

The Brunnsviken Brief Quality of Life Scale (BBQ): Development and Psychometric Evaluation

Philip Lindner^{a,b}, Ola Frykheden^a, David Forsström^a, Erik Andersson^b, Brjánn Ljótsson^b, Erik Hedman^b, Gerhard Andersson^{b,c} and Per Carlbring^a

^aDepartment of Psychology, Stockholm University, Stockholm, Sweden; ^bDepartment of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden; ^cDepartment of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

ABSTRACT

Measurements of subjective quality of life (QoL) are an important complement to symptom ratings in clinical research and practice. Despite there being several established QoL self-rating scales, we identified a need for a freely accessible, easy-to-use inventory, validated for use with both clinical and non-clinical samples, based on the overall life satisfaction conceptualization of QoL. The Brunnsviken Brief Quality of life scale (BBQ) was designed to meet these requirements. Items were selected by performing a factor analysis on a large data-set of QoL ratings collected previously. Six life areas (Leisure time, View on life, Creativity, Learning, Friends and Friendship, and View of self) were identified as important for overall QoL and were included in the BBQ. A psychometric evaluation was performed using two independent samples: healthy undergraduate students ($n = 163$), and a sample seeking treatment for social anxiety disorder ($n = 568$). Results suggested a unifactorial structure, with good concurrent and convergent validity, high internal and test-retest reliability, and accurate classification ability. We conclude that the BBQ is a valid and reliable measure of subjective QoL for use in clinical and research settings. The BBQ is presently available in 31 languages and can be freely downloaded from www.bbqscale.com.

ARTICLE HISTORY

Received 25 November 2015
Accepted 13 January 2016

KEYWORDS

Psychometrics; quality of life; self-rating scale; life satisfaction

1. Introduction

Quality of life (QoL) has received increased attention within medical and psychological research and practice (Hofmann, Wu, & Boettcher, 2014). Traditionally, psychological well-being has often been conceptualized in the negative sense as the absence of mental illness. Subjective QoL on the other hand, refers to a wider construct, conceptualized as an overall, self-perceived satisfaction with life areas of deemed importance, e.g. work, friendship, leisure time, possibilities to be creative and learn new things (Katschnig, 2006). While mental illness symptoms can be expected to correlate with lower life satisfaction, these two variables represent different concepts and should be

measured independently and not inferred from each other (Frisch, 1998). In psychiatric diagnostics, this distinction corresponds to the clinically significant distress and/or impairment required alongside symptoms to meet diagnostic criteria (American Psychiatric Association, 2013). QoL measures thereby provide information beyond what is conveyed by symptom measures, making them meaningful complements in clinical trials, psychiatric practice and psychological research.

There are numerous valid and reliable self-rating scales for the assessment of subjective QoL (Berzon, Donnelly, Simpson, Simeon, & Tilson, 1995). However, substantial differences in the underlying conceptualization of QoL, even in the generic QoL measures (i.e. not disorder-specific), means that comparing QoL measurements across studies is difficult. Popular QoL scales such as the EuroQol (1990), the RAND-36 (Hays, Sherbourne, & Mazel, 1993), and the Sheehan Disability Scale (Sheehan, Harnett-Sheehan, & Raj, 1996) are primarily measures of lack of symptoms or lack of symptom impact, making them inappropriate for use with healthy populations. Further, this QoL conceptualization departs markedly from the common psychological definition of QoL as an overall life satisfaction (Katschnig, 2006). Although commonly used in clinical psychology research (Hofmann et al., 2014), many of these scales place a heavy psychometric emphasis on somatic functioning and general health, making them ill-suited for use as screening and outcome measures in psychological research and practice. A self-rating scale based on the overall life satisfaction conceptualization of QoL would arguably be more appropriate in fields of psychology, psychiatry, and behavioral medicine.

There are several established QoL measures based on this QoL conceptualization, such as the Satisfaction With Life Scale (Diener, Emmons, Larsen & Griffin, 1985), Quality Of Life Inventory (QOLI) (Frisch, Cornell, Villanueva, & Retzlaff, 1992), Quality of Life Scale (Burckhardt & Anderson, 2003; Flanagan, 1978), Quality of Life Index (Ferrans & Powers, 1985) and QoL Enjoyment and Satisfaction Questionnaire (Endicott, Nee, Harrison, & Blumenthal, 1993). While shown to be valid and reliable measures of QoL, we identified several key constraints that limit their usage and spread. In order to be used in-session in clinical practice and research settings, a scale should be brief (i.e. few items) and presented in an accessible manner. Items should be selected on the basis of empirically shown importance for overall life satisfaction. The scoring procedure should be simple and intuitive, to allow immediate score calculations. To avoid unnecessary barriers to access and usage, a rating scale should be easily accessible, with no licensing costs, copyright limitations, or restrictions to translate.

In sum, although there are several published QoL measures, few are based on the overall life satisfaction conceptualization of QoL and there is no empirically derived, brief and easy-to-use scale based on this conceptualization. The aim of the present study was to develop and test the psychometric properties of a new QoL scale, the Brunnsvikén Brief Quality of life scale (BBQ), which was specifically designed to meet these requirements. We used a large ($n = 4016$) data-set of previously collected QoL ratings from both clinical and non-clinical samples (Lindner, Andersson, Öst, & Carlbring, 2013) to identify the six areas of life showing the greatest importance for a unified, latent QoL factor. We then constructed novel item-phrasings and a novel item response-format, and evaluated the BBQ's performance in both clinical and non-clinical samples using classic psychometric methods.

2. Methods

Item selection for the BBQ was done by performing novel analyses on previously collected data. After development, a psychometric evaluation was performed using two independent samples.

2.1. Development and description of the BBQ

Life areas to include in the BBQ were identified by performing a forced unifactorial maximum-likelihood factor analysis of weighted QOLI items (satisfaction \times importance) in the data-set collected for the Lindner et al. (2013) study (see publication for more details). Six life areas (items) showed factor loadings above .50 when analyzing the screening data ($n = 4016$): Recreation (.65), Philosophy of life (0.63), Creativity (.60), Learning (.59), Self-regard (.55) and Friends and friendship (.51). The same six life areas had the top six factor loadings when a similar factor analysis was performed on post-treatment data ($n = 1268$), albeit with slightly different loadings and order (Recreation: .69; Philosophy of life: .69; Self-regard: .63; Creativity: .63; Learning: .59; Friendship: .56).

These six life areas were used to construct novel items for BBQ. Previous research by our group has shown that importance ratings of life areas may impact comparisons of QoL scores across clinical groups (Lindner et al., 2013). Moreover, in clinical settings, importance ratings may convey significant information, and may be of particular interest in psychological treatments emphasizing that behavior change should be implemented in accordance with one's values (Hayes, Strosahl, & Wilson, 1999). For these reasons, we decided to include not only satisfaction-rating items but also importance-rating items for each life area. Thus, BBQ has a total of 12 items covering six life areas: Leisure, View on life, Creativity, Learning, Friends and Friendship, and View on self. See Appendix A for the full scale. Item-pairs appear sequentially, with the Satisfaction-item for each life area preceding the Importance-item. All items are scored using the same response format, consisting of a five-step Likert rating scale, visually scored 0–4 with written anchor points at 0 (Strongly disagree) and 4 (Strongly agree). The BBQ total score is computed by summing the weighted satisfaction ratings, i.e. by multiplying the Satisfaction and Importance items for each life area and summing the six products for a total score (possible score range 0–96). Item-level data (used e.g. for Cronbach's alpha calculations and factor analyses) thus correspond to item-pairs (i.e. weighted satisfaction ratings). The weighting procedure used for the BBQ resembles the one use for the QOLI, which instead calculates an average of non-zero weighted satisfaction ratings. The weighted sum approach was chosen for the BBQ to simplify the scoring calculation, while maximizing the score impact of the importance weighting. This scoring procedure is also closer to the psychometric standard of summing item scores to provide a scale score.

2.3. Samples and procedure

BBQ data were collected from two independent samples: a non-clinical sample consisting of psychology undergraduate students, and a clinical sample seeking treatment for social anxiety disorder (SAD), recruited from the general public for an ongoing treatment study (Miloff, Marklund, & Carlbring, 2015). The undergraduate students ($n = 163$) provided data

on two occasions (test-retest), 7 ± 1 days apart, with no intervention in between. Participants in the clinical sample provided screening data upon study registration ($n = 568$). A subset of participants in the clinical sample ($n = 123$) who were subsequently included in the study and completed treatment also provided post-treatment data. There was a significant difference in age between samples ($F[1,667] = 24.95, p < .001$), such that the clinical sample was older ($M = 31.84, SD = 12.04$) than the non-clinical sample ($M = 26.82, SD = 7.77$). Samples were matched on sex, with more females in both groups (75.5% in the non-clinical sample; 79.4% in the clinical sample; Fisher's exact test $p = .32, n = 667$).

All data were collected online, a method shown to give valid data comparable to traditional pen-and-paper administration (Carlbring et al., 2007; Hedman et al., 2010; Holländare, Andersson, & Engström, 2010; Lindner et al., 2013; Thorndike et al., 2009). All items were mandatory and the different instruments were presented in a random order to avoid possible bias.

2.4. Other measures

All participants also answered the QOLI (Frisch et al., 1992) and the Patient Health Questionnaire 9-item (PHQ9; (Kroenke, Spitzer, & Williams, 2001)), a measure of depression symptoms, to enable assessment of concurrent and negative convergent validity of the BBQ (respectively).

2.5. Statistical analyses

Data processing and statistical analyses were performed using the R (3.1.2) statistical environment (R Core Team, 2015). We used parallel analysis (Hayton, Allen, & Scarpello, 2004) to determine number of factors to extract, then performed maximum likelihood explorative factor analysis using the *psych* R package (Revelle, 2015) on the full sample ($n = 731$) to estimate factor loadings of items. Concurrent validity of the BBQ was assessed by regressing BBQ scores on QOLI scores using both ordinary least-squares regression (OLS) and quantile regression (performed using the *quantreg* R package (Koenker, 2015)). Quantile regression estimates predictor impact along any given quantile (i.e. outcome range, e.g. the median) (Cade & Noon, 2003), making it useful to assess the impact of predictors in the extreme ends of the sample, which is often of interest in clinical studies. In this study, we calculated quantile steps of .1, ranging from .1 to .9. Negative convergent validity was investigated by calculating the Pearson's correlation coefficient between the BBQ and the PHQ9. This correlation coefficient was then statistically compared to the QOLI-PHQ9 correlation using the confidence interval (CI) approach (Zou, 2007) (as implemented in the *cocor* R package (Diedenhofen & Musch, 2015)). Responsiveness to change among participants who received treatment was assessed by calculating the standardized response mean, transformed using the pre-post correlation to equal between-group effect sizes (Middel & van Sonderen, 2002).

Internal reliability of the BBQ was assessed by calculating Cronbach's α and inter-item Pearson's correlations using the whole sample. Test-retest reliability was investigated in the undergraduate sample by creating a Bland–Altman plot (Bland & Altman, 1986) and using bootstrapping (5000 resamples) to provide a population estimate of the test-retest score difference. An intra-class correlation coefficient (ICC, agreement type) was also calculated with bootstrapped CI.

The ability of BBQ scores to differentiate between clinical and non-clinical groups was assessed by calculating F -tests and Cohen's d between-group effect sizes, under the assumption that a sample with SAD should present decreased QoL (as manifested in the diagnostic criteria). Further, receiver operating characteristics (ROC) were calculated using the *pROC* R package (Robin et al., 2011) in order to assess the classification accuracy along different discrimination thresholds. The positive state was defined as belonging to the SAD group and having been included in the intervention study, i.e. meeting SAD criteria ($n = 207$), while the undergraduate sample ($n = 163$) served as negative cases. Sensitivity and specificity were weighted equally to calculate a maximizing cut-off. CIs were calculated using bootstrapping with 5000 resamples. Area under the curve (AUC) was computed for both the BBQ and QOLI, and statistically compared using bootstrapping.

2.6. Ethics

Data from the clinical sample were collected as part of a larger intervention study that was approved by the Stockholm Regional Ethics Board (2014/680-31/3). Data collection from the non-clinical sample received departmental approval and was collected anonymously.

3. Results

3.1. Effects of demographics

There was no effect of sex on BBQ scores ($F[1,665] = .12, p = .734$), nor any correlations with age, either in the clinical ($r = .03, n = 504, p = .575$) or the non-clinical group ($r = .12, n = 161, p = .121$). BBQ total scores followed an approximate normal distribution (see Supplementary figure S1).

3.2. Factor structure

Inspection of the parallel analysis scree plot for the BBQ revealed a marked, consistent drop in eigenvalues following the first factor and a large difference vis-à-vis the resampling-derived eigenvalue only for the first factor, suggesting a satisfactory and interpretable unifactorial solution (see Figure 1). When extracting a single factor, item factor loadings ranged from .51 to .71 (Leisure: .6, View on life: .63, Creativity: .52, Learning: .57, Friends and Friendship .51, and View on self: .72). The root mean square of the residuals (RMSR) for this factor solution was calculated to .06.

Extracting a two-factor solution (with varimax rotation) resulted in an RMSR of .03, with the Friends and Friendship-item loading onto a separate factor (loading .85; see Supplementary figure S2). However, performing factor analyses sample-wise revealed more evenly distributed factor loading across the two factors, suggesting that the unique Friends and Friendship factor found when combining samples was likely an artifact of combining two samples with large differences in scores on specific items.

3.3. Convergent validity and sensitivity to change

Quantile regression showed significant, positive associations between BBQ and QOLI scores at all investigated quantiles. Inspection of the quantile function plot (see Figure 2) revealed

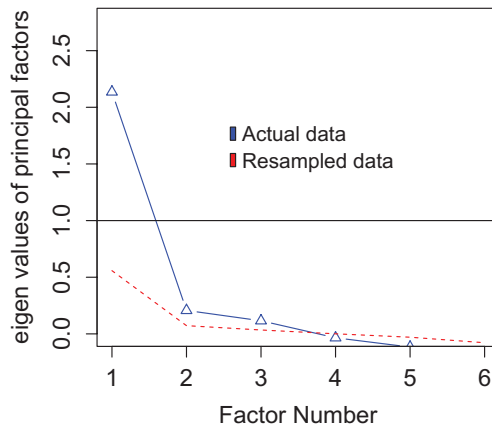


Figure 1. Parallel analysis scree plot.

fairly uniform coefficients at low-to-mid quantiles, with a trend of decreasing coefficients for higher quantiles. The OLS-derived parameter estimate was calculated to $B = .065$, with a BBQ-QOLI correlation coefficient of $r = .65$ ($p < .001$) and a corresponding explained variance of $r^2 = .42$. Both BBQ scores and QOLI scores correlated with PHQ9 scores ($r_{\text{BBQ-PHQ9}} = -.51$, $n = 731$, $p < .001$; $r_{\text{QOLI-PHQ9}} = -.60$, $n = 731$, $p < .001$). The 95% CI for the difference in correlation coefficients ($r_{\Delta} = -.08$ [95% CI: $-.13$ to $-.03$]) did not include zero, indicating a statistical difference between the coefficients. SRM for the BBQ was calculated to 0.43 ($n = 123$), corresponding to a near-medium-sized effect of treatment. The corresponding SRM for the QOLI was $.59$.

3.4. Internal and test-retest reliability

The Cronbach's α of BBQ was $.76$ ($n = 731$), and was similar across the two subsamples (undergraduate $\alpha = .69$, SAD $\alpha = .71$). BBQ inter-item correlations ranged from $r = .20$ to $.51$ (all $p < .001$; see Figure 3).

ICC for the BBQ was calculated to $.82$ (95% CI: $.75$ – $.89$) indicating high test-retest reliability. Bootstrap estimates of the population mean test-retest difference was calculated to $.58$, with a 95% CI covering zero ($-.91$ – 2.08). A Bland–Altman test-retest plot of average score versus difference in scores is presented as Figure 4 and revealed no systematic difference between test and retest scores.

3.5. Classification accuracy

The clinical group scored significantly lower on all BBQ items (see Table 1). There was a large difference in BBQ total scores between the clinical and non-clinical group (Cohen's $d = 1.26$ [95% CI: 1.07 – 1.45]). The ROC curves for the BBQ and QOLI are presented in Figure 5. AUC for the BBQ was 79.65% (95% CI: 75.12 – 84.18%). Maximized classification accuracy was reached at a score of 52 , with a sensitivity of $.75$ (95% CI: $.70$ – $.81$) and a specificity of $.71$ (95% CI: $.63$ – $.78$). The comparable AUC obtained with the QOLI was 67% (95% CI: 61.43 – 72.56%) with a sensitivity of $.73$ (95% CI: $.66$ – $.78$) and specificity $.58$.

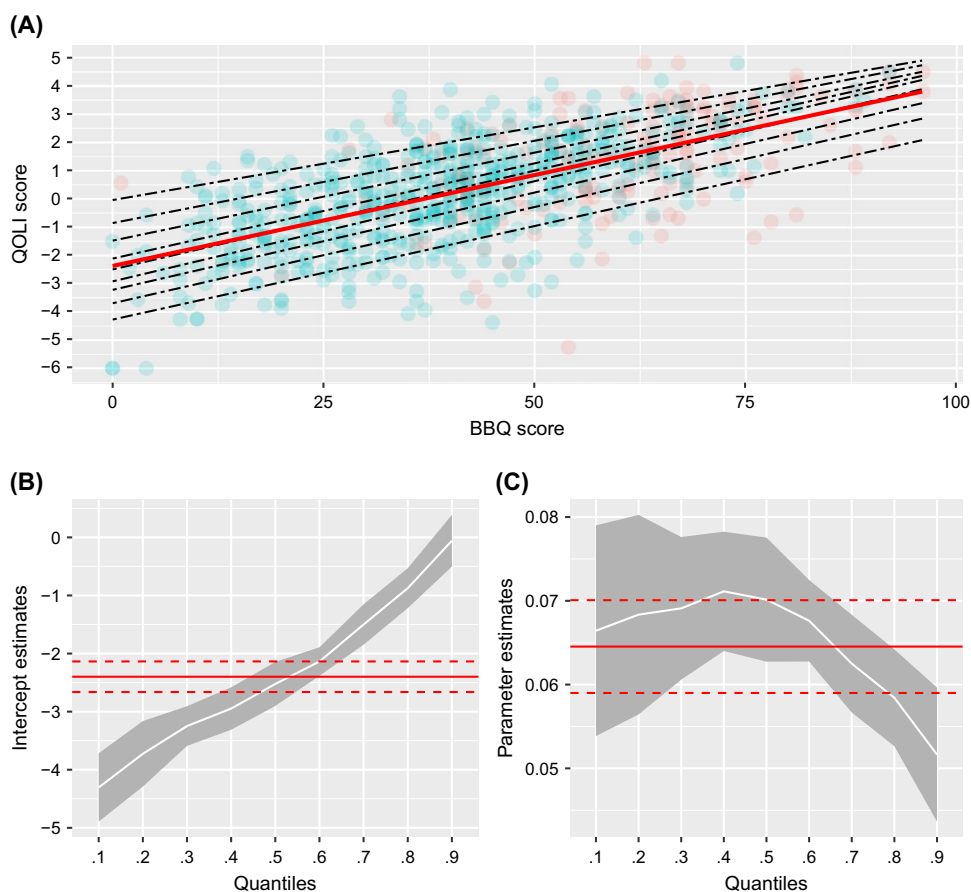


Figure 2. Regression results (A) Full sample scatter plot (SAD group in blue, undergraduate in red) along with OLS (in red) and quantile (in black, dashed) regression lines. (B) Intercept plot for quantiles (white line; CIs in dark gray). Red line: OLS intercept (with CI). (C) Parameter (slope) plot for quantiles (white lines; CI in dark gray). Red line: OLS regression slope (with CI).

(95% CI: .50–.66) at the optimal cut-off score threshold of 1.41. A bootstrap test between the AUCs of the BBQ and QOLI revealed a significant difference ($p < .001$), indicating that BBQ was superior to the QOLI in classification of cases.

3.6. Assessment of alternative total score

To assess the impact of the specific weighting procedure chosen for the BBQ total score, we repeated our primary, scale-level analyses of concurrent validity, test-retest reliability and classification accuracy using an alternative total score (BBQ*), defined as the regular BBQ score divided by the total sum of importance ratings. Rather than having importance ratings contribute directly to the total score, this alternative total score in effect adjusts satisfaction ratings according to importance ratings. Two cases were excluded from comparative analyses due to having importance sums equal to zero. The BBQ*–QOLI correlation was calculated to $r = .68$ (compared to $r = .64$ when using the regular BBQ total score), with

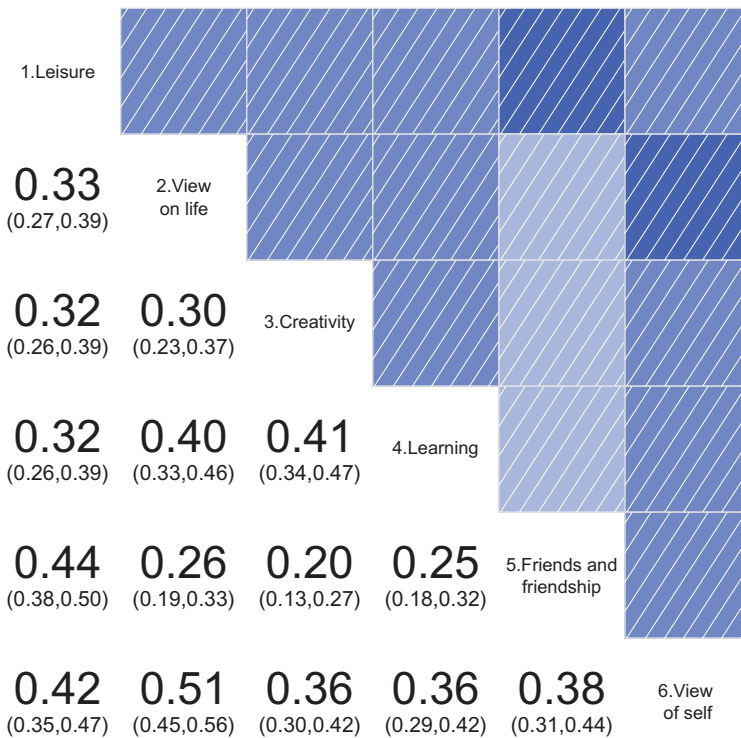


Figure 3. BBQ inter-item correlation matrix Pearson correlation coefficients (with 95% CIs) calculated using the full sample ($n = 731$). All correlations significant ($p < .001$).

a BBQ* test-retest ICC of .86 (compared to .82 when using the regular BBQ total score). Comparing the SAD and healthy groups, the between-group effect size was calculated to Cohen's $d = 1.22$ (compared to $d = 1.26$ when using the regular BBQ total score).

4. Discussion

There is growing interest in measurement of subjective QoL as a complement to symptom measures in clinical psychology and psychiatry. The BBQ was developed to be a brief, easy-to-use and freely accessible self-rating scale of QoL concordant with the overall life satisfaction conceptualization thereof. The results of this initial validation study revealed that the BBQ is a valid and reliable measure of QoL that is sensitive to difference between clinical and non-clinical groups, with a psychometric performance on par with longer, more complex QoL instruments.

Men and women did not differ in BBQ scores and scores did not correlate with age, suggesting that item selection for the BBQ was not biased toward any sex or age group. The BBQ was designed to measure a single latent factor corresponding to overall subjective QoL, and results of the parallel analysis strongly suggested a single-factor solution. All six item loadings were above .50, suggesting a solid factor (Costello & Osborne, 2005). However, future studies featuring larger, more diverse samples will be required to provide a better estimate of the factor structure and whether the structure differs between clinical groups.

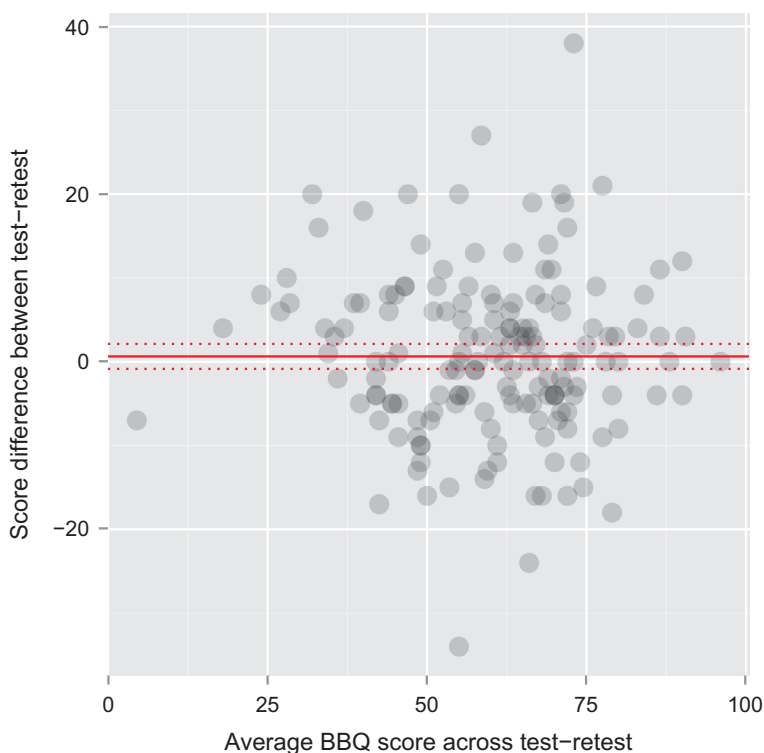


Figure 4. BBQ test-retest Bland–Altman plot Red solid line shows mean test-retest score difference (dotted lines are 95% bootstrapped confidence intervals).

Table 1. BBQ item and total scores.

Item	Non-clinical sample		Clinical sample		Total sample		Statistics
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
1. Leisure time	10.36	3.92	7.07	4.33	7.80	4.46	$F[1,729] = 76.36, p < .001$
2. View on life	10.52	3.91	7.16	4.60	7.91	4.67	$F[1,729] = 71.90, p < .001$
3. Creativity	7.53	4.44	6.13	4.56	6.44	4.57	$F[1,729] = 12.19, p = .001$
4. Learning	10.39	4.39	6.86	4.50	7.64	4.71	$F[1,729] = 78.71, p < .001$
5. Friends and friendship	10.87	4.62	5.99	4.86	7.08	5.22	$F[1,729] = 129.79, p < .001$
6. View of self	10.45	4.24	5.53	4.22	6.63	4.69	$F[1,729] = 171.6, p < .001$
Total Score	60.12	16.01	38.74	17.23	43.50	19.15	$F[1,729] = 201.11, p < .001$

As expected, BBQ scores correlated positively with QoL scores obtained with the QOLI, and correlated negatively with depression ratings obtained using the PHQ9. There was a large (Cohen's $d = 1.26$) difference in BBQ scores between the clinical and non-clinical samples, with high classification accuracy found at a cut-off of 52 points (near the theoretical half of 96 points). In line with previous findings revealing lower QoL in individuals with SAD (Hambrick, Turk, Heimberg, Schneier, & Liebowitz, 2003), these results show that the BBQ is sensitive to QoL impairments prevalent in psychiatric populations. The results of this initial psychometric evaluation also suggests that the BBQ is sensitive to improvements

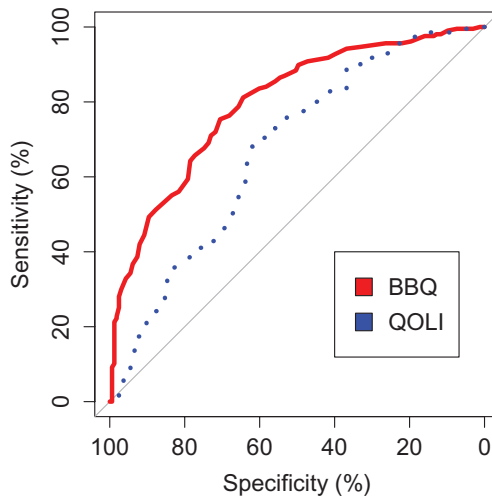


Figure 5. Classification accuracy of the BBQ and QOLI in identifying cases of SAD.

in QoL following an intervention, although data from more intervention studies will be required to draw more thorough conclusion as to sensitivity to change.

The BBQ showed good reliability in terms of both internal and test-retest. The Bland–Altman plot revealed no pattern of systematic differences between test and retest scores. Likewise, the ICC of .82 was high. While the Cronbach's alpha measure of internal consistency of .76 may appear merely satisfactory in light of the often reported, yet misquoted (Lance, Butts, & Michels, 2006) threshold of .70, this interpretation guideline is well argued (Briggs & Cheek, 1986) to be inappropriately applied in cases such as the BBQ. Scales with few items (such as the BBQ) will tend to have lower alphas, in which case inter-item correlations, preferably between .20 and .40, are considered more appropriate measure of scale reliability (Briggs & Cheek, 1986). Inter-item correlations for the BBQ found in the current study ranged from $r = .20$ to $.51$ and were thus within the appropriate range.

We used the well-established QOLI scale to compare the psychometric performance of the BBQ with. The correlation between the QOLI and PHQ-9 was statistically higher than that between the BBQ and PHQ-9. While this suggests that the QOLI better measures QoL impairments associated with depressive symptoms, the superior accuracy of the BBQ in classifying the clinical and non-clinical samples suggests that the BBQ may better capture generic QoL impairments not specifically associated with depression symptoms. Unlike the QOLI, the BBQ was designed to be a unidirectional measure of QoL, i.e. individual items cannot have a negative impact on the total score. The BBQ response format does not feature negative numbers that correspond to negative life area satisfaction, in contrast to the QOLI. Using the BBQ's sum-based approach for total score calculation also ensures that even comparably lower item scores will have a positive impact on the total score. This is in contrast to the QOLI's mean-based approach for calculating final score, where comparably lower scores on some items will have a negative impact on the total score. The choice of a unidirectional approach for the BBQ was primarily a conceptual one, although in this initial psychometric evaluation, we found no empirical indications of a limitation in this approach to QoL measurement. Distribution plots for the clinical sample showed no floor-effect and

the quantile function plot revealed an association between QOLI and BBQ scores that was significant and uniform across the low and mid-ranges of QOLI scores. At higher quantiles however, there was a pattern of decreasing coefficients. Thus, for individuals scoring very high on the QOLI, increases in BBQ scores appear to correspond to relatively low changes in QOLI scores. Hence, the BBQ and QOLI appear to measure similar, but not identical concepts.

Weighting satisfaction ratings in QoL measures according to rated importance is a controversial topic, with some studies reporting no additional gains beyond satisfaction-only measures (Hsieh, 2011). Different weighting methods have different underlying conceptual assumptions, yet may have little impact on the final psychometric properties used for evaluation. Multiplying satisfaction and importance ratings for each life domain, scored using the same item response format, and summing products, as done for the BBQ, implies that operand order is unimportant: a respondent who rates Satisfaction = 4 and Importance = 1 on any given life domain will receive the same item product (score) as a respondent who rates Satisfaction = 1 and Importance = 4. Thus, although the manifestations differ from a phenomenological perspective (being very satisfied with a domain of low importance, compared to not being satisfied with a domain of high importance), their contribution to the total score is the same. However, life domains were included in the BBQ based on solid empirical findings on their importance for a single, latent QoL factor. We believe this supports defining high QoL as rating not only high satisfaction with these specific domains, but also rating them as highly important. The good psychometric properties of the BBQ found and reported in this first psychometric study support the utilized scoring method. In addition, we found only negligible differences with regards to concurrent validity, test-retest reliability and classification accuracy between the regular BBQ total score and an alternative, weighted arithmetic mean total score. Considering the practical inconvenience that the alternative scoring procedure introduces, using the regular weighted product sum total score is recommended.

In this first validation study, we included only one clinical sample ($n = 568$ seeking treatment for SAD) and a non-clinical sample that was comprised of $n = 163$ psychology undergraduate students. Although suitable for the primary research questions and statistical methods of the current study, and in line with samples used in other QoL psychometric studies (Burckhardt & Anderson, 2003; Frisch et al., 1992), larger and more heterogeneous samples will be necessary to develop disorder-specific norms. Since QoL ratings differ somewhat across psychiatric disorder groups (Lindner et al., 2013), future research should develop disorder-specific norms and clinically relevant cut-offs for the BBQ. The BBQ is currently available in over 30 languages; whether the psychometric properties are similar in all language versions and cultures remain to be investigated (Guillemin, Bombardier, & Beaton, 1993). Future studies should also investigate whether the BBQ is as valid and reliable when used outside psychiatry and psychology, e.g. in fields of medicine where patients experience reduced QoL without meeting criteria for psychiatric disorders.

4.1. Availability

BBQ is freely available for non-commercial use in research and clinical practice, and can be downloaded at <http://www.bbqscale.com>. At time of writing, the BBQ has been translated by professional translators into the following languages, freely available for download:

Albanian, Arabic, Bosnian, Bulgarian, Czech, Danish, Dari, Dutch, English, Estonian, Farsi, Finnish, French, German, Greek, Hungarian, Icelandic, Italian, Latvian, Lithuanian, Norwegian, Polish, Portuguese, Rumanian, Russian, Slovakian, Slovenian, Somali, Sorani, Spanish, Swedish, and Thai.

5. Conclusions

The Brunnsviken Brief Quality of life scale measures importance-adjusted satisfaction across six life areas, corresponding to a unified, latent quality of life factor. Using traditional psychometric tools, we show that the BBQ displays high concurrent and convergent validity, high internal and test-retest reliability, good ability to differentiate between clinical and healthy groups, and is sensitivity to change, making it suitable for as a screening and outcome measure in psychological and psychiatric research and practice.

Acknowledgments

Preliminary results of this study formed part of the MSc thesis of author OF and were presented at the Annual Congress of the Swedish Behavior Therapy Association in March 2015 and at the third meeting of the European Society for Research on Internet Interventions in September 2015. The authors wish to thank Professor Matteo Bottai (Karolinska Institutet, Sweden) for statistical consultation.

Disclosure statement

All authors report no conflicts of interest.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, TX: American Psychiatric Publishing.
- Berzon, R., Donnelly, M. A., Simpson, R. L., Simeon, G. P., & Tilson, H. H. (1995). Quality of life bibliography and indexes: 1994 update. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 4, 547–569. [10.1007/BF00634750](https://doi.org/10.1007/BF00634750)
- Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1, 307–310. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2868172>
- Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, 54, 106–148. [10.1111/j.1467-6494.1986.tb00391.x](https://doi.org/10.1111/j.1467-6494.1986.tb00391.x)
- Burckhardt, C. S., & Anderson, K. L. (2003). The quality of life scale (QOLS): Reliability, validity, and utilization. *Health and Quality of Life Outcomes*, 1, 60. [10.1186/1477-7525-1-60](https://doi.org/10.1186/1477-7525-1-60)
- Cade, B. S., & Noon, B. R. (2003). A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, 412–420. [10.1890/1540-9295\(2003\)001\[0412:AGITQR\]2.0.CO;2](https://doi.org/10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2)
- Carlbring, P., Brunt, S., Bohman, S., Austin, D., Richards, J., Öst, L. -G., & Andersson, G. (2007). Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. *Computers in Human Behavior*, 23, 1421–1434. [10.1016/j.chb.2005.05.002](https://doi.org/10.1016/j.chb.2005.05.002)
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10.
- Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, 10, e0121945. [10.1371/journal.pone.0121945](https://doi.org/10.1371/journal.pone.0121945)

- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of personality assessment*, 49, 71–75.
- Endicott, J., Nee, J., Harrison, W., & Blumenthal, R. (1993). Quality of life enjoyment and satisfaction questionnaire: A new measure. *Psychopharmacology Bulletin*, 29, 321–326. [10.1111/j.1365-2850.2011.01735.x](https://doi.org/10.1111/j.1365-2850.2011.01735.x)
- EuroQol. (1990). EuroQol – A new facility for the measurement of health-related quality of life. *Health Policy*, 16, 199–208. [10.1016/0168-8510\(90\)90421-9](https://doi.org/10.1016/0168-8510(90)90421-9)
- Ferrans, C. E., & Powers, M. J. (1985). Quality of life index: Development and psychometric properties. *ANS. Advances in Nursing Science*, 8, 15–24.
- Flanagan, J. C. (1978). A research approach to improving our quality of life. *American Psychologist*, 33, 138–147. [10.1037//0003-066X.33.2.138](https://doi.org/10.1037//0003-066X.33.2.138)
- Frisch, M. B. (1998). Quality of life therapy and assessment in health care. *Clinical Psychology: Science and Practice*, 5, 19–40. [10.1111/j.1468-2850.1998.tb00132.x](https://doi.org/10.1111/j.1468-2850.1998.tb00132.x)
- Frisch, M. B., Cornell, J., Villanueva, M., & Retzlaff, P. J. (1992). Clinical validation of the quality of life inventory. A measure of life satisfaction for use in treatment planning and outcome assessment. *Psychological Assessment*, 4, 92–101. [10.1037//1040-3590.4.1.92](https://doi.org/10.1037//1040-3590.4.1.92)
- Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *Journal of Clinical Epidemiology*, 46, 1417–1432. [10.1016/0895-4356\(93\)90142-N](https://doi.org/10.1016/0895-4356(93)90142-N)
- Hambrick, J. P., Turk, C. L., Heimberg, R. G., Schneier, F. R., & Liebowitz, M. R. (2003). The experience of disability and quality of life in social anxiety disorder. *Depression and Anxiety*, 18, 46–50. [10.1002/da.10110](https://doi.org/10.1002/da.10110)
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York, NY: Guilford Press. Retrieved from <http://books.google.com/books?hl=en&lr=&id=CTgSzAdxc8cC&pgis=1>
- Hays, R. D., Sherbourne, C. D., & Mazel, R. M. (1993). The RAND 36-item health survey 1.0. *Health Economics*, 2, 217–227. [10.1002/hec.4730020305](https://doi.org/10.1002/hec.4730020305)
- Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, 7, 191–205. [10.1177/1094428104263675](https://doi.org/10.1177/1094428104263675)
- Hedman, E., Ljótsson, B., Rück, C., Furmark, T., Carlbring, P., Lindefors, N., & Andersson, G. (2010). Internet administration of self-report measures commonly used in research on social anxiety disorder: A psychometric evaluation. *Computers in Human Behavior*, 26, 736–740. [10.1016/j.chb.2010.01.010](https://doi.org/10.1016/j.chb.2010.01.010)
- Hofmann, S. G., Wu, J. Q., & Boettcher, H. (2014). Effect of cognitive-behavioral therapy for anxiety disorders on quality of life: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 82, 375–391. [10.1037/a0035491](https://doi.org/10.1037/a0035491)
- Holländare, F., Andersson, G., & Engström, I. (2010). A comparison of psychometric properties between internet and paper versions of two depression instruments (BDI-II and MADRS-S) administered to clinic patients. *Journal of Medical Internet Research*, 12, e49. [10.2196/jmir.1392](https://doi.org/10.2196/jmir.1392)
- Hsieh, C. -M. (2011). Importance is not unimportant: The role of importance weighting in QOL measures. *Social Indicators Research*, 109, 267–278.
- Katschnig, H. (2006). Quality of life in mental disorders: Challenges for research and clinical practice. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, 5, 139–145. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1636133&tool=pmcentrez&rendertype=abstract>
- Koenker, R. (2015). Quantreg: Quantile regression. Retrieved from <http://cran.r-project.org/package=quantreg>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16, 606–613.
- Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, 202–220. [10.1177/1094428105284919](https://doi.org/10.1177/1094428105284919)

- Lindner, P., Andersson, G., Öst, L.-G., & Carlbring, P. (2013). Validation of the Internet-administered quality of life inventory (QOLI) in different psychiatric conditions. *Cognitive Behaviour Therapy*, 42, 315–327. [10.1080/16506073.2013.806584](https://doi.org/10.1080/16506073.2013.806584)
- Middel, B., & van Sonderen, E. (2002, December). Statistical significant change versus relevant or important change in (quasi) experimental design: Some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. *International Journal of Integrated Care*, 2, e15. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1480399&tool=pmcentrez&rendertype=abstract>
- Miloff, A., Marklund, A., & Carlbring, P. (2015). The challenger app for social anxiety disorder: New advances in mobile psychological treatment. *Internet Interventions*. [10.1016/j.invent.2015.08.001](https://doi.org/10.1016/j.invent.2015.08.001)
- R Core Team. (2015). *R: A language and environment for statistical computing*, 2, 382–391. Vienna. Retrieved from <http://www.r-project.org/>
- Revelle, W. (2015). *psych: Procedures for psychological, psychometric, and personality research*. Evanston, IL. Retrieved from <http://cran.r-project.org/package=psych>
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12, 77. [10.1186/1471-2105-12-77](https://doi.org/10.1186/1471-2105-12-77)
- Sheehan, D. V., Harnett-Sheehan, K., & Raj, B. A. (1996). The measurement of disability. *International Clinical Psychopharmacology*, 11(Suppl. 3), 89–95.
- Thorndike, F. P., Carlbring, P., Smyth, F. L., Magee, J. C., Gonder-Frederick, L., Ost, L.-G., & Ritterband, L. M. (2009). Web-based measurement: Effect of completing single or multiple items per webpage. *Computers in Human Behavior*, 25, 393–401. [10.1016/j.chb.2008.05.006](https://doi.org/10.1016/j.chb.2008.05.006)
- Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, 12, 399–413. [10.1037/1082-989X.12.4.399](https://doi.org/10.1037/1082-989X.12.4.399)