
Article:
Tatham, M., Turner, H., Mountford, V.A. et al. (3 more authors) (2015) Development, psychometric properties and preliminary clinical validation of a brief, session-by-session measure of eating disorder cognitions and behaviors: The ED-15. International Journal of Eating Disorders. ISSN 0276-3478

https://doi.org/10.1002/eat.22430

Development, psychometric properties and preliminary clinical validation of a brief, session-by-session measure of eating disorder cognitions and behaviors: The ED-15

Madeleine Tatham, DClinPsy (1)

Hannah Turner, PhD, DClinPsy (2,3)

Victoria A. Mountford, DClinPsy (4,5)

Ashley Tritt (6,7)

Rebecca Dyas (8)

Glenn Waller (8)

1. Norfolk Community Eating Disorders Service, UK
2. Eating Disorders Service, Southern Health NHS Foundation Trust, UK
3. Psychology Academic Unit, University of Southampton, UK
4. Institute of Psychiatry, King's College London, UK
5. South London and Maudsley NHS Foundation Trust, UK
6. McGill University, Canada
7. Project Heal, Canada
8. Department of Psychology, University of Sheffield, UK

Corresponding author: Glenn Waller, Clinical Psychology Unit, Department of Psychology, University of Sheffield, Sheffield S10 2NT, UK. email: g.waller@sheffield.ac.uk

Acknowledgement: A brief version of this paper was presented at the Eating Disorders Research Society meeting, San Diego, 2014

Word count for paper (excluding references, tables, title page and abstract): 5524

Word count for Abstract: 250

Running head: THE ED-15 QUESTIONNAIRE

Development, psychometric properties and preliminary clinical validation of a brief, session-by-session measure of eating disorder cognitions and behaviors: The ED-15

Abstract

Objective: In the treatment research literature on other psychological disorders, there is a move towards session-by-session symptom measurement. The necessary measures need to be brief, focused on core features since the last session, and readily available to clinicians. There is no measure in the eating disorders that meets those criteria. This research reports the development and validation of such a self-report questionnaire.

Method: The authors generated and refined a brief set of attitudinal and behavioral items. The resulting questionnaire (the ED-15) and an existing measure (Eating Disorders Examination-Questionnaire; EDE-Q) were completed by a large non-clinical adult sample ($N = 531$), a group of self-reported eating disorder sufferers ($N = 63$), and a group of women ($N = 33$) diagnosed with bulimia nervosa or atypical bulimia nervosa and undertaking cognitive-behavioral therapy.

Results: Factor analysis identified two scales (Weight & Shape Concerns; Eating Concerns), with strong internal consistency and test-retest reliability. Correlations with the EDE-Q ($r = .889$) indicates that the ED-15 and EDE-Q measure near-identical constructs. The ED-15 differentiated self-reported eating-disordered and non-clinical groups to the same degree as the longer EDE-Q. Session-by-session analysis of the CBT treatment group demonstrated that the different ED-15 scales changed in different patterns across therapy.

Discussion: The ED-15 is not proposed as an alternative to existing measures, but as a complementary tool, used to measure session-by-session change for clinical and research purposes. Future research will track changes in ED-15 scores across therapy, to determine the importance of very early response to therapy and sudden changes.

Keywords: Eating disorders; symptoms; measurement; validation; self-report; questionnaire

Development, psychometric properties and preliminary clinical validation of a brief, session-by-session measure of eating disorder cognitions and behaviors: The ED-15

The evaluation of therapy outcomes is an important task, whether in research or in routine clinical settings. In other fields, there has been a move towards session-by-session symptom monitoring as part of this evaluation process, reflecting an understanding that it is crucial that clinicians and researchers should monitor what happens during the process of therapy. For example, early change can predict remission (1), while 'sudden gains' in therapy have a positive impact on alliance and outcome (2-3). Such session-by-session analysis requires measures that are brief, freely available and psychometrically robust, and which focus on core features and reflect anticipated patterns of change during therapy.

Such measures are well established in depression and the anxiety disorders. For example, the Generalised Anxiety Disorder Assessment (GAD-7 [4]) is a brief self-report questionnaire that can be used as a screening tool and as a measure of severity of anxiety. Similarly, the Patient Health Questionnaire (PHQ-9 [5]) is a brief measure of severity of depression and response to treatment. The utility of those measures is demonstrated by their inclusion as core outcome indices in the UK-wide Improving Access to Psychological Therapies (IAPT) programme. Other measures have been developed to measure specific facets of anxiety, including the Penn State Worry Questionnaire-Short (PSWQ) for pathological worry (6), and the Social Phobia Inventory (SPIN) for social phobia (7). These measures meet the clinical utility and availability criteria outlined above. Such questionnaires can provide ready evidence of progress in therapy (or lack of change, or even deterioration), guiding clinicians on the targets for treatment as it progresses. They also afford opportunities for more detailed examination of the mechanisms of change during therapy (e.g., early change; sudden change; links to the working alliance).

These brief, session-by-session measures are not yet available for all disorders. In the field of the eating disorders, most self-report questionnaires are designed to measure change over a longer time frame (typically, a month), are too long for weekly use, or do not reflect the

changes that one would expect in the short-term. For example, different versions of the Eating Attitudes Test (EAT-40 [8] and EAT-26 [9]) are relatively long and have not been designed to reflect short-term change. They are more often used as screening tools to assess eating disorder risk in non-clinical populations. Other measures are designed to generate diagnoses (10), and are therefore relatively insensitive to change during therapy. Another group of measures provide a more comprehensive, multidimensional assessment of eating pathology, such as the Eating Pathology Symptoms Inventory (11), the Eating Disorder Inventory (12), and the Stirling Eating Disorders Scales (13). However, these are relatively lengthy, making them impractical to use on a session-by-session basis. Furthermore, concerns have been raised regarding the psychometric properties of some of these multidimensional measures – particularly the Stirling Eating Disorder Scales and the Eating Disorders Inventory (14-16). Those brief measures that have been developed for the eating disorders focus on diagnostic and screening purposes (e.g., Short Evaluation of Eating Disorders [17]), and are not designed or able to track cognitive and symptom change during therapy.

One of the best-established self-report measures of eating disorder pathology is the Eating Disorders Examination-Questionnaire (EDE-Q [18-19]). Developed as a self-report version of the Eating Disorders Examination (20), the EDE-Q generates frequency ratings for key eating disorder behaviors (e.g., binge eating, self-induced vomiting, and laxative misuse) as well as four attitudinal subscales (Restraint, Weight Concern, Shape Concern, and Eating Concern). The EDE-Q is widely used to assess treatment effectiveness (21-23), and has the advantage of being free to use for clinicians. However, the EDE-Q has two drawbacks when compared to the development of such instruments in other disorders, as outlined above. First, it measures change over the previous 28 days, making it unsuitable for registering shorter-term change during therapy. Second, it is relatively long compared with measures such as the GAD-7 and the PHQ-9, making it less likely to be completed on a session-by-session basis. Furthermore, recent post hoc assessments of its psychometric properties have shown that the factor structure of the EDE-Q does not match or approximate to the item groupings used in its sub-scales (24-26). Its measure of objective bingeing has also been called into question (27).

Thus, the clinical utility of the EDE-Q is limited for a variety of reasons, though it remains useful as a measure of outcome of therapy for the eating disorders overall (e.g., used monthly, or for pre-post therapy change).

To summarise, there is a need for a brief, clinically accessible measure of eating pathology, which can be used for the specific purpose of monitoring session-by-session change in core eating disorder behaviors and attitudes. Moreover, that measure needs to be psychometrically sound and clinically valid, showing correspondence with established measures such as the EDE-Q. Allowing the measurement of session-by-session changes in therapy for eating disorders would give such an instrument the potential to determine the importance and sequencing of changes in therapy (e.g., early change; sudden gains; link to alliance), as demonstrated in other disorders.

This aim of this study is to report the development and preliminary validation of a brief, session-by-session measure of eating pathology that meets the criteria outlined above. The factor structure and test-retest reliability of the measure will be established, and it will be validated using established clinical indices (diagnosis; levels of eating pathology; depression; anxiety). Finally, changes in attitudinal patterns across the course of therapy will be examined, to demonstrate the responsiveness of the measure to psychological intervention.

Method

Design

The study used a mixed comparative and correlational design, with a combination of cross-sectional and longitudinal data.

Ethical Considerations

This study was reviewed and approved by the Ethics Committee of the Department of Psychology, University of Sheffield, UK.

Initial Development of the ED-15 Items

The principle was to develop a brief measure of the core attitudinal items, plus a set of behavioral items, to reflect eating pathology over the past week. To ensure content validity,

the items were developed through an iterative process of generation and reduction, carried out by four of the authors (MT, HT, VM and GW). The attitudinal items were generated using a corpus of clinical notes and cognitive records that the authors had accumulated in their experience of the delivery of psychological therapies for eating disorders, to ensure representativeness and validity of the items. The initial pool was generated by each author independently, then they were collated. Where two or more items were found to reflect similar attitudes, these four authors reached a consensus about which to remove or how to merge them. This process resulted in eleven attitudinal items (see Table 1) in the version to be used in the analyses below, accompanied by weekly ratings of the frequency of objective binge-eating and vomiting (episodes across the week) and of restriction, laxative use and exercise to lose weight (days per week that these were undertaken). The final set of items is given in Appendix A.

Participants

Three samples were used. The first was a large non-clinical group of males and females, some of whom completed the measure on two occasions. The second was a self-reported group of eating-disordered women, where no formal diagnoses were available. The final group consisted of formally diagnosed eating disorder patients, undertaking cognitive-behavioral therapy.

The non-clinical sample consisted of 438 woman and 93 men (mean age = 30.4 years; *SD* = 12.1; range = 18-71) who volunteered to take part in an online survey relating to eating attitudes. They were recruited via a university-wide email that went to all relevant students and staff members, asking them to participate by clicking on a weblink that took them directly to the survey. Thus, they were all students (undergraduate or postgraduate) or staff members in a UK university. Their reported weight and height gave them a mean body mass index (BMI) of 23.0 (*SD* = 4.64). Although the males tended to have a higher BMI than the females, this difference did not achieve significance (mean = 24.1 vs 22.8; $t = 1.92$, $P = .051$). Of these participants, 149 (129 women; 20 men) agreed to participate a second time (one to three weeks later) in order to examine test-retest reliability. There were no exclusion criteria for this

group, other than being under 18 years of age.

The self-reported eating disorder group consisted of 63 women who reported currently suffering from eating disorders. These women had a mean reported age of 28.7 years (*SD* = 9.97; range = 18-59) and a mean BMI of 22.2 (*SD* = 6.51; range = 12.7 to 48.2). They were recruited via Project Heal, Canada (a support group for eating disorder sufferers), rather than via clinics. Therefore, their diagnoses were self-reported rather than clinically confirmed (though each reported having been given her diagnosis by a clinician). Each member of that organisation's listserv was sent an email, asking them to participate by clicking on the relevant weblink. Anybody who reported that they did not have a current eating disorder was excluded from the study. One male was also excluded, as he was the only such case and comparison would not be possible or meaningful. Of the 63 women, 23 reported that they had a current diagnosis of anorexia nervosa (mean age = 28.8 years, *SD* = 9.00); 13 reported a diagnosis of bulimia nervosa (mean age = 26.6 years, *SD* = 9.54); and 27 reported a diagnosis of eating disorder not otherwise specified (EDNOS), including four with binge eating disorder (mean age = 30.8 years, *SD* = 11.2).

The final group consisted of 33 women (mean BMI = 23.8, *SD* = 3.66; mean age = 30.8 years, *SD* = 6.20) with formal diagnoses of bulimia nervosa or atypical bulimia nervosa, who undertook and completed a ten-session course of cognitive behavioral therapy for their eating disorder. They were recruited as a case series of patients who were offered that therapy, with the only exclusion criteria being severe self-harm or suicide risk. Each was diagnosed using the Eating Disorders Examination (20) by an appropriately trained clinician. They completed the ED-15 weekly across the course of the therapy. This group's data were used to demonstrate the feasibility of weekly completion of the ED-15, to validate the scores attained by the self-reported eating-disordered group, to provide test-retest reliability scores for that group, and to test the pattern of change in different aspects of psychopathology across therapy.

Measures and Procedure

The survey was distributed using Qualtrics survey software. The participants in the

non-clinical sample completed four measures at the first time point, and repeated one at the second time point for retest purposes. The self-reported eating-disordered group completed the same four measures (once only). The clinically diagnosed eating disorder group completed the ED-15 weekly over the ten weeks of therapy (alongside other measures completed for clinical purposes), with no failures to complete the ED-15. Twenty-three of this last group completed the ED-15 twice over the week prior to starting therapy, with a mean and median gap of seven days.

Eating Disorder-15 (ED-15). This measure was devised for the purposes of this research, using item generation methods and aims as outlined above. The original measure consisted of 11 attitudinal items and a further five behavioral items (see Table 1 and Appendix A). It was completed by all participants, and repeated by the test-retest subsamples.

Eating Disorder Examination-Questionnaire (EDE-Q, version 6 [19]**).** The EDE-Q is a well-established 28-item measure of eating attitudes and behaviors over the past 28 days. Its psychometric properties are generally acceptable, though they vary across studies (25). Its scoring system yields four attitudinal scales (Restraint; Eating Concern; Weight Concern; Shape Concern) and a number of behavioral measures (e.g., objective and subjective bingeing, vomiting, laxative use). However, factor analysis of the attitudinal items does not reflect these four scales, suggesting instead that there are two or three scales (26,28). The validity of the behavioral items is also variable when compared with the interview version of the measure, with particular concerns regarding the accuracy of the EDE-Q objective bingeing measure in some studies (27).

Patient Health Questionnaire (PHQ-9 [5]**).** The PHQ-9 is a nine-item self-report measure of depression, designed to be used repeatedly with the same patient to screen and to monitor severity of mood state. It is used routinely as a session-by-session tool in measurement of depression in the UK's IAPT scheme. It has well-established psychometric properties.

Generalised Anxiety Disorder Questionnaire (GAD-7 [4]**).** The GAD-7 is a seven-item measure of anxiety, used to screen and monitor cases. It is also used as a session-by-

session measure within the IAPT programme. Its psychometric properties are satisfactory.

Data Analysis

The data from the first cohort of women (N = 438) was used to establish the factor structure of the attitudinal scales of the ED-15. The focus on females for this purpose reflected the gender bias in the eating disorders, making the females' scores more pertinent to understanding eating pathology (without variance caused by inclusion of males). Principal components analysis was used, with Varimax and Direct Oblimin rotations, to establish the most meaningful potential scales. Cronbach's *alpha* was used to establish the internal consistency of the resultant scales, and Spearman-Brown coefficients were used to calculate their split-half reliability. The data were split into the first and second halves of the listed items (rather than the 'even/odd' item method, which can result in different outcomes). The scores of non-clinical women (N = 438) and men (N = 93) were compared using independent sample *t*-tests, in order to demonstrate test the validity of the ED-15 (where one would expect higher scores among female participants). Test-retest reliability was established using Pearson's correlation coefficients, separately for men (N = 20) and women (N = 129) who completed the measure twice. In this case, it was anticipated that similar levels of reliability would be demonstrated, so that the measure's characteristics could be shown to be robust, regardless of gender. The concurrent validity of the ED-15 relative to the EDE-Q and to the measures of anxiety and depression was tested using Pearson's correlations. Finally, the concurrent validity of the ED-15 was compared to that of the EDE-Q, contrasting the scores of the 438 non-clinical women with those of the 63 self-reported eating-disordered women, using one-way ANOVAs (with post hoc Least Significant Difference tests). The self-reported eating-disordered women were divided into their self-reported diagnostic groups, to determine whether there were inter-diagnosis differences or not (the latter possibility being more consistent with a transdiagnostic model, where no difference would be expected). All analyses were conducted using SPSS v.21.

Results

Factor Structure of the ED-15 Attitudinal Items

In keeping with the exploratory nature of the factor analysis at this initial developmental stage, principal components analysis was used to determine the factor structure of the ED-15 among the 438 women who completed the measure in the main phase of data collection. The most meaningful factor structure emerged with a Direct Oblimin rotation, meaning that the factors were intercorrelated (component correlation = .551). This association is a common pattern with measures of psychopathology, where the resultant dimensions are often correlated rather than being orthogonal. The results of this analysis are presented in Table 1. The analysis identified two factors, using the criteria of scree analysis and an eigenvalue greater than 1.0. They accounted for 62.0% and 10.3% of the variance in scores, respectively.

Insert Table 1 about here

Ten of the items each loaded clearly onto one of the two factors (factor loading > .5; difference of at least .2 between loadings). However, the remaining item (number 5) did not load uniquely onto either factor, and was therefore excluded from all further consideration. Six items (2, 4, 6, 7, 10, and 11) loaded onto one factor, which was labelled 'Weight & Shape Concerns'. The other factor consisted of four items (1, 3, 8, and 9) and was labelled 'Eating Concerns'. The internal consistencies of the scales were both strong (Cronbach's *alpha* = .938 and .802 respectively).

The two subscales were scored by taking the item means of the six and four items in each scale, and a total score was calculated using the mean of the full set of ten items. The mean scores for this female sample are given in Table 1 (possible range of scores = 0-6, with higher scores indicating greater eating pathology). The final measure and scoring system are given in Appendix A.

Split-Half Reliability of the ED-15 Attitudinal scores

These analyses were undertaken for the first sample as a whole, as results were very

similar for males and females separately. The split-half reliability of the scale as a whole was satisfactory (Spearman-Brown coefficient = .926). The same was true for the Weight & Shape Concerns and the Eating Concerns scales (Spearman-Brown coefficients = 0.926 and .704, respectively).

Norms for the ED-15 and Other Measures

Table 2 shows the mean scores for the ED-15, EDE-Q, PHQ-9 and GAD-7 for the 438 non-clinical women and 93 non-clinical men who completed the main study. In all cases apart from the ED-15 behavioral items (which are reported only for those who used the each behavior at all), the scores show that the females had significantly greater levels of eating pathology, depression and anxiety than the males. The mean scores on the existing measures were in the normal range.

Insert Table 2 about here

In the case of the ED-15 behavioral items, the pattern was that proportionally more females than males undertook the behavior. However, in the case of those who undertook the behavior at all, there were no differences between men and women in the frequency of the behavior (number of episodes; days used per week).

Test-Retest Reliability of the ED-15 Attitudinal Scales

Of the participants in the first wave of data collection, 149 (20 male, 129 female) undertook the second wave, completing the ED-15 again. The mean gap between completions was 18 days (range = 14-22). Pearson's correlations between time 1 and time 2 scores were as follows: Overall score: r = .908 (females – r = .901; males – r = .934); Weight & Shape Concerns: r = .903 (females – r =.896; males – r = .922); and Eating Concerns r = .860 (females – r = .852; males – r = .881). All correlations were significant at the $P < .001$ level.

Among the 23 women with formally-diagnosed eating disorders who completed the measure twice prior to therapy, the test-retest reliability of the Overall ED-15 score was r =

.785. Test-retest reliability on the Weight & Shape Concerns scale was r = .788, and the test-retest reliability of the Eating Concerns scale was r = .806. In all cases, the association was strong ($P < .001$). Given this evidence of the stability of ED-15 scores, it can be concluded that changes as a result of therapy are unlikely to be due to random fluctuations.

Concurrent Validity of the ED-15 Attitudinal and Behavioral Items

In order to establish the concurrent validity of the ED-15 among the non-clinical group, scores on its scales were correlated with those on the existing, validated EDE-Q scales. Furthermore, the ED-15 and EDE-Q scales were each correlated with participants' BMIs. Only the findings for the sample as a whole are reported, because they were almost identical to those for the males and females separately. Table 3 shows the resulting Pearson correlation coefficients. It is noteworthy that the correlation between the total scores of the two measures was very strong ($r = +.889$), suggesting that they measure very similar constructs, despite their time frames being different.

Insert Table 3 about here

Considering the correlations between attitudinal subscales, the strongest correlations were between the ED-15 Weight & Shape Concerns scale and the EDE-Q Weight Concern and Shape Concern scales, supporting the concurrent validity of this ED-15 scale. However, the ED-15 Eating Concerns scale was more evenly associated with the four EDE-Q scales. For both the EDE-Q and the ED-15, concerns about weight and shape were associated with a higher BMI, but other eating attitudes were not, suggesting similarity in the links between weight and specific attitudes (and thus, comparability of the EDE-Q and ED-15).

For the behavioral scales that could be compared (objective bingeing, vomiting, laxative abuse, excessive exercise), the number of non-clinical individuals reporting some behaviors differed across measures. On the EDE-Q, 121 individuals reported objective binges, but only 100 did so on the ED-15. In contrast, more reported excessive exercise on the ED-

15 than on the EDE-Q (181 vs 117). However, there was greater concordance between the ED-15 and EDE-Q in the numbers who vomited (14 each) and those who used laxatives (six vs ten). For those who reported each of the behaviors on both measures, the correlation coefficients for frequencies of the behaviors were: objective bingeing - $r = +.601$; vomiting - $r = +.789$; laxative use - $r = +.971$; and exercise - $r = +.628$ ($P < .001$, in all cases).

To summarise, in the non-clinical group, the ED-15 had strong concurrent validity (relative to the EDE-Q) in terms of eating attitudes scales, purging behaviors (vomiting and laxatives), and BMI. The two measures differed in the number of identified cases where the individual binged objectively and where there was excessive exercise, with the EDE-Q showing more of the former (in keeping with existing literature [26]) and the ED-15 identifying more of the latter. However, in those cases where there was concordance over the presence of the relevant behavior, the measures showed a strong agreement over the level of those behaviors.

Convergent Validity of the ED-15 Relative to Non-Eating Pathology Measures

In the non-clinical group, Pearson's correlations were used to determine the associations between the two measures of eating attitudes (ED-15 and EDE-Q) and the measures of anxiety and depression (GAD and PHQ), in order to demonstrate whether either eating measure was superior in reflecting wider psychopathology. These analyses are reported for the non-clinical sample as a whole, as there were no differences in outcome when conducted separately for females and males. Table 4 shows that the patterns of association were similar for the ED-15 and the EDE-Q, with moderate, highly significant correlations between the measures of eating and the measures of other anxiety and depression. In both cases, the correlation with depression tended to be higher than the association with anxiety. To summarise, the ED-15 has comparable concurrent validity to the EDE-Q, based on the two measures' associations with more general psychopathology.

Insert Table 4 about here

Comparability of Scores on the ED-15 and EDE-Q for the Non-Clinical Sample

In order to determine the comparability of scores on the two eating measures among the first sample, Table 5 shows the scores that represent different percentiles. While these scores are not intended to represent any form of translation across measures, it is noteworthy that the EDE-Q tended to have lower scores in the normal range for females than the ED-15, but that the scores became more similar at the higher end of the distribution. However, that move towards similarity was not found among the males.

Insert Table 5 about here

Clinical Validation of the ED-15

Table 6 shows the mean scores on the ED-15 and the EDE-Q attitudinal scales of the 438 non-clinical women and the 63 women with self-reported eating disorders. There were similar results for both measures, with the three self-reported diagnostic subgroups (among those with a self-reported eating disorder) having significantly higher scores on all scales than the non-clinical group. There were no differences between the anorexia nervosa and bulimia nervosa groups on most scales. However, there was one pattern of differences between diagnoses, with the anorexia nervosa group scoring higher on the EDE-Q Restraint and Eating Concern scales and on the ED-15 Eating Concerns scale. To summarise, the ED-15 and EDE-Q had comparable patterns of clinical validity.

Insert Table 6 about here

Considering the 33 formally diagnosed eating-disordered women, Table 7 provides weekly, session-by-session ED-15 scores across the course of a ten-session cognitive-

behavioral therapy for individuals with bulimic or atypical normal-weight eating disorders. It is noteworthy that the pre-treatment scores of the formally diagnosed group were similar to those of the comparable self-reported clinical subgroups in Table 6. The ANOVAs demonstrate that the scores on the two ED-15 scales fell over time, as one would hope. However, post-hoc multiple comparison tests ($P < .05$) show that the two scores changed at different time points. The Eating Concerns score showed its greatest change between sessions 1-4 and then from session 8-10), while the Weight & Shape Concerns score changed most at a later point in the therapy (between sessions 3-6, and then again from sessions 8-10). While preliminary, this difference in patterns of change indicates the distinctive clinical utility of each ED-15 scale.

Insert Table 7 about here

In order to contextualise these overall changes in ED-15 scores across brief CBT, Table 5 shows that the mean scores of the formally-diagnosed eating-disordered group changed from between the 90th and 95th centile for non-clinical women (session 1) to below the 75th centile (session 10), suggesting substantial reduction in eating disorder cognitions. However, it is also important to consider how these changes in ED-15 scores were or were not mirrored in other clinical indices. First, the change in the same patients' EDE-Q scores between sessions (session 1: $M = 3.99$, $SD = 1.45$; session 1: $M = 2.42$, $SD = 1.32$; paired $t = 3.79$, $P = .002$) was of a comparable degree to that for the ED-15, with the patients' mean scores falling from between the 90th and 95th centile to just below the 75th centile (Table 5). Second, the behavioral remission rate (no use of any bingeing or compensatory behaviors by the end of therapy) was 48%, suggesting that the change in cognitions was associated with a relatively high level of reduction in behaviors. This abstinence level is nearly identical to that reported elsewhere for CBT for comparable eating disorders (22). To summarise, the change during therapy in ED-15 eating disorder cognitions has been shown to be clinically meaningful, relative to other, established outcome indices.

Discussion

This study has reported on the development and initial psychometric and clinical validation of the Eating Disorder-15 (ED-15), a brief self-report measure of eating attitudes and behaviors, developed for use on a weekly basis in clinical settings. The long-term aim is that the measure should be tested for its clinical utility in monitoring therapy process and outcomes on a session-by-session basis. As a basis for that work, the present study has addressed core psychometric and clinical properties, comparing the ED-15's properties to those of the well-established EDE-Q (which is longer, and is used for measuring attitudes and behaviors over the previous month).

A large non-clinical female sample was used to determine that the ED-15 attitudinal items have a two-factor structure, resulting in two subscales - 'Weight & Shape Concerns' and 'Eating Concerns'. Each subscale had satisfactory internal consistency, split-half reliability and test-retest reliability. There was a similar difference between men's and women's scores to that found on the EDE-Q. The ED-15 scales correlated very highly with the existing, validated EDE-Q scales, suggesting that they measure very similar constructs. There was similar concordance between the ED-15 and EDE-Q when it came to associations with BMI and levels of eating behaviors (though most strongly for purging behaviors). The ED-15 and EDE-Q demonstrated similar patterns of association with measures of anxiety (GAD-7) and depression (PHQ-9). Overall, these findings indicate that the ED-15 has comparable concurrent validity to the EDE-Q. Finally, clinical validation was demonstrated by comparing the scores of the original sample with those given by two groups of women with eating disorders – one self-reported, and one formally diagnosed (though with similar scores). Again, the level of differentiation between the non-clinical group and the self-reported eating-disordered group was comparable for the ED-15 and the EDE-Q. As a further indication of the ED-15's clinical utility, it was shown that changes across a brief course of CBT were comparable to those shown using other clinical indices, resulting in more normative cogntions. Tables 5, 6 and 7 provide norms for use in future research and clinical work. The limitation

imposed by the self-reported nature of the eating disorders in the relevant group is important to note, though the similarity of scores across the self-reported and comparable formally-diagnosed group attenuates that concern to a degree.

In clinical terms, the ED-15 is designed to complement measures such as the EDE-Q, rather than to supplant them. Those more detailed questionnaires provide a level of clinical detail that is likely to be more appropriate pre- and post-treatment, measuring overall outcome. In contrast, the brevity of the ED-15 makes it more suitable for session-by-session use. It is free to use, rapidly completed (e.g., at the start of a session or while waiting), and easy to score, and norms have been provided to allow clinicians to interpret scores. In addition, the ED-15 has a scoring system that has been derived from factor analysis, making it more robust than questionnaires such as the EDE-Q and EDI, where factor analysis was applied following the implementation of the measure (and not always successfully). As with similar measures of other disorders (e.g., depression, anxiety), the ED-15 is intended to be useful for session-by-session measurement of key eating disorder attitudes and behaviors, allowing the clinician to be responsive to change (or lack of it) in the early and later stages of treatment, and explicitly to link the tasks of therapy with changes in psychopathology. The scores can be discussed in supervision, to ensure that the clinician remains on track with the delivery of evidence-based approaches (19,30). The ED-15 is potentially suitable for use in other languages, dependent on appropriately rigorous translation processes. A related consideration in future research is whether the wording is appropriate use in all English speaking countries, though different variants should be used only if justified. Modification might also be needed for children and adolescents. A further issue that the ED-15 has in common with other measures (e.g., the EDE-Q) is that it is not fully reflective of the range of anorexia nervosa presentations. Its focus on issues of weight and shape means that its utility is limited with non-fat-phobic anorexia nervosa. There is a need for most such measures to be developed to assess the pathology of this subset of anorexia nervosa patients.

This study found little difference across diagnoses in the ED-15 scores of the self-reported eating-disordered groups (as well as showing that the self-reported eating-disordered

groups' scores were similar to those of the comparable women in the formally-diagnosed clinical group). This finding is similar to the outcome of other studies using different measures of eating pathology. While this conclusion might appear to be in keeping with the transdiagnostic model of the eating disorders (29), it is equally possible that the self-selected, self-diagnosed nature of the individuals in question masked true group differences. Therefore, this study needs to be replicated with a diagnostically diverse clinical group with confirmed diagnoses, which is large enough to allow for confirmatory factor analysis.

As per the aim of this study, the ED-15 has the potential to inform research into the outcome and process of change in therapy. As well as allowing comparison between pre- and post-therapy scores (as has been done using measures such as the EDE and EDE-Q [22-23]), a session-by-session measure can be used to identify change during therapy and its implications. Examples might include detecting early cognitive and behavioral responses to therapy and sudden changes in behavior, both of which have been found to be important in other disorders (1-3). While this measure has a largely cognitive-behavioral origin, so does the EDE-Q, and that measure has been used to show effects of other treatments for the eating disorders (31-32). Therefore, the ED-15 can be suggested as a within-treatment measure of the impact of a range of therapies for the eating disorders. In that context, it could be used to validate clinicians' judgements regarding session-by-session progress. It could also be used in combination with real-time assessment measures (33) to determine the possible interaction of symptom fluctuation and overall symptom change. Finally, in clinical terms, the ED-15 can be used to identify the most effective components of treatment and sequences of changes (e.g., does cognitive change predict behavioral change or vice versa; what is the link between change in the working alliance and symptoms [34]), thus informing the optimum sequence of therapy delivery. However, it should be stressed that the ED-15 is not designed for diagnostic or screening purposes, and that other measures are more likely to be useful in this regard (10,17). The other potential use of the ED-15 is as a brief measure of eating pathology in experimental studies (e.g., priming, the impact of food intake, cognitive dissonance effects), where a brief, focused measure is needed as the dependent variable, though this possibility

clearly needs to be established empirically.

Conclusions

The ED-15 is a brief, reliable and valid measure of core eating disorder features that are commonly addressed in therapy. It is not suggested as an alternative to existing pre- and post-treatment measures such as the EDE-Q, but as a complementary tool for measuring session-by-session impact of treatment for eating disorders.

Declaration of interests

The authors have no competing interests.

References

1. Crits-Christoph P, Connolly MB, Gallop R, Barber JP, Tu X, Gladis M, Siqueland L. Early improvement during manual-guided cognitive and dynamic psychotherapies predicts 16-week remission status. J Psychotherapy Pract Res 2001; 10:145-154.
2. Tang TZ, DeRubeis RJ. Sudden gains and critical sessions in cognitive-behavioral therapy for depression. J Consult Clin Psychol 1999; 67:894-904.
3. Tang TZ, DeRubeis RJ, Beberman R, Pham T. Cognitive changes, critical sessions, and sudden gains in cognitive-behavioral therapy for depression. J Consult Clin Psychol 2005; 73:168-172.
4. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Int Med, 2006; 166:1092-1097.
5. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. Psychiatric Annals 2002; 32:1-7.
6. Behar E, Alcaine O, Zuellig AR, Borkovec TD. Screening for generalized anxiety disorder using the Penn State Worry Questionnaire: A receiver operating characteristic analysis. Journal of Beh Ther Exp Psychiatry 2003; 34:25-43.
7. Connor KM, Davidson JR, Churchill LE, Sherwood A, Weisler RH, Foa E. Psychometric properties of the Social Phobia Inventory (SPIN): New self-rating scale. Br J Psychiatry 2000; 176:379-386.
8. Garner DM, Garfinkel PE. The Eating Attitudes Test: An index of the symptoms of anorexia nervosa. Psychol Med 1979; 9:273-279.
9. Garner DM, Olmsted MP, Bohr Y, Garfinkel PE. The Eating Attitudes Test: psychometric features and clinical correlates. Psychol Med 1982; 12:871-878.
10. Stice E, Telch CF, Rizvi SL. Development and validation of the Eating Disorder Diagnostic Scale: A brief self-report measure of anorexia, bulimia, and binge-eating disorder. Psychol Assess 2000; 12:123-131.
11. Forbush KT, Wildes JE, Pollack LO, Dunbar D, Luo J, Patterson K, Petruzzi L, Pollpeter M, Miller H, Stone A, Bright A, Watson D. Development and validation of the Eating

Pathology Symptoms Inventory (EPSI). Psychol Assess 2013; 25:859-878.

12. Garner DM, Olmsted MP, Polivy J. Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. Int J Eat Disord, 1983; 2:15-34.
13. Williams GJ, Power KG, Miller HR, Freeman CP, Yellowlees A, Dowds T, Walker M, Parry-Jones WL. Development and validation of the Stirling Eating Disorder Scales. Int J Eat Disord 1994; 16:35-43.
14. Gamble C, Bryant-Waugh R, Turner H, Jones C, Mehta R, Graves A. An investigation into the psychometric properties of the Stirling Eating Disorder Scales. Eat Beh, 2006; 7:395-403.
15. Openshaw C, Waller G. Psychometric properties of the Stirling Eating Disorder Scales with bulimia nervosa patients. Eat Behav 2005; 6:165-168.
16. Tasca GA, Illing V, Lybanon-Daigle V, Bissada H, Balfour L. Psychometric properties of the Eating Disorders Inventory-2 among women seeking treatment for binge eating disorder. Assessment 2003; 10:228-236.
17. Bauer S, Winn S, Schmidt U, Kordy H. Construction, scoring and validation of the Short Evaluation of Eating Disorders (SEED). Eur Eat Disord Rev 2005; 13:191-200.
18. Fairburn CG, Beglin SJ. Assessment of eating disorders: Interview or self-report questionnaire? Int J Eat Disord 1994; 16:363-370.
19. Fairburn CG. Cognitive behavior therapy and eating disorders. New York, NY: Guilford; 2008.
20. Cooper Z, Fairburn C. The Eating Disorder Examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. Int J Eat Disord 1987; 6:1-8.
21. Byrne SM, Fursland A, Allen KL, Watson H. The effectiveness of enhanced cognitive behavioural therapy for eating disorders: An open trial. Beh Res Ther 2011; 49:219-226.
22. Fairburn C, Cooper Z, Doll H, O'Connor M, Bohn K, Hawker D, Wales JA, Palmer R. Transdiagnostic cognitive-behavioral therapy for patients with eating disorders: a two-

site trial with 60-week follow-up. Am J Psychiatry 2009; 166:311-319.

23. Fairburn CG, Cooper Z, Doll HA, O'Connor ME, Palmer RL, Dalle Grave R. Enhanced cognitive behaviour therapy for adults with anorexia nervosa: A UK–Italy study. Beh Res Ther 2013; 51:R2-R8.
24. Allen KL, Byrne SM, Lampard A, Watson H, Fursland A. Confirmatory factor analysis of the Eating Disorder Examination-Questionnaire (EDE-Q). Eat Beh 2011; 12:143-151.
25. Berg KC, Peterson CB, Frazier P, Crow SJ. Psychometric evaluation of the Eating Disorder Examination and Eating Disorder Examination-Questionnaire: A systematic review of the literature. Int J Eat Dis 2012; 45:428-438.
26. Waller G, Hamill M, van Gerko K, Hughes ML. Psychometric properties and discriminant validity of the Eating Disorders Examination-Questionnaire form. Paper presented at the Eating Disorders Research Society. Port Douglas; September 2006.
27. Black CM, Wilson GT. Assessment of eating disorders: Interview versus questionnaire. Int J Eat Disord 1996; 20:43-50.
28. Peterson CB, Crosby RD, Wonderlich SA, Joiner T, Crow SJ, Mitchell JE, Bardone-Cone AM, Klein M, le Grange D. Psychometric properties of the Eating Disorder Examination-Questionnaire: Factor structure and internal consistency. Int J Eat Disord 2007; 40:386-389.
29. Fairburn CG, Cooper Z, Shafran R. Cognitive behaviour therapy for eating disorders: A “transdiagnostic” theory and treatment. Beh Res Ther 2003; 41:509-528.
30. Waller G, Cordery H, Corstorphine E, Hinrichsen H, Lawson R, Mountford V, Russell K. Cognitive behavioral therapy for eating disorders: A comprehensive treatment guide. Cambridge, UK: Cambridge University Press, 2007.
31. Poulsen S, Lunn S, Daniel SI, Folke S, Mathiesen BB, Katznelson H, Fairburn CG. A randomized controlled trial of psychoanalytic psychotherapy or cognitive-behavioral therapy for bulimia nervosa. Am J Psychiatry 2014; 171:109-116.
32. Zipfel S, Wild B, Groß G, Friederich HC, Teufel M, Schellberg D, Giel KE, de Zwaan M, Dinkel A, Herpertz S, Burgmer M, Löwe B, Tagay S, von Wietersheim J, Zeeck A,

Schade-Brittinger C, Schauenburg H, Herzog W. Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study): randomised controlled trial. Lancet 2014; 383(9912):127-137.

33. Smyth J, Wonderlich S, Crosby R, Miltenberger R, Mitchell J, Rorty M. The use of ecological momentary assessment approaches in eating disorder research. Int J Eat Disord 2001;30: 83-95.
34. Brown A, Mountford VA, Waller G. An idea worth researching: Is the therapeutic alliance overvalued in the treatment of eating disorders? Int J Eat Disord 2013; 46:779-782.

Table 1

Principal components analysis (direct oblimin rotation) of ED-15 items for non-clinical females (N = 438), with item mean scores and internal consistency of resulting scales

	Items	Factor 1	Factor 2
	Over the past week, how often have I:	**Weight & Shape Concerns**	**Eating Concerns**
1	Worried about losing control over my eating	.363	**.588**
2	Avoided activities or people because of the way I look	**.911**	-.157
3	Been preoccupied with thoughts of food and eating	.272	**.636**
4	Compared my body negatively with others'	**.784**	.111
5	Been worried that whatever I ate, I would gain lots of weight	.503	.467
6	Avoided looking at my body (e.g., in mirrors; wearing baggy clothes) because of the way it makes me feel	**.921**	-.072
7	Felt distressed about my weight	**.763**	.234
8	Checked my body to reassure myself about my appearance (e.g., weighing myself; using mirrors)	-.092	**.841**
9	Followed strict rules about my eating	-.035	**.804**
10	Felt distressed about my body shape	**.814**	.152
11	Worried that other people were judging me as a person because of my weight and appearance.	**.864**	-.005
	Eigenvalue	6.82	1.14
	Variance explained	62.0%	10.3%
	Cronbach's *alpha*	.938	.802
	Item mean (*SD*)	1.79 (1.49)	2.44 (1.37)

Table 2

Mean scores on measures of eating, anxiety and depression for the non-clinical males and females (first sample only)

	Female (*N* = 438)		Male (*N* = 93)		t-test	
ED-15 scale	*M*	*(SD)*	*M*	*(SD)*	*t*	*P*
Attitudinal scales						
Weight & Shape Concerns	1.79	(1.49)	1.02	(1.28)	4.38	.001
Eating Concerns	2.44	(1.37)	1.75	(1.22)	4.98	.001
Total	2.05	(1.33)	1.31	(1.15)	5.32	.001
Behavioral items (per week)						
Objective binges (82 F; 18 M)	2.57	(1.96)	1.78	(1.06)	1.68	*NS*
Vomiting episodes (13 F; 1 M)	2.56	(2.14)	1.00	(-)	0.73	*NS*
Laxative use days (6 F; 0 M)	2.83	(1.72)	-	(-)	-	-
Exercise days (143 F; 28 M)	3.28	(1.68)	3.61	(1.64)	0.95	*NS*
Restriction days (188 F; 26 M)	4.35	(2.04)	3.88	(2.05)	1.09	*NS*
EDE-Q scale						
Restraint	1.59	(1.44)	1.11	(1.33)	2.65	.01
Weight Concern	1.99	(1.62)	0.92	(1.20)	6.56	.001
Eating Concern	1.07	(1.31)	0.46	(0.88)	4.93	.001
Shape Concern	1.79	(1.47)	0.98	(1.22)	4.99	.001
Global	1.61	(1.32)	0.87	(1.00)	5.45	.001
PHQ						
Depression	6.68	(5.76)	5.08	(5.71)	2.27	.025
GAD						
Anxiety	6.24	(5.23)	4.52	(5.21)	2.67	.01

Table 3

Pearson's correlations (*r*) between ED-15 scales and EDE-Q scales for the whole non-clinical sample (results were near identical for males and females separately)

	EDE-Q scale					
ED-15 scale	Restraint	Eating Concern	Weight Concern	Shape Concern	Global	**BMI**
Weight & Shape Concerns	.550***	.728***	.860***	.884***	.840***	.169***
Eating Concerns	.744***	.700***	.721***	.726***	.800***	.043
Total	.674***	.773***	.870***	.888***	.889***	.131**
BMI	.033	.090	.214***	.184***	.152**	-

** P < .01; *** P < .001

Table 4

Pearson's correlations (*r*) between the ED-15 and EDE-Q scales and the GAD anxiety and PHQ depression scores for the whole non-clinical sample

	GAD anxiety	PHQ depression
ED-15 scales		
Weight & Shape Concerns	.520***	.633***
Eating Concerns	.412***	.456***
Total	.517***	.612***
EDE-Q scales		
Restraint	.305***	.356***
Eating Concern	.490***	.599***
Weight Concern	.481***	.568***
Shape Concern	.504***	.595***
Global	.492***	.585***

*** P < .001

Table 5

ED-15 centile scores for a non-clinical population of adults (range of scores = 0-6), compared to similar centiles for EDE-Q scores (range = 0-6)

	Centile point on measure					
Whole group	*25th*	*50th*	*75th*	*90th*	*95th*	*99th*
EDE-Q Global	0.39	1.14	2.23	3.48	4.15	5.00
ED-15 Total	0.80	1.65	1.80	2.75	4.50	5.15
Females						
EDE-Q Global	0.49	1.27	2.47	3.62	4.30	5.20
ED-15 Total	0.90	1.75	2.95	4.00	4.45	5.20
Males						
EDE-Q Global	0.06	0.40	1.36	2.25	2.90	3.25
ED-15 Total	0.40	0.90	1.80	1.85	4.65	5.00

Table 6

Mean scores on ED-15 and EDE-Q scales for non-clinical and self-diagnosed eating-disordered women

	Non-clinical (*N* = 438)		Anorexia nervosa (*N* = 23)		Bulimia nervosa (*N* = 13)		EDNOS (*N* = 27)		One-way ANOVA		Least Significant Difference tests
	M	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*F*	*P*	(*P* < .05)
ED-15 scales											
Weight & Shape Concerns	1.79	(1.49)	3.93	(1.35)	4.58	(0.87)	3.77	(1.27)	42.6	.001	NC<AN=BN=EDNOS
Eating Concerns	2.44	(1.37)	4.66	(0.91)	4.33	(1.10)	3.84	(0.86)	35.5	.001	NC<EDNOS<AN; NC<BN
Total	2.05	(1.33)	4.22	(1.14)	4.48	(0.90)	3.80	(0.89)	46.7	.001	NC<AN=BN=EDNOS
EDE-Q scales											
Restraint	1.59	(1.44)	4.23	(1.47)	3.60	(1.62)	3.24	(1.62)	37.2	.001	NC<EDNOS<AN; NC<BN
Weight Concern	1.99	(1.62)	4.49	(1.28)	4.75	(1.34)	4.15	(1.43)	39,9	.001	NC<AN=BN=EDNOS
Eating Concern	1.07	(1.31)	3.88	(1.39)	3.42	(1.23)	2.89	(1.29)	54.4	.001	NC<EDNOS<AN; NC<BN
Shape Concern	1.79	(1.47)	4.91	(1.41)	4.95	(1.17)	4.63	(1.20)	43.7	.001	NC<AN=BN=EDNOS
Global	1.61	(1.32)	4.38	(1.26)	4.18	(1.18)	3.73	(1.16)	53.0	.001	NC<AN=BN=EDNOS

Table 7

Mean session-by-session ED-15 scores across the course of ten-session cognitive-behavioral therapy for formally-diagnosed bulimia nervosa and atypical cases (*N* = 33)

	Session										ANOVA	
	1	2	3	4	5	6	7	8	9	10	F	P
Eating concerns	4.68	4.17	4.33	3.75	3.77	3.73	3.75	3.77	2.93	2.77	7.49	.001
(SD)	(0.93)	(1.19)	(1.01)	(1.21)	(1.31)	(1.43)	(1.24)	(1.35)	(1.44)	(1.30)		
Weight and shape concerns	3.94	3.79	3.90	3.46	3.33	3.12	3.36	3.09	2.81	2.55	5.64	.001
(SD)	(1.19)	(0.88)	(1.03)	(1.26)	(1.38)	(1.51)	(1.39)	(1.31)	(1.42)	(1.39)		

Appendix 1 - ED-15 and scoring key

ED-15

This questionnaire considers your eating attitudes and behaviors over the last week. Please complete this measure by ticking the appropriate answers for all items.

	Over the past week, how often have I:	Not at all	Rarely	Occasionally	Sometimes	Often	Most of the time	All the time
1	Worried about losing control over my eating.	0	1	2	3	4	5	6
2	Avoided activities or people because of the way I look	0	1	2	3	4	5	6
3	Been preoccupied with thoughts of food and eating	0	1	2	3	4	5	6
4	Compared my body negatively with others'	0	1	2	3	4	5	6
5	Avoided looking at my body (e.g., in mirrors; wearing baggy clothes) because of the way it makes me feel	0	1	2	3	4	5	6
6	Felt distressed about my weight	0	1	2	3	4	5	6
7	Checked my body to reassure myself about my appearance (e.g., weighing myself; using mirrors)	0	1	2	3	4	5	6
8	Followed strict rules about my eating	0	1	2	3	4	5	6
9	Felt distressed about my body shape	0	1	2	3	4	5	6
10	Worried that other people were judging me as a person because of my weight and appearance.	0	1	2	3	4	5	6

If you have never used any of the following behaviors, please respond with N/A.

	For those that you have used, over the past week, how many times have you:	*Number of times*
a	Binged (felt out of control of your eating, and eaten far more than a person normally would at one go)	
b	Vomited to control your weight (whether you had to make yourself sick or not) *	
	Finally, on how many days in the past week have you:	*Number of days*
c	Used laxatives to control your weight or shape	
d	Restricted or dieted in order to control your weight	
e	Exercised hard in order to control your weight	

** i.e., Using your fingers or medicines to make yourself sick, or vomiting without such aids*

ED-15 scoring key

- All items are positively scored from 0-6.
- The ED-15 includes two attitudinal subscales, scored as follows:
    - Weight & Shape Concerns = mean of items 2, 4, 5, 6, 9, and 10 (add the six scores and divide by 6)
    - Eating Concerns = mean of items 1, 3, 7 and 8 (add the four scores and - divide by 4)
- The Overall attitudinal score is the mean of the scores on all ten items (total the ten items and divide by 10).
- Up to one item can be missed from either scale, and the item mean can be corrected accordingly. If more are missing, then the scores are invalid.