Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher’s definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher’s website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
An Investigation into the Factor Structure of the Cognitive Therapy Scale – Revised (CTS-R) in a CBT Training Sample

Sarah Beale (ORCiD: https://orcid.org/0000-0002-4038-7460)
Silia Vitoratou (ORCiD: https://orcid.org/0000-0003-1209-8167)
Sheena Liness (ORCiD: https://orcid.org/0000-0003-1141-9782)

* Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, UK

* Corresponding author.
Email address: sarah.beale@kcl.ac.uk
Postal address: Department of Psychology
PO77 Henry Wellcome Building, IoPPN/King’s College London
De Crespigny Park, London SE5 8AF
Abstract

**Background:** Effective monitoring of cognitive behaviour therapy (CBT) competence depends on psychometrically robust assessment methods. While the UK Cognitive Therapy Scale – Revised (CTS-R; Blackburn et al. 2001) has become a widely-used competence measure in CBT training, practice, and research, its underlying factor structure has never been investigated.

**Aims:** This study aimed to present the first investigation into the factor structure of the CTS-R based on a large sample of postgraduate CBT trainee recordings.

**Method:** Trainees (n=382) provided 746 mid-treatment audio recordings for depression (n =373) and anxiety (n =373) cases scored on the CTS-R by expert markers. Tapes were split into two equal samples counterbalanced by diagnosis and with one tape per trainee. Exploratory factor analysis was conducted. The suggested factor structure and a widely-used theoretical two-factor model were tested with confirmatory factor analysis. Measurement invariance was assessed by diagnostic group (depression versus anxiety).

**Results:** Exploratory factor analysis suggested a single-factor solution (98.68% explained variance), which was supported by confirmatory factor analysis. All twelve CTS-R items were found to contribute to this single factor. The univariate model demonstrated full metric invariance and partial scalar invariance by diagnosis, with one item (Item 10 – Conceptual Integration) demonstrating scalar non-invariance.

**Conclusions:** Findings indicate that the CTS-R is a robust homogenous measure and do not support division into the widely used theoretical generic versus CBT-specific competency subscales. Investigation into the CTS-R factor structure in other populations is warranted.

**Keywords:** cognitive behaviour therapy (CBT), competence, factor structure, training, depression, anxiety
Introduction

The large-scale implementation of evidence-based psychotherapies such as CBT relies on accurate measurement of therapy competence. Effective CBT delivery is widely regarded to depend on the related constructs of therapist adherence – implementation of the correct therapeutic procedures – and therapist competence – the skillful delivery of these interventions (Barber et al. 2003; Blackburn et al. 2001; Fairburn & Cooper, 2011; Sharpless and Barber 2009). Adherence is widely incorporated into assessments of competence as it is regarded as a necessary element of competent therapy delivery (Fairburn & Cooper, 2011) and the two constructs are strongly correlated (Barber et al., 2003). Gold-standard assessments of therapist competence for training and evaluation involve an assessor observing a therapist delivering CBT and rating their performance on a competence assessment scale (Clark, 2018; McHugh & Barlow, 2010; Rosen et al., 2016). Evaluating the psychometric properties of these competence rating scales is consequently important given their key role in CBT training and delivery, and as their widespread use in training and clinical practice may influence therapists’ conceptualisation of competence.

The Cognitive Therapy Scale - Revised (CTS-R), has become an established measure of competence in CBT training (Branson, Shafran, & Myles, 2015; Keen & Freeston, 2008; Liness et al., 2018; Liness et al., 2019), routine practice and clinical trials (e.g. Freeman et al., 2015; Wilkinson et al., 2009). Adapted from the related but distinct original Cognitive Therapy Scale (CTS; Young & Beck, 1980), the CTS-R comprises twelve items that assess various elements of skilful therapy delivery, including interpersonal and time management skills, adherence to CBT protocols, and effective use of CBT-specific session features. Competence is scored on a scale between 0 and 6 for each item (0=non-competent, 3=competent, 6= expert).

The measure has demonstrated promising psychometric properties, namely high internal consistency (α range= .75-.97; Blackburn et al., 2001; Kazantzis et al., 2018; Reichelt, James, & Blackburn, 2003), responsiveness to increased competence with training (Branson et al., 2015; Liness et al., 2019), and some evidence of clinical predictive validity in depression (Liness et al., 2019). While
inter-rater reliability estimates vary (Blackburn et al., 2001; James et al., 2001; Kazantzis et al., 2018), agreement improves with rater training (ICC=.38 to .76, Gordon, 2007; \( r = .44 \) to .67, Reichelt et al., 2003) and recent estimates are high (Finn’s \( r = .88 \), Kazantzis et al., 2018). The CTS-R is consequently a measure of successful CBT training and an important tool to guide supervision and clinical practice. Both the CTS-R and the original CTS are often divided into two subscales that are believed to reflect two hypothetical factors underlying competence: general interpersonal competence (CTS-R Items 1-5) and skill at delivering CBT-specific interventions (CTS-R Items 6-12) (James, Blackburn, & Reichelt, 2001). The factor structure of the CTS-R, however, has never been empirically investigated.

Factor analytic investigation into the original CTS has generated surprising findings. A two-factor structure (“quality of therapy” and “session structure”) was indicated based on depression trial data (Vallis, Shaw, & Dobson, 1986). Notably, this did not correspond with the widely-used hypothetical two-factor structure previously discussed (general vs CBT-specific competencies), nor any of the similar but less common hypothetical three subscale divisions previously proposed for the CTS (McManus et al, 2010 & Westbrook et al., 2008: interpersonal effectiveness, general therapeutic skills, and cognitive-behavioural skills; Simons et al., 2010: general therapy skills, CBT skills, and structure]). A recent factor analysis of the CTS in community clinicians treating a range of disorders (Affrunti & Creed, 2019) supported neither Vallis et al.’s (1986) two-factor solution nor the hypothetical two-factor solution, and instead endorsed distinct structures in CBT for adults (two factor solution) and children (one factor solution). The CTS-R is also divided into general versus CBT-specific competency subscales (James et al., 2001) despite factor analyses of the CTS demonstrating that the underlying structure of competence measures does not correspond to this theoretical understanding and can vary across populations. Consequently, investigation into the factor structure of the CTS-R is warranted.

The variability in CTS factor structures suggested across studies may reflect systematic differences in how assessors conceptualise competence and score the measure in different contexts and with different patient presentations. Vallis et al. (1986) was based on a highly-monitored clinical trial of
depression – the original context for which the CTS was developed – while Affrunti & Creed (2019) was based on community treatment of adults and children with a variety of clinical presentations. Measurement invariance analysis is consequently indicated for factor analysis studies of competence measures, including the CTS-R, where different therapist and/or patient populations are assessed. This is particularly relevant for populations that are commonly collapsed together or compared to one another in studies involving competence assessment (Van de Shoot et al., 2015), e.g. patient diagnostic groups which are often combined in studies of therapist competence.

This paper aimed to conduct the first investigation into the factor structure of the CTS-R – including exploratory factor analysis and confirmatory factor analysis of the suggested solution and of the hypothetical two-factor model. Data were drawn from postgraduate CBT training for adult depression and anxiety disorders. Given that structure may vary across populations (Affrunti & Creed, 2019), we also assessed measurement invariance across diagnostic groups (depression and anxiety cases).

**Methods**

**Ethics Statement**

This study received ethical clearance from the King’s College London Psychiatry, Nursing, and Midwifery Research Ethics Committee: reference number PNM/12/13-50. Authors abided by the Ethical Principles of Psychologists and Code of Conduct as set out by the BABCP and BPS.

**Participants**

Participants were 382 trainees from the High-Intensity IAPT Postgraduate Diploma at King’s College London. All available data from ten training cohorts (2008-2018) were included in the present study. Reported ethnicity was 81% (n=311) white and 19% (n=71) Black, Asian, or minority ethnic. Gender distribution was 76% (n=292) female and 24% (n=90) male. Median age was 32 years (IQR=7). Trainees’ professions were psychological wellbeing practitioner (38%, n=145), clinical psychologist (26%, n=101), counselling psychologist (12%, n=46), mental health nurse (9%, n=36),
counsellor/psychotherapist (8%, n=31), occupational therapist (3%, n=12), social worker (1%, n=4) and other (2%, n=7).

**Measures**

Therapist competence was assessed using the Cognitive Therapy Scale – Revised (CTS-R; Blackburn et al., 2001). The CTS-R comprises 12 items (detailed in Table 1 of the results) which assess general therapeutic skills (Items 1-5) and CBT-specific competencies (Items 6-12). Items are rated between 0 (non-competent) and 6 (expert), producing a total score out of 72. The recommended competence threshold is item mean score ≥3 or total score ≥ 36 (James et al., 2001). Raters in the present study were course staff with training and experience in delivering and assessing CBT (see Liness et al., 2019). Estimates of inter-rater reliability based on randomly-selected tapes from the current sample were examined separately as part of a course assessment study (Liness et al., 2019) and were excellent between internal course raters (one-way random single measures ICC with absolute agreement (114, 113) =.95) and good between internal and external expert raters (two-way random single measures ICC with absolute agreement (42, 41) =.68). CTS-R internal consistency and factor structure are reported in the results.

**Procedure**

Data were collected as part of course assessments and a larger programme of CBT training research. The one-year course is delivered as part of the UK IAPT initiative and focuses on treatment of adult anxiety disorders and depression in primary care. The content of the training course has been described in detail elsewhere (see Liness et al., 2019). Trainees submit five audio-recordings of mid-treatment therapy sessions. One baseline tape within the first month of training (any depression/ anxiety presentation) and two mid-module tapes (one depression and one anxiety presentation) provide informal feedback for skills development. Two end-of-module tapes (one depression and one anxiety presentation) provide formal examination of therapist competence. While course staff mark assessments due to resource limitations, several steps are taken to blind markers to trainee identity. Tapes are
submitted without identifying information. Course staff also have varying degrees of trainee contact, and tapes are allocated to markers with limited familiarity with the given trainee. To assess the CTS-R factor structure at the end of training (i.e. factor structure of the CTS-R for therapists entering the workforce), we included available end-of-training anxiety and depression tapes in the present study. All available recordings were included in the present analysis, including those available for trainees who did not complete the course. A total of 746 recordings were included, comprising 373 depression cases and 373 anxiety disorder cases. Trainees provided informed consent to use data from their course submissions for research purposes. Patient recordings were not directly handled as part of the current study, as only anonymised CTS-R result data obtained from course records were analysed.

**Statistical Analysis**

As trainees submitted both an end-of-training anxiety tape and an end-of-training depression tape, we ensured that the same trainees were not included multiple times in a given analysis by splitting the data by diagnosis for exploratory and confirmatory factor analysis (i.e. a given trainee’s anxiety tape would be included in the exploratory factor analysis and their depression tape in the confirmatory factor analysis or vice versa). This also ensured counterbalance of diagnostic groups. Four trainees provided only one submission (due to dropouts, 2 anxiety and 2 depression) and these were divided between the two datasets. The remaining missing tapes \(n=14\) occurred due to withdrawals/fails and human error. The full exploratory and confirmatory datasets each comprised 373 tapes counterbalanced by diagnosis, with 186 depression tapes and 187 anxiety tapes in the exploratory factor analysis diagnostic subsample, and 187 depression tapes and 186 anxiety tapes in the confirmatory factor analysis diagnostic subsamples.

We performed exploratory factor analysis using principal axis factoring due to skew observed in some variables. As we expected factors underlying competence to be correlated, we decided a-priori to apply oblique factor rotation (Promax) where required. Exploratory factor analysis was conducted on the full sample and on diagnostic subsamples, as the CTS-R is widely used to measure competence regardless
of the presentation being treated, but we were interested in uncovering any potential differences in the factor structure by diagnosis.

We planned to assess goodness of fit via confirmatory factor analysis for a univariate model, the model generated by exploratory factor analysis, and for the hypothetical two-factor model commonly used in the literature (Generic Subscale Items 1-5, and Specific Subscale Items 6-12). Only two models - the univariate and two-factor models – were tested as exploratory factor analysis indicated a univariate model. We assessed the following measures of absolute and relative fit: the relative $\chi^2$ (close fit: values from close to 2 to no larger than 5; Schumacker & Lomax, 2004; Ullman, 2001), the Root Mean Square Error of Approximation (RMSEA; close fit: <.05, adequate fit <0.8; Browne & Cudeck, 1993), the Taylor-Lewis Index (TLI; close fit: >.9; Bentler & Bonett, 1980), and the Comparative Fit Index (CFI; close fit: .9; Bentler, 1990). We compared the fit between models based on their Akaike’s Information Criterion (AIC; Akaike, 1974) and Bayesian Information Criterion (BIC; Schwarz, 1978) values, with better model fit indicated with lower values. Measurement invariance of CTS-R items across diagnostic groups (depression and anxiety) was assessed using the multiple independent groups confirmatory factor analysis model.

**Results**

**Exploratory Factor Analysis**

High Kaiser-Meyer-Olkin values (.92 full and anxiety subsample, .90 depression subsample) and significant Bartlett’s tests (all $p<.001$) suggested that the items were suitably factorable for the full sample, and for the anxiety and depression subsamples.

Exploratory factor analysis (Table 1) suggested a one-factor solution for the full sample and subsamples based on Kaiser’s criterion and Scree plots, respectively explaining 98.68% (full sample), 92.81% (depression), and 95.16% (anxiety) of the variance for the set of twelve variables. Rotation was not performed due to the one-factor structure. Factor loadings (.39-.82) indicated that all items clustered well onto the single factor and were similar for the full sample and subsamples. Communalities range from .15-.60, and were relatively low ($\leq30\%$ common variance) across the full
sample and both subsamples for Items 5 and 8, and additionally for Item 12 on the full and anxiety subsamples and Items 1 and 4 in the depression subsample only.

**INSERT TABLE 1**

**Confirmatory Factor Analysis**

High Kaiser-Meyer-Olkin values (.93 full scale) and significant Bartlett’s test ($p<.001$) suggested that the items were again suitably factorable. Item means and factor loadings for the one and two-factor models are reported in Table 2.

**INSERT TABLE 2**

Confirmatory factor analysis yielded similar indices of relative and absolute fit to the data (Table 3) for both the one-factor and two-factor models. The one-factor model demonstrated lower AIC and BIC values, supporting the univariate structure suggested by the exploratory factor analysis. Relative $\chi^2$ values indicate close fit for either model (one-factor = 3.01, two-factor = 3.04). RMSEA indicated adequate fit (.07 for both models) and TLI values exceeded .9, also indicating adequate fit.

**INSERT TABLE 3**

**Measurement Invariance**

We assessed measurement invariance for the CTS-R items across diagnostic groups using multiple groups confirmatory factor analysis (Table 4), which applies a series of confirmatory factor analysis models with increasingly strict invariance assumptions at each step and compares the fit to the less constrained model using a likelihood ratio test. The CTS-R demonstrated full metric invariance (equal corresponding factor loadings), and partial scalar invariance (equal corresponding thresholds), with only Item 10 (Conceptual Integration) demonstrating scalar non-invariance. Raters appeared to endorse higher scores on Item 10 in anxiety than depression (unstandardized thresholds 3.24 for anxiety and 3.07 for depression). However, as the invariance occurs in one item only, we conclude that the CTS-R is not biased with respect to diagnosis.
Reliability and Item-Level Descriptive Statistics

Descriptive statistics and internal consistency are reported for the exploratory factor analysis dataset in Table 1 and for the confirmatory factor analysis dataset in Table 2. Based on both full samples, the scale demonstrated high Cronbach’s alpha ($\alpha=.88-90$) and strong item-total correlations ($r=.51-.83$) supporting retention of all items. There was no evidence that removing any item would increase Cronbach’s alpha values for either dataset. Item means ranged from 2.84 ($SD=.55$) to 3.72 ($SD=.45$) for the exploratory factor analysis dataset and 2.92 ($SD=.59$) to 3.72 ($SD=.47$) for the confirmatory factor analysis dataset, with overall means of 3.15 ($SD=.35$) and 3.18 ($SD=.37$) respectively.

Discussion

This study is an important first investigation into the factor structure of the CTS-R in a large sample of IAPT CBT trainees treating depression and anxiety disorders. Competence – the capable delivery of appropriate therapeutic procedures (Blackburn et al. 2001; Sharpless and Barber 2009) – is proposed to underlie effective CBT provision. Regular assessment using validated rating scales, such as the CTS-R (Blackburn et al., 2001), is recommended to evaluate and guide trainee and qualified therapists (McHugh & Barlow, 2010; Rosen et al., 2016), and to ensure the availability of quality therapy for patient benefit (Clark, 2018; Strunk et al. 2010). Exploratory factor analysis suggested a one-factor solution including all twelve items, which was supported by confirmatory factor analysis and was largely invariant across diagnostic categories. As in previous studies (Blackburn et al., 2001; Kazantzis et al., 2018; Reichelt et al., 2003), internal consistency was high. Strong item-total correlations and no gain in reliability with the removal of any item supported retention of all twelve CTS-R items.

The CTS-R has been hypothetically divided into two constructs underlying CBT competence, i.e. generic interpersonal/therapeutic competence (Items 1-5) and specific competence at delivering CBT
FACTOR STRUCTURE OF THE CTS-R

interventions (Items 6-12) (James et al. 2001). Consistent with investigations into the factor structure of the original CTS (Affrunti & Creed, 2019; Vallis et al., 1986), findings from the present study did not support division of the scale into subscales based on this hypothetical two-factor model. This departure from the hypothetical two-factor model is notable, considering this theoretical division has been used to conceptualise competence as scored on the CTS-R in both therapist and supervisor training (Liness et al., 2018; Liness et al., 2019; Loades & Armstrong, 2016; Loades & Myles, 2016). The current findings support conceptualisation of trainee competence scored on the CTS-R as a single construct comprising all items of the measure, which appear closely related rather than clustering into separate dimensions of general vs CBT-specific skill. Consequently, CBT training and supervisor/rater training should focus on assessment across all items of the CTS-R, including both generic and CBT-specific elements.

This unidimensional CTS-R structure was also distinct to that of the CTS, which has previously demonstrated varying two factor structures for adult CBT (Affrunti & Creed, 2019; Vallis et al., 1986). The current populations of therapists and patients varied from CTS factor analysis studies, however, with Vallis et al. (1986) based on a clinical trial for depression involving highly experienced and monitored therapists, and Affrunti and Creed (2019) based on a sample of community clinicians treating a range of disorders with varying severity. The present study was based on a sample of CBT trainees at the end of training, who were treating patients with depression and anxiety disorders with limited psychosocial complexity or comorbidity outside of these diagnostic categories. How competence is conceptualised and scored on the CTS-R, and therefore its factor structure, may vary systematically by therapists’ level of experience (i.e. trainee versus experienced therapist), the context in which they practice, and the population being treated. For example, the factor structure of the CTS appeared non-invariant for therapists treating children (one-factor structure) versus adults (two-factor structure) (Affrunti & Creed, 2019). The one-factor CTS-R structure identified in the current study was suggested and confirmed on a large sample of CBT trainees at the end of long-term training. Corresponding investigation into the CTS-R factor structure in other populations and contexts is consequently recommended.
Notably, the scale was not fully invariant across diagnostic categories. Item 10 – ‘Conceptual Integration’ – demonstrated diagnosis-related scalar non-invariance. This item assesses competence at CBT conceptualisation, including whether a presentation-relevant conceptualisation was used and whether it was incorporated into treatment in an appropriate manner. Findings indicated that raters endorsed higher scores for this item in anxiety than depression (unstandardized thresholds respectively 3.24 and 3.07), though these scores were within the boundaries of the same CTS-R competence category (3= ‘competent’). Systematically higher ratings on Item 10 for anxiety presentations may reflect the strong role of disorder-specific conceptualisations in CBT protocols for specific anxiety presentations. However, strong invariance held for the majority (11/12) of CTS-R items, indicating that diagnostic category was not a substantial source of bias in competence ratings.

Additionally, means and standard deviations across items in both the exploratory factor analysis and confirmatory factor analysis samples were similar, tending towards mean values around 3 (‘competent’) with relatively limited variation. Trainees received highly structured and intensive training over one year (please see Liness et al., 2018 for details), likely contributing to relatively homogenous post-training scores. While this indicates an appropriate level of post-training competence across the sample, the relatively homogenous CTS-R scores may have attenuated the correlations between items. However, overall findings are unlikely to be substantially affected as a single-factor structure is likely to remain consistent in a more heterogenous sample due to higher variance. While selection of training cases for end-of-training evaluation occurred prior to recording the sessions submitted in the present study, self-selection of tapes may have influenced overall scores and score variability if trainees selected their strongest session recordings.

Data in the present study were drawn from a single training course and generalizability of this factor structure should be assessed both through confirmation with other training courses and investigation in other populations and contexts, such as assessment of experienced therapists practicing in various settings common to CBT practice. Furthermore, while the CTS-R appeared to be largely invariant across the depression and anxiety diagnostic categories in the present study, these findings may not be
consistent when the scale is used to rate therapists treating other diagnoses or at higher levels of
patient severity or complexity.

While tapes were submitted without identifying information and scored by markers with limited
familiarity with individual trainees, we could not ensure that markers were always blind to trainee
identity. Recognition of a trainees’ identity and recollection of personal factors, e.g. prior training,
interpersonal factors etc., could affect CTS-R ratings. However, this was unlikely to systematically
affect ratings as course markers were allocated tapes from less familiar or unfamiliar trainees and
varied in their contact with trainees overall.

Conclusion

This study adds to the CBT competence literature with a focus on an important missing element in the
psychometric evaluation of the CTS-R. Measures such as the CTS-R influence how competence is
conceptualized in training and practice. Results based on a large sample of CBT trainees demonstrate
the reliability of CTS-R and its component items, and support its continued use to evaluate
competence. The one-factor structure identified in this study indicates that conceptualising and
supporting competence as a holistic construct comprised of both general therapeutic skills and CBT-
specific techniques when using the CTS-R may be beneficial in CBT training. Further investigation is
warranted to confirm the structure underlying the CTS-R and its stability across the various contexts
and populations in which CBT is applied.

Acknowledgements: With many thanks to the course staff, trainees and patients who contributed to
this research.

Funding: This work received no financial support for research, authorship, and/or publication. Silia
Vitoratou is funded by the Biomedical Research Centre for Mental Health at South London and
Maudsley NHS Foundation Trust and King’s College London. The views expressed are those of the
author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social
Care.

Disclosure Statement: The authors report no conflict of interest.
References


Clark, D. M. (2018). Realizing the Mass Public Benefit of Evidence-Based Psychological Therapies:


http://dx.doi.org/10.1037/0022-006X.54.3.381


http://doi.org/10.3389/fpsyg.2015.01064


https://doi.org/10.1017/S1352465808004608

Table 1. Factor Structure and Internal Consistency of the Cognitive Therapy Scale -Revised

<table>
<thead>
<tr>
<th>Item</th>
<th>Full Sample (n=373)</th>
<th>Depression (n=186)</th>
<th>Anxiety (n=187)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>α if item deleted</td>
<td>Item-Total Correlation</td>
</tr>
<tr>
<td>1 – Agenda Setting and Adherence</td>
<td>3.18 (.55)</td>
<td>.88</td>
<td>.61</td>
</tr>
<tr>
<td>2 – Feedback</td>
<td>3.14 (.53)</td>
<td>.87</td>
<td>.74</td>
</tr>
<tr>
<td>3 – Collaboration</td>
<td>3.38 (.56)</td>
<td>.86</td>
<td>.78</td>
</tr>
<tr>
<td>4 – Pacing and Efficient Use of Time</td>
<td>2.95 (.58)</td>
<td>.87</td>
<td>.63</td>
</tr>
<tr>
<td>5 – Interpersonal Effectiveness</td>
<td>3.72 (.45)</td>
<td>.88</td>
<td>.51</td>
</tr>
<tr>
<td>6 – Eliciting of Appropriate Emotional Expression</td>
<td>2.96 (.50)</td>
<td>.87</td>
<td>.62</td>
</tr>
<tr>
<td>7 – Eliciting Key Cognitions</td>
<td>3.18 (.52)</td>
<td>.87</td>
<td>.66</td>
</tr>
<tr>
<td>8 – Eliciting and Planning Behaviours</td>
<td>3.07 (.49)</td>
<td>.88</td>
<td>.53</td>
</tr>
<tr>
<td>9 – Guided Discovery</td>
<td>3.18 (.55)</td>
<td>.86</td>
<td>.78</td>
</tr>
<tr>
<td>10 – Conceptual Integration</td>
<td>3.16 (.55)</td>
<td>.87</td>
<td>.70</td>
</tr>
<tr>
<td>11 – Application of Change Methods</td>
<td>3.13 (.61)</td>
<td>.86</td>
<td>.77</td>
</tr>
<tr>
<td>12 – Homework Setting</td>
<td>2.84 (.55)</td>
<td>.88</td>
<td>.55</td>
</tr>
<tr>
<td>Overall Mean</td>
<td>3.16 (.35)</td>
<td>.88</td>
<td>.55</td>
</tr>
<tr>
<td>Cronbach’s α</td>
<td>.88</td>
<td>.88</td>
<td>.89</td>
</tr>
<tr>
<td>Eigenvalue</td>
<td>4.73</td>
<td>4.67</td>
<td>4.92</td>
</tr>
<tr>
<td>Total variance</td>
<td>98.68%</td>
<td>92.81%</td>
<td>95.16%</td>
</tr>
</tbody>
</table>

Note: $h^2$ = communalities
### Table 2. Confirmatory Factor Analysis and Internal Consistency of the Cognitive Therapy Scale - Revised

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Mean ($SD$)</th>
<th>α if item deleted</th>
<th>Item-Total Correlation</th>
<th>One Factor Model CFA Loading (EFA Loading)</th>
<th>Two Factor Model CFA Loading Factor 1 - Generic</th>
<th>Two Factor Model CFA Loading Factor 2 (Specific)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Agenda Setting and Adherence</td>
<td>3.18 (.55)</td>
<td>.89</td>
<td>.63</td>
<td>1.00 (.55)</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>Feedback</td>
<td>3.18 (.53)</td>
<td>.89</td>
<td>.74</td>
<td>1.25 (.72)</td>
<td>1.25</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>Collaboration</td>
<td>3.41 (.57)</td>
<td>.88</td>
<td>.79</td>
<td>1.44 (.76)</td>
<td>1.45</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>Pacing and Efficient Use of Time</td>
<td>2.92 (.59)</td>
<td>.89</td>
<td>.63</td>
<td>1.06 (.57)</td>
<td>1.06</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>Interpersonal Effectiveness</td>
<td>3.72 (.47)</td>
<td>.89</td>
<td>.56</td>
<td>.82 (.48)</td>
<td>.82</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>Eliciting of Appropriate Emotional Expression</td>
<td>3.04 (.48)</td>
<td>.89</td>
<td>.62</td>
<td>.88 (.58)</td>
<td>-</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td>Eliciting Key Cognitions</td>
<td>3.18 (.52)</td>
<td>.89</td>
<td>.70</td>
<td>1.12 (.63)</td>
<td>-</td>
<td>1.28</td>
</tr>
<tr>
<td>8</td>
<td>Eliciting and Planning</td>
<td>3.12 (.46)</td>
<td>.89</td>
<td>.62</td>
<td>.84 (.48)</td>
<td>-</td>
<td>.96</td>
</tr>
<tr>
<td>9</td>
<td>Guided Discovery</td>
<td>3.19 (.57)</td>
<td>.88</td>
<td>.83</td>
<td>1.56 (.77)</td>
<td>-</td>
<td>1.77</td>
</tr>
<tr>
<td>10</td>
<td>Conceptual Integration</td>
<td>3.15 (.59)</td>
<td>.89</td>
<td>.69</td>
<td>1.26 (.66)</td>
<td>-</td>
<td>1.43</td>
</tr>
<tr>
<td>11</td>
<td>Application of Change Methods</td>
<td>3.14 (.59)</td>
<td>.88</td>
<td>.77</td>
<td>1.39 (.74)</td>
<td>-</td>
<td>1.59</td>
</tr>
<tr>
<td>12</td>
<td>Homework Setting</td>
<td>2.93 (.58)</td>
<td>.89</td>
<td>.64</td>
<td>1.06 (.48)</td>
<td>-</td>
<td>1.20</td>
</tr>
<tr>
<td>Overall Mean</td>
<td>3.18 (.37)</td>
<td>.89</td>
<td>.64</td>
<td>1.06 (.48)</td>
<td>1.20</td>
<td>1.20</td>
<td>1.20</td>
</tr>
<tr>
<td>Cronbach’s α</td>
<td>.90</td>
<td></td>
<td></td>
<td>.77</td>
<td>.85</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** All factor loadings significant $p<.0001$; EFA= exploratory factor analysis; CFA = confirmatory factor analysis
Table 3. *Goodness of Fit for One Factor versus Theoretical Two Factor Model of the Cognitive Therapy Scale -Revised*

<table>
<thead>
<tr>
<th>Model</th>
<th>$X^2$</th>
<th>df</th>
<th>Relative $X^2$</th>
<th>CFI</th>
<th>RMSEA (90% CI)</th>
<th>TLI</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>One Factor</td>
<td>162.38*</td>
<td>54</td>
<td>3.01</td>
<td>.94</td>
<td>.07 (.06-.09)</td>
<td>.93</td>
<td>5465.82</td>
<td>5607.00</td>
</tr>
<tr>
<td>Two Factor</td>
<td>160.97**</td>
<td>53</td>
<td>3.04</td>
<td>.94</td>
<td>.07 (.06-.09)</td>
<td>.93</td>
<td>5466.41</td>
<td>5611.51</td>
</tr>
</tbody>
</table>

*Note: CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; TLI = Taylor-Lewis Index; AIC = Akaike’s Information Criterion; BIC = Bayesian Information Criterion*

*p < .0005, **p < .001*
Table 4. *Multiple Groups Confirmatory Factor Analysis by Diagnosis for the Cognitive Therapy Scale -Revised*

<table>
<thead>
<tr>
<th>Model</th>
<th>Constraints</th>
<th>Non-Invariant Items</th>
<th>Likelihood Ratio $X^2$</th>
<th>df</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Configural)</td>
<td>None</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2 (Metric)</td>
<td>Equal loadings</td>
<td>-</td>
<td>10.92</td>
<td>12</td>
<td>0.54</td>
</tr>
<tr>
<td>3 (Scalar)</td>
<td>Equal loadings and thresholds</td>
<td>Item 10</td>
<td>17.02</td>
<td>11</td>
<td>0.11</td>
</tr>
</tbody>
</table>