Reconsidering the use of cut-off scores for the Eating Disorder Examination–Questionnaire

Adrian Meule
Department of Psychiatry and Psychotherapy, University Hospital, LMU Munich, Munich, Germany; Schoen Clinic Roseneck, Prien am Chiemsee, Germany

ABSTRACT
The Eating Disorder Examination–Questionnaire (EDE–Q) is one of the most widely used self-report measures for the assessment of eating disorder psychopathology. Numerous studies have provided norm data for different populations and suggested possible cut-off values for the EDE–Q global score that may indicate the presence of an eating disorder. This commentary argues against the unquestioned use of such cut-off scores as their application may often be unnecessary, disadvantageous, or inappropriate.

Clinical Implications
- Proposed cut-off scores of the EDE–Q should not be taken for granted.
- Relatively low scores between 2 and 3 may have low specificity in certain populations, e.g. obese persons.
- Relatively high scores (e.g., 4) may have low sensitivity in various demographic groups that are underrepresented in eating disorder research.
- Other assessment methods should be used when a clear eating disorder diagnostic score is needed.

The Eating Disorder Examination–Questionnaire (EDE–Q) is one of the most widely used self-report measures for the assessment of eating disorder psychopathology (Berg, Peterson, Frazier, & Crow, 2012). It consists of 28 items, six of which ask for the number of binge eating episodes and inappropriate compensatory behaviors in the past 28 days. The other 22 items assess restrained eating, eating concern, weight concern, and shape concern on a 7-point scale with different response labels (e.g., 0 = no days to 6 = every day). Besides the calculation of subscale scores, a total mean score of these 22 items is often used (which can, thus, range between 0 and 6), with higher scores indicating higher eating pathology.

The EDE–Q is not a diagnostic measure, that is, scoring of the scale does not allow for a clear decision whether an individual has an eating
disorder or not. Nevertheless, several cut-off scores have been proposed that indicate the likely presence of an eating disorder. These suggestions include a score of 1.7 (for males; Schaefer et al., 2018), scores between 2 and 3 (Mond et al., 2008; Rø, Reas, & Stedal, 2015), or a score of 4 (e.g., Lavender, De Young, & Anderson, 2010; Luce, Crowther, & Pole, 2008). A score that figures prominently in the literature is 2.3 (Mond, Hay, Rodgers, Owen, & Beumont, 2004). This score is often cited as an established, empirically derived cut-off score for discriminating between individuals with and without eating disorders (e.g., Brockmeyer, Grosse Holtforth, Bents, Herzog, & Friederich, 2013; Penelo, Negrete, Portell, & Raich, 2013; Ruwaard et al., 2013). Yet, there are several problems with this.

First, many researchers ignore the fact that Mond et al. (2004) did not propose a score of 2.3 to differentiate between individuals with and without an eating disorder per se, but this score was only able to discriminate well between groups in conjunction with at least one objective binge eating episode or excessive use of exercise in the past 28 days. Second, the study by Mond et al. (2004) included only a small number of individuals with eating disorders ($n = 13$). Analyses with such a low base rate can yield very high sensitivity and specificity values, although the measure may actually perform subpar for classifying individuals correctly (Rosenfeld, Sands, & Gorp, 2000). Third, a cut-off score of 2.3 may have high sensitivity to detect cases of eating disorders in none-obese persons, but it may have low specificity in certain populations such as persons with obesity. For example, while obese individuals with regular binge eating have higher EDE–Q scores than obese individuals without regular binge eating, those without regular binge eating usually still score higher than 2.3 on the EDE–Q (e.g., Elder et al., 2006; Müller et al., 2012). This can be explained by the fact that the EDE–Q assesses eating restraint, eating concern, weight concern, and shape concern, which are understandably elevated in obese individuals even if they do not have an eating disorder. Thus, as obese persons without an eating disorder have higher mean scores than 2.3, this cut-off seems to be too low in this population.

The latter concern is also part of a more general issue when it comes to applying cut-off scores for questionnaire measures: they may only be appropriate for the type of population that the original study was based on. For example, a score of 4 that is sometimes applied (e.g., Lavender et al., 2010; Luce et al., 2008) stems from a study by Mond, Hay, Rodgers, and Owen (2006), in which a global EDE–Q score of 4 represented the 95th percentile of the distribution of scores in a sample of young adult women in Australia. Thus, this value may be inappropriate for men and younger or older women, let alone possible cultural differences when using these norm data for people outside of Australia.
Another more general issue pertains to the fact that applying a cut-off value for the EDE–Q global score represents an artificial dichotomization of a continuous variable. For example, a researcher may use a cut-off score to categorize individuals in two groups with high and low EDE–Q scores and then compare these groups on some dependent variable. However, there are many drawbacks to this strategy such as loss of information about individual differences, loss of effect size and power or spurious statistical significance and over-estimation of effect size, the potential to overlook nonlinear relationships, and loss of measurement reliability (MacCallum, Zhang, Preacher, & Rucker, 2002; Maxwell & Delaney, 1993; McClelland, Lynch, Irwin, Spiller, & Fitzsimons, 2015; Rucker, McShane, & Preacher, 2015). Thus, researchers should consider carefully whether it is necessary to divide their sample into groups based on a cut-off score or if using continuous EDE–Q scores in the statistical analyses should be preferred.

In conclusion, numerous studies have proposed certain cut-off scores for the EDE–Q that may discriminate between individuals with and without eating disorders. Depending on the research question, methodological design, or therapeutic use, using such cut-off scores may be of practical value in certain circumstances. When used as a screening tool in clinical settings, for example, they may help to decide who should receive an earlier evaluation and potential signposting into treatment. Or they may be helpful in large-scale epidemiological studies to estimate the prevalence of disordered eating behavior. However, cut-off scores should not be considered established findings that can be applied in any case. For example, using these cut-offs may be particularly problematic in various demographic groups that are often underrepresented in eating disorder research (e.g., males, ethnic minorities, or older individuals). If it is not necessary to dichotomize EDE–Q data, they should rather be used as continuous scores in statistical analyses. If it is necessary to obtain a clear-cut dichotomous score (i.e., eating disorder present vs. absent), more complex algorithms based on the EDE–Q questions should be used instead of relying on the global score (Berg, Stiles-Shields et al., 2012) or other questionnaire- or interview-based instruments (e.g., the Eating Disorder Diagnostic Scale or the Eating Disorder Examination) should be employed.

**ORCID**

Adrian Meule [http://orcid.org/0000-0002-6639-8977](http://orcid.org/0000-0002-6639-8977)

**References**


