 96%, a positive predictive value of 56%, and a negative predictive value of 91%. An analysis confirmed 17 as the optimal cut-off score in order to maximize both sensitivity and specificity.

### Questionnaire for Eating and Weight Patterns (QEWP)

The QEWP (Spitzer et al., 1992) is a 13-item self-report questionnaire assessing the presence of BED according to the *DSM-IV* (American Psychiatric Association, 1994) criteria. Data regarding the usefulness of the QEWP as a screening instrument have been mixed. In a sample of 157 adults screened for a BED treatment study, the QEWP demonstrated a sensitivity of 74% and a specificity of 35%. Using the Structured Clinical Interview for DSM-IV (SCID) as

a comparison (First, Spitzer, Williams, & Gibbon, 1995), the QEWP demonstrated a sensitivity of 55%, specificity of 80%, and positive predictive value of 79%. Further research is necessary to clarify the validity of the QEWP.

## Overview of Tools for Screening for High-Risk Individuals

### The McKnight Investigators

The McKnight Investigators (2003) conducted a 4-year longitudinal study of 1,358 high-school students. The McKnight Risk Factor Survey IV includes seven different factors: thin-body preoccupation and social pressure; substance use; parental influences; general psychological influences; social support; number of negative life-events; and school performance. The items from the thin body preoccupation and social pressure factors were examined for their usefulness as a potential screen for ED cases. With the aim of deriving the most parsimonious screening tool, three criteria were derived from four questions:

1 A score of 4 or 5 in response to "How often have you been on a diet to lose weight?"
2 A score of 5 in response to "How often does your weight make boys not like you?"
3 An average score ≥3.5 in response to the two questions: "How often do you change your eating around boys?" and "How often do you change your eating around girls?"

When considering participants who scored above the cut-off for any of these three questions, the screen demonstrated a sensitivity of 72% and a specificity of 80%. The predictive value of a positive test was 10% (of a base rate of cases of 2.9%), and the predictive value of a negative test was 99% (of a base rate of non-cases of 97.1%).

### Weight Concerns Scale

Using data from the first wave of a large longitudinal study of 967 sixth- and seventh-grade girls, Killen et al. (1993) developed a five-item Weight Concerns measure assessing worry over weight and body shape, fears of gaining weight, importance of weight, perceived fatness, and diet history. Their initial analysis revealed that, compared to SCID diagnoses (First et al., 1995), a score of 57 on this scale yielded a sensitivity of 86% and specificity of 63%. Their measure demonstrated excellent 7-month stability and longitudinal validity, as scores were significantly associated with the onset of ED symptoms 3 years later (Killen et al., 1994). The Weight Concerns Scale is one of the few screening instruments available to identify high-risk individuals rather than cases and is supported by evidence for longitudinal validity.

### Stanford-Washington Eating Disorder Screen (SWED)

The SWED (Wilfley, Agras, & Taylor, 2013) contains 11 items taken from the Weight Concerns Scale, the EDE-Q, the EDDS, and the Clinical Impairment Assessment (Bohn et al., 2008). The scale properties have been examined among college women compared to the SCID and the EDE-Q, and the SWED demonstrated high sensitivity and specificity (>90%) for AN and BN when using *DSM-IV-TR* (American Psychiatric Association, 2000) and *DSM-5* criteria (Wilfley et al., 2013). Further investigations of the criterion validity of the instrument using *DSM-5* (2013) criteria are underway, and the SWED appears to be a very promising instrument for screening for ED risk.

## Specific Populations: Children, Adolescents, Athletes, and Men

### Children's Eating Attitudes Test (ChEAT)

The ChEAT is a modified version of the EAT (see above) with language adapted for children (Vacc & Rhyne, 1987). However, data on the properties of the ChEAT as a screening instrument are somewhat lacking. One study conducted among 152 preadolescents aged from 8 to 12, testing a variety of potential cut-offs, suggested 25 as an optimal cut-off, but concluded that the sensitivity of the ChEAT was overall somewhat low and varied with age (Erickson & Gerstle, 2007). Using the threshold of 25, overall sensitivity was 36%; however, this increased to 53% among girls aged between 10 and 12 years. In a sample of 409 girls aged 9–13 years, compared to the child version of the EDE using the cut-off score of 20, the ChEAT demonstrated a sensitivity of 17%, specificity of 98%, positive predictive value of 63%, and negative predictive value of 87% (Colton, Olmsted, & Rodin, 2007). Decreasing the cut-off score improved the sensitivity, which reached 76% with a cut-off of 5; however, the specificity then fell to 59%. Taken together these findings suggest that the usefulness of the ChEAT as a screening instrument among children may be limited. Translations of the ChEAT exist in Croatian, Portuguese, and Spanish.

### Children's Binge Eating Disorder Scale (C-BEDS)

The C-BEDS (Shapiro et al., 2007) is a seven-item self-report questionnaire designed to assess the core behaviors of BED as defined by Marcus and Kalarchian (2003), including eating in the absence of hunger, eating in response to strong emotions or external cues, and eating in secret. Among a sample of 55 children aged 5–13 years, using the SCID as a gold standard, the C-BEDS demonstrated a sensitivity of 71% and specificity of 89%. The authors concluded that the C-BEDS may be more useful than the SCID for screening for BED in children but additional evaluations in larger and more diverse samples are warranted.

### Athletic Milieu Direct Questionnaire (AMDQ)

Athletes have been reported to be two to three times more at risk for EDs compared to non-athletes (Bratland-Sanda & Sundgot-Borgen, 2013; see also Chapter 35) and may have specific risk factors (Torstveit, Rosenvinge, & Sundgot-Borgen, 2008). The AMDQ, a screening instrument for EDs specifically tailored to athletes, was developed and tested among 149 college athletes (Nagel, Black, Leverenz, & Coster, 2000). The initial pool of 119 items was pared down to three statistical subsets including 35, 19, and 9 items, respectively. These three subsets yielded sensitivity values ranging from 80% to 82%, specificity values ranging from 75.27% to 79.57%, positive predictive values ranging from 63.49% to 68.33%, and negative predictive values ranging from 87.50% to 89.16%, indicating good psychometric properties. Thus, the AMDQ may prove a highly useful tool among athletes.

### Eating Disorder Assessment for Men (EDAM)

EDs are disproportionately diagnosed among women, and the large majority of the assessment instruments were developed and normed among female populations (see Chapter 37). The EDAM (Stanford & Lemberg, 2012b), a new instrument for assessing EDs in men, includes 50 self-report items. The initial factor analysis revealed four underlying factors: BED, Muscle Dysmorphic Disorder, Body Dissatisfaction, and Disordered Eating. In a subsequent study, a

sample of 108 individuals (78% male) completed the EDAM and the Eating Disorder Inventory-3 (EDI-3; Stanford & Lemberg, 2012a). Criterion validity was not formally assessed; however, using logistic regression the authors concluded that the EDAM successfully predicted ED status above and beyond the EDI-3. Further research on the usefulness of the EDAM as a screening tool among men is warranted.

## The Importance of Setting

### Primary Care

General practitioners are in an ideal position to identify EDs that might otherwise go undetected (Mond et al., 2008). However, EDs often are overlooked and underdiagnosed in primary care practice, in part due to lack of time and the impact of nonclinical demographic variables such as ethnicity and gender on the likelihood of being diagnosed (Currin, Schmidt, & Waller, 2007; see also Chapter 25). One study found that screening was relatively acceptable, with approximately one half of the female primary-care patients aged 16–35 years agreeing to be screened (Johnston, Fornai, Cabrini, & Kendrick, 2007). However, general practitioners were found to be uncertain of the procedure to follow in the case of a positive screen, and referrals were rarely made. These data suggest that, while primary care settings provide a highly useful context for screening for EDs, there is a need for education surrounding EDs among general practitioners.

In a study exploring the comparative usefulness of the SCOFF versus the EDE-Q in general practice settings, Mond et al. (2008) concluded that the two instruments had relative merits according to the situation. In the context of routine screening, the brevity and ease of the SCOFF might make it more suitable; however, when seeking to confirm suspicion of an ED diagnosis, the EDE-Q could be preferable due to its higher specificity and positive predictive value, as well as the more detailed information provided, which may guide referral. Cotton, Ball, and Robinson (2003) compared the usefulness, in general practice, of the four-question Eating Disorder Screen for Primary Care (ESP) and the SCOFF. They found that, using a cut-off of two positive responses, the sensitivity and specificity of the SCOFF were 78% and 88% respectively, while those of the ESP were 100% and 71%. The authors concluded that, while both instruments were useful in detecting cases, the ESP performed somewhat better at ruling out potential cases.

### Schools

Schools and college campuses provide an ideal setting for screening for EDs, due to the practical advantages of having youths concentrated in one place and the on-site presence of support staff (Austin et al., 2008). The first National Eating Disorders Screening Program (NEDSP) was launched on U.S. college campuses in 1996 as a confidential volunteer two-stage screening program for students (Becker, Franko, Speck, & Herzog, 2003). The first stage included a written self-report questionnaire, which was followed by a face-to-face interview with a counselor to review responses. In the 1996 wave of NEDSP, 9,069 students from 409 colleges returned questionnaires, and 5,787 of these (64%) met with a counselor to discuss their responses. Counselors were provided with a manual for identifying participants reporting potentially clinically significant symptoms (purging or binge eating at least once a month or significant distress), and were instructed to then make an appropriate referral for further evaluation and treatment, with three possible outcomes: no referral at this time, professional evaluation, and professional evaluation-urgent. Of the 5,787 program participants who met with a counselor,

4,398 (76%) met the criteria for clinically significant symptoms and were therefore referred for professional evaluation, including 71 (1.2%) who were referred for urgent evaluation.

Subsequently NEDSP was successfully extended to high schools, and in 2000 over 35,000 high-school students from 152 schools completed a NEDSP screening measure, including attitudinal and behavioral questions and the EAT-26 (Austin et al., 2008). The NEDSP screening measure was compared against the EAT within each gender separately (Haines et al., 2011). Among girls, the combination of the attitudinal and behavioral items performed well, with sensitivity ranging from 88% to 91%, specificity ranging from 73% to 82%, positive predictive values ranging from 36% to 45%, and negative predictive value = 98%. Among boys, the combination of attitudinal and behavioral items also performed well, although the positive predictive values were somewhat low: sensitivity ranging from 88% to 91%, specificity ranging from 87% to 94%, positive predictive values ranging from 22% to 36%, and negative predictive value = 99%. The authors reiterated the recommendations of Jacobi et al. (2004) that screening take place in two phases, with an initial screening (e.g., using the NEDSP screen) followed by a more specific instrument.

## Cultural and Ethnic Considerations

Rates of EDs have been found to vary according to ethnicity and cultural context (Franko, Becker, Thomas, & Herzog, 2007; see also Chapters 23 & 25); however, to date, there has been little focus on the impact of culture or race and ethnicity in screening for EDs. As previously noted, the overwhelming majority of existing assessment instruments for EDs were developed among White English-speaking young women. While many of these scales have been translated and validated in different languages and different cultural contexts, it should not be assumed that they perform identically across cultural contexts (see Chapter 16). For example, Mumford, Whitehouse, and Choudry (1992) validated the EAT among a sample of 369 schoolgirls in Pakistan and reported that, despite a factor structure almost identical to that of the original instrument, the optimal cut-off score might be lower among this population. Similarly, a validation of the EAT among a sample of 67 patients suffering from BN and 109 patients suffering from AN in Hong Kong revealed that up to 41.5% of patients with typical AN scored below the traditional cut-off score on the EAT, suggesting that a much lower cut-off score would be appropriate among this population (Lee, Kwok, Liau, & Leung, 2002). Furthermore, the presentation of AN in this sample was somewhat different to that usually seen in Western countries, with a high proportion of patients described as non-fat phobic, who attributed their restriction to reasons other than fear of fatness. These considerations are not well assessed by the items of the EAT, which focus strongly on fear of fatness as a core dimension of AN. Taken together, these studies suggest that existing screening measures for EDs may indeed be ethnocentric, and should only be transposed to other cultural contexts with caution.

Regarding race and ethnicity, findings from the high-school NEDSP in 2000 documented few ethnic differences in the prevalence of EDs (Austin et al., 2008). The authors concluded that screening for EDs was necessary in schools with and without ethnic diversity and argued that targeting specific ethnic groups for screening would be inappropriate. The necessity of screening among all ethnic groups is particularly important in view of findings that Native Americans and Hispanic students were less likely to be referred than White students with comparable scores in the national college screening study (Becker et al., 2003).

## Conclusions and Future Directions

Screening for EDs has a number of benefits for the mental health of individuals suffering from or at risk for EDs. While a large number of screening instruments have been developed to date, many of them still lack evidence of validity, both criterion and external. In addition, screening instruments tailored to particular groups such as athletes or males are lacking. Overall, the last decade has seen little effort to improve, refine, or seek to better establish the validity of existing screening instruments. In view of these shortcomings, the field would benefit from increased consensus regarding the use of screening instruments so as to facilitate the examination of psychometric properties as well as to conduct comparisons in terms of prevalence.

The success of programs such as NEDSP also highlights the usefulness of routine screening in school and college settings. However, widespread screening would give rise to a number of challenges, including the processing of the screening forms, and the allocation of necessary resources for the follow-up of potential cases for referral and treatment where necessary. The ethical implications of widespread screening in the absence of sufficient resources downstream are numerous.

In conclusion, screening offers a valuable opportunity for the early identification of EDs; however, efforts to further explore the psychometric properties of existing instruments and develop appropriate measures for specific vulnerable populations are warranted. In addition, greater efforts should be made to screen for EDs among youth in school and college settings.

## References

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.; DSM-III). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.; DSM-III-R). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.; DSM-IV). Washington, DC: Author.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.; text rev.; DSM-IV-TR). Washington, DC: Author.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.; DSM-5). Washington, DC: Author.

Austin, S. B., Ziyadeh, N. J., Forman, S., Prokop, L. A., Keliher, A., & Jacobs, D. (2008). Screening high school students for eating disorders: Results of a national initiative. *Preventing Chronic Disease*, *5*(4). Retrieved from http://www.cdc.gov/pcd/issues/2008/oct/07_0164.htm

Becker, A.E., Franko, D.L., Nussbaum, K., & Herzog, D.B. (2004). Secondary prevention for eating disorders. *International Journal of Eating Disorders*, *36*, 157–162.

Becker, A. E., Franko, D. L., Speck, A., & Herzog, D. B. (2003). Ethnicity and differential access to care for eating disorder symptoms. *International Journal of Eating Disorders*, *33*, 205–212.

Berg, K. C., Stiles-Shields, E. C., Swanson, S. A., Peterson, C. B., Lebow, J., & Le Grange, D. (2012). Diagnostic concordance of the interview and questionnaire versions of the eating disorder examination. *International Journal of Eating Disorders*, *45*, 850–855.

Bohn, K., Doll, H. A., Cooper, Z., O'Connor, M., Palmer, R. L., & Fairburn, C. G. (2008). The measurement of impairment due to eating disorder psychopathology. *Behaviour Research and Therapy*, *46*, 1105–1110.

Bratland-Sanda, S., & Sundgot-Borgen, J. (2013). Eating disorders in athletes: Overview of prevalence, risk factors and recommendations for prevention and treatment. *European Journal of Sport Science*, *13*, 499–508.

Colton, P. A., Olmsted, M. P., & Rodin, G. M. (2007). Eating disturbances in a school population of preteen girls: Assessment and screening. *International Journal of Eating Disorders*, *40*, 435–440.

Cooper, Z., & Fairburn, C. (1987). The eating disorder examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders*, *6*, 1–8.

Cotton, M.A., Ball, C., & Robinson, P. (2003). Four simple questions can help screen for eating disorders. *Journal of General Internal Medicine*, *18*, 53–56.

Currin, L., Schmidt, U., & Waller, G. (2007). Variables that influence diagnosis and treatment of the eating disorders within primary care settings: A vignette study. *International Journal of Eating Disorders*, *40*, 257–262.

Erickson, S. J., & Gerstle, M. (2007). Developmental considerations in measuring children's disordered eating attitudes and behaviors. *Eating Behaviors*, *8*, 224–235.

Fairburn, C. G., Cooper, Z., Doll, H. A., & Davies, B. A. (2005). Identifying dieters who will develop an eating disorder: a prospective, population-based study. *American Journal of Psychiatry*, *162*, 2249–2255.

First, M., Spitzer, R., Williams, J., & Gibbon, M. (1995). *Structured clinical interview for DSM-IV (SCID)*. Washington, DC: American Psychiatric Association.

Franko, D. L., Becker, A. E., Thomas, J. J., & Herzog, D. B. (2007). Cross-ethnic differences in eating disorder symptoms and related distress. *International Journal of Eating Disorders*, *40*, 156–164.

Franko, D. L., Mintz, L. B., Villapiano, M., Green, T. C., Mainelli, D., Folensbee, L.,…Budman, S. H. (2005). Food, Mood, and Attitude: Reducing risk for eating disorders in college women. *Health Psychology*, *24*, 567–578.

Garner, D. M., & Garfinkel, P. E. (1979). The Eating Attitudes Test: An index of the symptoms of anorexia nervosa. *Psychological Medicine*, *9*, 273–279.

Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The Eating Attitudes Test: Psychometric features and clinical correlates. *Psychological Medicine*, *12*, 871–878.

Gormally, J., Black, S., Daston, S., & Rardin, D. (1982). The assessment of binge eating severity among obese persons. *Addictive Behaviors*, *7*, 47–55.

Greeno, C. G., Markus, M. D., & Wing, R. R. (1995). Diagnosis of binge eating disorder: Discrepancies between a questionnaire and clinical interview. *International Journal of Eating Disorders*, *17*, 153–160.

Grupski, A. E., Hood, M. M., Hall, B. J., Azarbad, L., Fitzpatrick, S. L., & Corsica, J. A. (2013). Examining the Binge Eating Scale in screening for binge eating disorder in bariatric surgery candidates. *Obesity Surgery*, *23*, 1–6.

Haines, J., Ziyadeh, N. J., Franko, D. L., McDonald, J., Mond, J. M., & Austin, S. B. (2011). Screening high school students for eating disorders: Validity of brief behavioral and attitudinal measures. *Journal of School Health*, *81*, 530–535.

Hautala, L., Junnila, J., Alin, J., Grönroos, M., Maunula, A.-M., Karukivi, M.,…Saarijärvi, S. (2009). Uncovering hidden eating disorders using the SCOFF questionnaire: Cross-sectional survey of adolescents and comparison with nurse assessments. *International Journal of Nursing Studies*, *46*, 1439–1447.

Henderson, M., & Freeman, C. P. (1987). A self-rating scale for bulimia. The "BITE". *British Journal of Psychiatry*, *150*, 18–24.

Jacobi, C., Abascal, L., & Taylor, C. B. (2004). Screening for eating disorders and high-risk behavior: Caution. *International Journal of Eating Disorders*, *36*, 280–295.

Johnston, O., Fornai, G., Cabrini, S., & Kendrick, T. (2007). Feasibility and acceptability of screening for eating disorders in primary care. *Family Practice*, *24*, 511–517.

Killen, J. D., Taylor, C. B., Hammer, L. D., Litt, I., Wilson, D. M.,…Varady, A. (1993). An attempt to modify unhealthful eating attitudes and weight regulation practices of young adolescent girls. *International Journal of Eating Disorders*, *13*, 369–384.

Killen, J. D., Taylor, C. B., Hayward, C., Wilson, D. M., Haydel, K. F., Hammer, L. D.,…Varady, A. (1994). Pursuit of thinness and onset of eating disorder symptoms in a community sample

of adolescent girls: A three-year prospective analysis. *International Journal of Eating Disorders*, *16*, 227–238.

Krabbenborg, M. A., Danner, U. N., Larsen, J. K., Veer, N., Elburg, A. A., Ridder, D. T.,…Engels, R. C. (2012). The Eating Disorder Diagnostic Scale: Psychometric features within a clinical population and a cut-off point to differentiate clinical patients from healthy controls. *European Eating Disorders Review*, *20*, 315–320.

Lee, S., Kwok, K., Liau, C., & Leung, T. (2002). Screening Chinese patients with eating disorders using the Eating Attitudes Test in Hong Kong. *International Journal of Eating Disorders*, *32*, 91–97.

Luck, A. J., Morgan, J. F., Reid, F., O'Brien, A., Brunton, J., Price, C.,…Lacey, J. H. (2002). The SCOFF questionnaire and clinical interview for eating disorders in general practice: Comparative study. *BMJ: British Medical Journal*, *325*, 755–756.

Mann, A., Wakeling, A., Wood, K., Monck, E., Dobbs, R., & Szmukler, G. (1983). Screening for abnormal eating attitudes and psychiatric morbidity in an unselected population of 15-year-old schoolgirls. *Psychological Medicine*, *13*, 573–580.

Marcus, M. D., & Kalarchian, M. A. (2003). Binge eating in children and adolescents. *International Journal of Eating Disorders*, *34*, S47–S57.

Mintz, L. B., & O'Halloran, M. S. (2000). The Eating Attitudes Test: Validation with DSM-IV eating disorder criteria. *Journal of Personality Assessment*, *74*, 489–503.

Mintz, L. B., O'Halloran, M. S., Mulholland, A. M., & Schneider, P. A. (1997). Questionnaire for Eating Disorder Diagnoses: Reliability and validity of operationalizing DSM-IV criteria into a self-report format. *Journal of Counseling Psychology*, *44*, 63–79.

Mond, J. M., Hay, P., Rodgers, B., Owen, C., & Beumont, P. (2004). Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour Research and Therapy*, *42*, 551–567.

Mond, J. M., Myers, T. C., Crosby, R. D., Hay, P. J., Rodgers, B., Morgan, J. F.,…Mitchell, J. E. (2008). Screening for eating disorders in primary care: EDE-Q versus SCOFF. *Behaviour Research and Therapy*, *46*, 612–622.

Morgan, J. F., Reid, F., & Lacey, J. H. (1999). The SCOFF questionnaire: Assessment of a new screening tool for eating disorders. *BMJ: British Medical Journal*, *319*, 1467–1468.

Mumford, D. B., Whitehouse, A. M., & Choudry, I. Y. (1992). Survey of eating disorders in English-medium schools in Lahore, Pakistan. *International Journal of Eating Disorders*, *11*, 173–184.

Nagel, D. L., Black, D. R., Leverenz, L. J., & Coster, D. C. (2000). Evaluation of a screening test for female college athletes with eating disorders and disordered eating. *Journal of Athletic Training*, *35*, 431–440.

Parker, S. C., Lyons, J., & Bonner, J. (2005). Eating disorders in graduate students: Exploring the SCOFF questionnaire as a simple screening tool. *Journal of American College Health*, *54*, 103–107.

Perry, L., Morgan, J., Reid, F., Brunton, J., O'Brien, A., Luck, A., & Lacey, H. (2002). Screening for symptoms of eating disorders: Reliability of the SCOFF screening tool with written compared to oral delivery. *International Journal of Eating Disorders*, *32*, 466–472.

Ricca, V., Mannucci, E., Moretti, S., Di Bernardo, M., Zucchi, T., Cabras, P. L., & Rotella, C. M. (2000). Screening for binge eating disorder in obese outpatients. *Comprehensive Psychiatry*, *41*, 111–115.

Shapiro, J. R., Woolson, S. L., Hamer, R. M., Kalarchian, M. A., Marcus, M. D., & Bulik, C. M. (2007). Evaluating binge eating disorder in children: Development of the children's binge eating disorder scale (C-BEDS). *International Journal of Eating Disorders*, *40*, 82–89.

Spitzer, R. L., Devlin, M., Walsh, B. T., Hasin, D., Wing, R., Marcus, M.,…Nonas, C. (1992). Binge eating disorder: A multisite field trial of the diagnostic criteria. *International Journal of Eating Disorders*, *11*, 191–203.

Spitzer, R. L., Endicott, J., & Robins, E. (1978). Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry*, *35*, 773–782.

Spitzer, R. L., Kroenke, K., & Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD. *Journal of the American Medical Association*, *282*, 1737–1744.

Stanford, S. C., & Lemberg, R. (2012a). A clinical comparison of men and women on the Eating Disorder Inventory-3 (EDI-3) and the Eating Disorder Assessment for Men (EDAM). *Eating Disorders: The Journal of Treatment and Prevention*, *20*, 379–394.

Stanford, S. C., & Lemberg, R. (2012b). Measuring eating disorders in men: Development of the Eating Disorder Assessment for Men (EDAM). *Eating Disorders: The Journal of Treatment and Prevention*, *20*, 427–436.

Stice, E., Fisher, M., & Martinez, E. (2004). Eating Disorder Diagnostic Scale: Additional evidence of reliability and validity. *Psychological Assessment*, *16*, 60–71.

Stice, E., Ng, J., & Shaw, H. (2010). Risk factors and prodromal eating pathology. *Journal of Child Psychology and Psychiatry*, *51*, 518–525.

Stice, E., Telch, C. F., & Rizvi, S. L. (2000). Development and validation of the Eating Disorder Diagnostic Scale: A brief self-report measure of anorexia, bulimia, and binge-eating disorder. *Psychological Assessment*, *12*, 123–131.

Striegel-Moore, R. H., Perrin, N., DeBar, L., Wilson, G. T., Rosselli, F., & Kraemer, H. C. (2010). Screening for binge eating disorders using the Patient Health Questionnaire in a community sample. *International Journal of Eating Disorders*, *43*, 337–343.

Thelen, M. H., Farmer, J., Wonderlich, S., & Smith, M. (1991). A revision of the Bulimia Test: The BULIT-R. *Psychological Assessment*, *3*, 119–124.

Thelen, M. H., Mintz, L. B., & Vander Wal, J. S. (1996). The Bulimia Test-Revised: Validation with DSM-IV criteria for bulimia nervosa. *Psychological Assessment*, *8*, 219–221.

The McKnight Investigators. (2003). Risk factors for the onset of eating disorders in adolescent girls: Results of the McKnight longitudinal risk factor study. *American Journal of Psychiatry*, *160*, 248–254.

Torstveit, M. K., Rosenvinge, J. H., & Sundgot-Borgen, J. (2008). Prevalence of eating disorders and the predictive power of risk models in female elite athletes: A controlled study. *Scandinavian Journal of Medicine and Science in Sports*, *18*, 108–118.

Vacc, N. A., & Rhyne, M. (1987). The Eating Attitudes Test: Development of an adapted language form for children. *Perceptual and Motor Skills*, *65*, 335–336.

Vander Wal, J. S., Stein, R. I., & Blashill, A. J. (2011). The EDE-Q, BULIT-R, and BEDT as self-report measures of binge eating disorder. *Eating Behaviors*, *12*, 267–271.

Waller, G. (1992). Bulimic attitudes in different eating disorders: Clinical utility of the BITE. *International Journal of Eating Disorders*, *11*, 73–78.

Welch, G., Thompson, L., & Hall, A. (1993). The BULIT-R: Its reliability and clinical validity as a screening tool for DSM-III-R bulimia nervosa in a female tertiary education population. *International Journal of Eating Disorders*, *14*, 95–105.

Wilfley, D. E., Agras, W. S., & Taylor, C. B. (2013). Reducing the burden of eating disorders: A model for population-based prevention and treatment for university and college campuses. *International Journal of Eating Disorders*, *46*, 529–532.

Williams, P., Hand, D., & Tarnopolsky, A. (1982). The problem of screening for uncommon disorders—a comment on the Eating Attitudes Test. *Psychological Medicine*, *12*, 431–434.