

MAIN

# An investigation into the factor structure of the Cognitive Therapy Scale – Revised (CTS-R) in a CBT training sample

Sarah Beale<sup>1,\*</sup>, Silia Vitoratou<sup>2</sup> and Sheena Liness<sup>1</sup>

<sup>1</sup>Institute of Psychiatry, Psychology & Neuroscience, King's College London, London SE5 8AF, UK and <sup>2</sup>Psychometrics & Measurement Lab, Department of Biostatistics and Health Informatics, King's College London, London, UK

\*Corresponding author. Email: [sarah.beale@kcl.ac.uk](mailto:sarah.beale@kcl.ac.uk)

(Received 4 October 2019; revised 24 November 2020; accepted 2 December 2020; first published online 18 January 2021)

## 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 12 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:** anxiety; cognitive behaviour therapy (CBT); competence; depression; factor structure; training

## Introduction

The large-scale implementation of evidence-based psychotherapies such as cognitive behaviour therapy (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 – and the skilful delivery of these interventions (Barber *et al.*, 2003; Blackburn *et al.*, 2001; Fairburn and 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 and 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 and 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 *et al.*, 2015; Keen and 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 and Beck, 1980), the CTS-R consists of 12 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 ( $\alpha$  range = .75–.97; Blackburn *et al.*, 2001; Kazantzis *et al.*, 2018; Reichelt *et al.*, 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 *et al.*, 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 *et al.*, 1986). Notably, this did not correspond with the widely used hypothetical two-factor structure previously discussed (general versus CBT-specific competencies), nor any of the similar but less common hypothetical three subscale divisions previously proposed for the CTS [McManus *et al.* (2010) and 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 and 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 and 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 with one another in studies

involving competence assessment (Van de Schoot *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 and Creed, 2019), we also assessed measurement invariance across diagnostic groups (depression and anxiety cases).

## Method

### 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 consists of 12 items (detailed in Table 1) 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  $\geq 3$  or total score  $\geq 36$  (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

**Table 1.** Factor structure and internal consistency of the Cognitive Therapy Scale-Revised

Item	Full sample ( <i>n</i> = 373)					Depression ( <i>n</i> = 186)					Anxiety ( <i>n</i> = 187)				
	Mean (SD)	$\alpha$ if item deleted	Item-total correlation	Factor loading	$h^2$	Mean (SD)	$\alpha$ if item deleted	Item-total correlation	Factor loading	$h^2$	Mean (SD)	$\alpha$ if item deleted	Item-total correlation	Factor loading	$h^2$
1 – Agenda Setting and Adherence	3.18 (.55)	.88	.61	.55	.31	3.21 (.55)	.87	.57	.50	.25	3.16 (.55)	.88	.65	.60	.37
2 – Feedback	3.14 (.53)	.87	.74	.72	.52	3.17 (.53)	.86	.75	.73	.53	3.12 (.54)	.87	.73	.71	.51
3 – Collaboration	3.38 (.56)	.86	.78	.76	.58	3.40 (.54)	.86	.78	.77	.59	3.36 (.57)	.87	.78	.76	.58
4 – Pacing and Efficient Use of Time	2.95 (.58)	.87	.63	.57	.33	2.95 (.56)	.87	.57	.49	.24	2.94 (.60)	.88	.69	.64	.41
5 – Interpersonal Effectiveness	3.72 (.45)	.88	.51	.48	.23	3.71 (.45)	.87	.53	.50	.25	3.74 (.44)	.89	.51	.47	.22
6 – Eliciting of Appropriate Emotional Expression	2.96 (.50)	.87	.62	.58	.33	2.97 (.46)	.87	.63	.60	.36	2.96 (.53)	.88	.61	.56	.31
7 – Eliciting Key Cognitions	3.18 (.52)	.87	.66	.63	.40	3.23 (.49)	.87	.65	.63	.39	3.14 (.55)	.88	.68	.65	.42
8 – Eliciting and Planning Behaviours	3.07 (.49)	.88	.53	.48	.23	3.08 (.50)	.88	.52	.45	.20	3.06 (.48)	.88	.55	.51	.26
9 – Guided Discovery	3.18 (.55)	.86	.78	.77	.59	3.16 (.50)	.86	.77	.76	.58	3.21 (.59)	.87	.79	.76	.60
10 – Conceptual Integration	3.16 (.55)	.87	.70	.66	.44	3.10 (.51)	.87	.64	.58	.37	3.32 (.58)	.87	.77	.75	.56
11 – Application of Change Methods	3.13 (.61)	.86	.77	.74	.55	2.13 (.56)	.86	.77	.76	.58	3.13 (.65)	.87	.77	.73	.54
12 – Homework Setting	2.84 (.55)	.88	.55	.48	.23	2.92 (.52)	.87	.65	.59	.35	2.76 (.56)	.89	.47	.39	.15
<b>Overall mean</b>	<b>3.16 (.35)</b>					<b>3.17 (.34)</b>					<b>3.15 (.37)</b>				
<b>Cronbach's <math>\alpha</math></b>	<b>.88</b>					<b>.88</b>					<b>.89</b>				
<b>Eigenvalue</b>	<b>4.73</b>					<b>4.67</b>					<b>4.92</b>				
<b>Total variance</b>	<b>98.68%</b>					<b>92.81%</b>					<b>95.16%</b>				

 $h^2$  = communalities.

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, consisting of 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 drop-outs, two anxiety and two 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 consisted of 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 and Lomax, 2004; Ullman, 2001), the root mean square error of approximation (RMSEA; close fit:  $<.05$ , adequate fit  $<.08$ ; Browne and Cudeck, 1993), the Taylor–Lewis index (TLI; close fit:  $>.9$ ; Bentler and 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 12 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 ranged from .15 to .60, and were relatively low ( $\leq 30\%$  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.

### **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.

Confirmatory factor analysis yielded similar indices of relative and absolute fit to the data (Table 3) for both the one- 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.

### **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 (unstandardised 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 data set in Table 1 and for the confirmatory factor analysis data set in Table 2. Based on both full samples, the scale demonstrated high Cronbach's alpha ( $\alpha = .88$ – $.90$ ) and strong item-total correlations (.51–.83) supporting retention of all items. There was no evidence that removing any item would increase Cronbach's alpha values for either data set. Item means ranged from 2.84 ( $SD = .55$ ) to 3.72 ( $SD = .45$ ) for the exploratory factor analysis data set and 2.92 ( $SD = .59$ ) to 3.72 ( $SD = .47$ ) for the confirmatory factor analysis data set, 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

**Table 2.** Confirmatory factor analysis and internal consistency of the Cognitive Therapy Scale-Revised

	Full sample ( <i>n</i> = 373)					
	Mean ( <i>SD</i> )	$\alpha$ if item deleted	Item-total correlation	Standardised factor loading		
				One-factor model CFA loading ( <i>EFA</i> loading)	Two-factor model CFA loading – Factor 1 (generic)	Two-factor model CFA loading – Factor 2 (specific)
1 – Agenda Setting and Adherence	3.18 (.55)	.89	.63	1.00 (.55)	1.00	—
2 – Feedback	3.18 (.53)	.89	.74	1.25 (.72)	1.25	—
3 – Collaboration	3.41 (.57)	.88	.79	1.44 (.76)	1.45	—
4 – Pacing and Efficient Use of Time	2.92 (.59)	.89	.63	1.06 (.57)	1.06	—
5 – Interpersonal Effectiveness	3.72 (.47)	.89	.56	.82 (.48)	.82	—
6 – Eliciting of Appropriate Emotional Expression	3.04 (.48)	.89	.62	.88 (.58)	—	1.00
7 – Eliciting Key Cognitions	3.18 (.52)	.89	.70	1.12 (.63)	—	1.28
8 – Eliciting and Planning Behaviours	3.12 (.46)	.89	.62	.84 (.48)	—	.96
9 – Guided Discovery	3.19 (.57)	.88	.83	1.56 (.77)	—	1.77
10 – Conceptual Integration	3.15 (.59)	.89	.69	1.26 (.66)	—	1.43
11 – Application of Change Methods	3.14 (.59)	.88	.77	1.39 (.74)	—	1.59
12 – Homework Setting	2.93 (.58)	.89	.64	1.06 (.48)	—	1.20
<b>Overall mean</b>	<b>3.18 (.37)</b>					
<b>Cronbach’s <math>\alpha</math></b>	<b>.90</b>				<b>.77</b>	<b>.85</b>

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

Model	$\chi^2$	d.f.	Relative $\chi^2$	CFI	RMSEA (90% CI)	TLI	AIC	BIC
One-factor	162.38**	54	3.01	.94	.07 (.06–.09)	.93	5465.82	5607.00
Two-factor	160.97**	53	3.04	.94	.07 (.06–.09)	.93	5466.41	5611.51

\*\* $p < .0001$ . CFI, comparative fit index; RMSEA, root mean square error of approximation; TLI, Taylor–Lewis index; AIC, Akaike’s information criterion; BIC, Bayesian information criterion.

**Table 4.** Multiple groups confirmatory factor analysis by diagnosis for the Cognitive Therapy Scale-Revised

Model	Constraints	Non-invariant Items	Likelihood ratio $\chi^2$	d.f.	$p$
1 (configural)	None	—	—	—	—
2 (metric)	Equal loadings	—	10.92	12	0.54
3 (scalar)	Equal loadings and thresholds	Item 10	17.02	11	0.11

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 and 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 12 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 12 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 interventions (items 6–12) (James *et al.*, 2001). Consistent with investigations into the factor structure of the original CTS (Affrunti and 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 and Armstrong, 2016; Loades and Myles, 2016). The current findings support conceptualisation of trainee competence scored on the CTS-R as a single construct containing all items of the measure, which appear closely related rather than clustering into separate dimensions of general versus 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 and 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 co-morbidity 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 and 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 (unstandardised thresholds respectively 3.24 and 3.07), although 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), probably 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 generalisability 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 practising 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 conceptualised 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 consisting 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.

**Data availability.** The datasets generated for this study are not currently publicly available as they comprise records of a clinical training course.

**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.

**Conflicts of interest.** The authors report no conflict of interest.

**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.

## References

- Affrunti, N. W., & Creed, T. A. (2019). The factor structure of the Cognitive Therapy Rating Scale (CTRS) in a sample of community mental health clinicians. *Cognitive Therapy and Research*, 43, 642–655. <http://doi.org/10.1007/s10608-019-09998-7>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723. <http://doi.org/10.1109/TAC.1974.1100705>
- Barber, J. P., Liese, B. S., & Abrams, M. J. (2003). Development of the cognitive therapy adherence and competence scale. *Psychotherapy Research*, 13, 205–221. <http://doi.org/10.1093/ptr/kpg019>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238. <http://doi.org/10.1037/0033-2909.107.2.238>
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588–606. <http://doi.org/10.1037/0033-2909.88.3.588>
- Blackburn, I.-M., James, I. A., Milne, D. L., Baker, C., Standart, S., Garland, A., & Reichelt, F. K. (2001). The Revised Cognitive Therapy Scale (CTS-R): psychometric properties. *Behavioural and Cognitive Psychotherapy*, 29, 431–446. <http://dx.doi.org/10.1017/S1352465801004040>
- Branson, A., Shafraan, R., & Myles, P. (2015). Investigating the relationship between competence and patient outcome with CBT. *Behaviour Research and Therapy*, 68, 19–26. <http://dx.doi.org/10.1016/j.brat.2015.03.002>
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (eds), *Testing Structural Equation Models* (pp. 136–62). Newbury Park, CA, USA: Sage.
- Clark, D. M. (2018). Realizing the mass public benefit of evidence-based psychological therapies: the IAPT Program. *Annual Review of Clinical Psychology*, 14, 1–25. <http://doi.org/10.1146/annurev-clinpsy-050817-084833>
- Fairburn, C. G., & Cooper, Z. (2011). Therapist competence, therapy quality, and therapist training. *Behaviour Research and Therapy*, 49, 373–378. <http://doi.org/10.1016/j.brat.2011.03.005>
- Freeman, D., Dunn, G., Startup, H., Pugh, K., Cordwell, J., Mander, H., ... Kingdon, D. (2015). Effects of cognitive behaviour therapy for worry on persecutory delusions in patients with psychosis (WIT): a parallel, single-blind, randomised controlled trial with a mediation analysis. *The Lancet Psychiatry*, 2, 305–313. [http://doi.org/10.1016/S2215-0366\(15\)00039-5](http://doi.org/10.1016/S2215-0366(15)00039-5)
- Gordon, P. K. (2007). A comparison of two versions of the Cognitive Therapy Scale. *Behavioural and Cognitive Psychotherapy*, 35, 343–353. <https://doi.org/10.1017/S1352465806003390>
- James, I. A., Blackburn, I. M., & Reichelt, F. K. (2001). *Manual of the Revised Cognitive Therapy Scale (CTSR)*. Newcastle, UK.
- Kazantzis, N., Clayton, X., Cronin, T. J., Farchione, D., Limburg, K., & Dobson, K. S. (2018). The Cognitive Therapy Scale and Cognitive Therapy Scale-Revised as measures of therapist competence in cognitive behavior therapy for depression: relations with short and long term outcome. *Cognitive Therapy and Research*, 42, 385–397. <http://dx.doi.org/10.1007/s10608-018-9919-4>
- Keen, A. J., & Freeston, M. H. (2008). Assessing competence in cognitive-behavioural therapy. *British Journal of Psychiatry*, 193, 60–64. <http://doi.org/10.1192/bjp.bp.107.038588>
- Liness, S., Beale, S., Lea, S., Byrne, S., Hirsch, C. R., & Clark, D. M. (2018). The sustained effects of CBT training on therapist competence and patient outcomes. *Cognitive Therapy and Research*, 43, 631–641.

- Liness, S., Beale, S., Lea, S., Byrne, S., Hirsch, C. R., & Clark, D. M. (2019). Multi-Professional IAPT CBT training: clinical competence and patient outcomes. *Behavioural and Cognitive Psychotherapy*, 47, 672–685.
- Loades, M. E., & Armstrong, P. (2016). The challenge of training supervisors to use direct assessments of clinical competence in CBT consistently: a systematic review and exploratory training study. *The Cognitive Behaviour Therapist*, 9. <http://doi.org/10.1017/S1754470X15000288>
- Loades, M. E., & Myles, P. J. (2016). Does a therapist's reflective ability predict the accuracy of their self-evaluation of competence in cognitive behavioural therapy? *The Cognitive Behaviour Therapist*, 9. <http://doi.org/10.1017/S1754470X16000027>
- McHugh, R. K., & Barlow, D. H. (2010). The dissemination and implementation of evidence-based psychological treatments: a review of current efforts. *American Psychologist*, 65, 73–84. <http://doi.org/10.1037/a0018121>
- McManus, F., Westbrook, D., Vazquez-Montes, M., Fennell, M., & Kennerley, H. (2010). An evaluation of the effectiveness of diploma-level training in cognitive behaviour therapy. *Behaviour Research and Therapy*, 48, 1123–1132. <https://doi.org/10.1016/j.brat.2010.08.002>
- Reichelt, F. K., James, I. A., & Blackburn, I. M. (2003). Impact of training on rating competence in cognitive therapy. *Journal of Behavior Therapy and Experimental Psychiatry*, 34, 87–99. [https://doi.org/10.1016/S0005-7916\(03\)00022-3](https://doi.org/10.1016/S0005-7916(03)00022-3)
- Rosen, C. S., Matthieu, M. M., Wiltsey Stirman, S., Cook, J. M., Landes, S., Bernardy, N. C., ... & Watts, B. V. (2016). A review of studies on the system-wide implementation of evidence-based psychotherapies for posttraumatic stress disorder in the Veterans Health Administration. *Administration and Policy in Mental Health and Mental Health Services Research*, 43, 957–977. <http://doi.org/10.1007/s10488-016-0755-0>
- Schumacker, R. E., & Lomax, R. G. (2004). *A Beginner's Guide to Structural Equation Modeling* (2nd edn). Mahwah, NJ, USA: Lawrence Erlbaum Associates.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464.
- Sharpless, B. A., & Barber, J. P. (2009). A conceptual and empirical review of the meaning, measurement, development, and teaching of intervention competence in clinical psychology. *Clinical Psychology Review*, 29, 47–56.
- Simons, A. D., Padesky, C. A., Montemaranano, J., Lewis, C. C., Murakami, J., Lamb, K., ... & Beck, A. T. (2010). Training and dissemination of cognitive behavior therapy for depression in adults: a preliminary examination of therapist competence and client outcomes. *Journal of Consulting and Clinical Psychology*, 78, 751. <https://doi.org/10.1037/a0020569>
- Strunk, D. R., Brotman, M. A., DeRubeis, R. J., & Hollon, S. D. (2010). Therapist competence in cognitive therapy for depression: predicting subsequent symptom change. *Journal of Consulting and Clinical Psychology*, 78, 429–437. <http://doi.org/10.1037/a0019631>
- Ullman, J. B. (2001). Structural equation modeling. In B. G. Tabachnick & L. S. Fidell (eds), *Using Multivariate Statistics* (4th edn, pp. 653–771). Needham Heights, MA, USA: Allyn & Bacon.
- Vallis, T. M., Shaw, B. F., & Dobson, K. S. (1986). The Cognitive Therapy Scale. Psychometric properties. *Journal of Consulting and Clinical Psychology*, 54, 381–385. <http://dx.doi.org/10.1037/0022-006X.54.3.381>
- Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijenburg, M. (2015). Measurement invariance. *Frontiers in Psychology*, 6, 1064. <http://doi.org/10.3389/fpsyg.2015.01064>
- Westbrook, D., Sedgwick-Taylor, A., Bennett-Levy, J., Butler, G., & McManus, F. (2008). A pilot evaluation of a brief CBT training course: impact on trainees' satisfaction, clinical skills and patient outcomes. *Behavioural and Cognitive Psychotherapy*, 36, 569–579. <https://doi.org/10.1017/S1352465808004608>
- Wilkinson, P., Alder, N., Juszczak, E., Matthews, H., Merritt, C., Montgomery, H., ... & Jacoby, R. (2009). A pilot randomised controlled trial of a brief cognitive behavioural group intervention to reduce recurrence rates in late life depression. *International Journal of Geriatric Psychiatry*, 24, 68–75. <http://doi.org/10.1002/gps.2076>
- Young, J. E., & Beck, A. T. (1980). *Cognitive Therapy Scale: Rating Manual*. Unpublished manuscript.

---

**Cite this article:** Beale S, Vitoratou S, and Liness S (2021). An investigation into the factor structure of the Cognitive Therapy Scale – Revised (CTS-R) in a CBT training sample. *Behavioural and Cognitive Psychotherapy* 49, 721–731. <https://doi.org/10.1017/S1352465820000983>