

MAIN

Multi-professional IAPT CBT training: clinical competence and patient outcomes

Sheena Liness^{1,*}, Sarah Beale¹, Susan Lea², Suzanne Byrne¹, Colette R. Hirsch^{1,§} and David M. Clark^{3,§}

¹Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK, ²Vice Chancellor, University of Hull, Hull, UK and ³Department of Experimental Psychology, University of Oxford, Oxford, UK

*Corresponding author. Email: sheena.liness@kcl.ac.uk

(Received 17 January 2018; revised 20 December 2018; accepted 15 January 2019; first published online 28 March 2019)

Abstract

Background: There is international interest in the training of psychological therapists to deliver evidence-based treatment for common mental health problems. The UK Improving Access to Psychological Therapies (IAPT) programme, one of the largest training initiatives, relies on competent therapists to successfully deliver cognitive behaviour therapy (CBT) and promote good patient outcome.

Aims: To evaluate an IAPT CBT training course by assessing if trainees' clinical skills improve during training and reach competency standards, and to report patient outcome for submitted training cases. To investigate a possible relationship between trainee competence and patient outcome. To explore professional differences during training.

Method: CBT trainee ($n = 252$) competence was assessed via audio recordings of therapy sessions at the beginning, middle and end of training. Patient pre- to post-treatment outcomes were extracted from submitted training cases ($n = 1927$). Differences in professional background were examined across competence, academic final grade and tutorial support.

Results: CBT trainees attained competence by the end of the course with 77% (anxiety recordings) and 72% (depression recordings) improving reliably. Training cases reported pre- to post-treatment effect sizes of 1.08–2.26 across disorders. CBT competence predicted a small variance in clinical outcome for depression cases. Differences in professional background emerged, with clinical psychologists demonstrating greater competence and higher academic grades. Trainees without a core professional background required more additional support to achieve competence.

Conclusions: Part of a new CBT therapist workforce was successfully trained to deliver relatively brief treatment effectively. Trainees without a core profession can be successfully trained to competence, but may need additional support. This has implications for workforce training.

Keywords: anxiety; depression; CBT training; clinical outcome; competence; IAPT

Introduction

Lack of access to appropriate treatment for common mental health problems has led to the development of large scale training initiatives worldwide aimed at disseminating evidence-based practice (Clark, 2018; Lund *et al.*, 2016; Rosen *et al.*, 2017). The success of these initiatives depends in part on the effective training of practitioners to deliver appropriate treatments competently and attain good patient outcomes. Success also depends on an increase in training numbers partly achieved through the expansion of training to less qualified practitioners (Fairburn and Patel, 2014; Singla *et al.*, 2017).

[§]Joint last authors due to equal contributions to the manuscript.

Improving Access to Psychological Therapies

The Improving Access to Psychological Therapies (IAPT) initiative, one of the largest training initiatives worldwide, set out to increase the availability of evidence-based National Institute for Health and Care Excellence (NICE) guidance treatment for depression and anxiety disorders (NICE, 2011) in England. By 2016, original targets were achieved with over 6500 new psychological therapists trained, over 50% of clinical cases recovered, and 15% of adults with depression and anxiety disorders across England accessing services (Department of Health, 2016a, 2016b, 2017). Given this success, new targets to train 10,500 psychological therapists to meet 25% of depression and anxiety disorder prevalence in England by 2020 were set (NHS England, 2016).

IAPT training

IAPT training is delivered as a joint venture between clinical services and universities. Trainees are employed in an IAPT service on a 1-year training contract, attending university 2 days a week, and delivering cognitive behaviour therapy (CBT) to patients they are trained to treat in IAPT services 3 days a week. A choice of evidence-based psychological therapies is delivered within a stepped-care treatment model, which aims to match patient severity to the appropriate intensity of intervention. Individual CBT is offered as a high-intensity intervention for moderate to severe depression and anxiety disorders. High-intensity CBT trainee therapists are the focus of this study.

To meet the requirements of an in-service training model and scale-up training at the level demanded by IAPT, entry criteria were widened to applicants without a core mental health professional background, namely Psychological Wellbeing Practitioners (PWP; Baguley *et al.*, 2010). Most PWPs are recent psychology graduates trained to deliver self-help, low-intensity IAPT interventions. Many PWPs have now been trained as IAPT high-intensity CBT therapists.

IAPT CBT training follows a national curriculum which specifies that courses cover two therapeutic models of depression – cognitive therapy and behavioural activation – and specific anxiety disorder protocols (Department of Health, 2011). Trainees submit eight training cases for formal assessment and receive weekly group supervision on training cases which includes live review of therapy clips and whole sessions. On course completion, trainees evidence 200 hours of CBT therapy plus 70 hours of CBT clinical supervision. An evaluation of the effectiveness of an IAPT CBT training course is timely and important given the international interest in workforce training and dissemination of evidence-based practice in real-world settings, and the large number of therapists trained and to be trained.

CBT training research

Previous CBT training evaluations indicate that trainees' clinical competence and knowledge improve over the training period (Barnfield *et al.*, 2007; Keen and Freeston 2008; McManus *et al.*, 2010), although trainee patient clinical outcomes were not reported. CBT trainees' patient outcomes evaluated in naturalistic clinical settings have reported comparable outcomes to those of qualified therapists (Forand *et al.*, 2011; Mason *et al.*, 2016; Öst *et al.*, 2012). Brief CBT training with the inclusion of a quasi-control condition of pre-training treatment as usual, had positive effects on clinical skills and patient outcome (Grey *et al.*, 2008; Westbrook *et al.*, 2008), with gains maintained at 12 months post-training (Simons *et al.*, 2010). These studies were, however, limited by small sample sizes and varied assessment methods.

Larger scale dissemination of CBT evidence-based training has shown promising results. In the USA, evidence-based treatment for veterans delivered by the US Department of Veteran Affairs has reported over 10,000 clinicians trained, improvement in therapist clinical competence, and medium to large effect sizes on patient outcome across diverse presentations and settings (Eftekhari *et al.*, 2013; Karlin and Cross, 2014; Karlin *et al.*, 2012). The expansion of psychological treatments in low- and middle-income countries has focused on specific treatments incorporating

CBT techniques adapted across cultures. These interventions report effect sizes for common mental disorders similar to those of high-income countries (Patel *et al.*, 2017; Singla *et al.*, 2017).

Studies on the relationship between therapist competence and clinical outcome have focused on experienced therapists within clinical trials and have produced mixed findings. No relationship was found in a meta-analysis of different therapeutic modalities (Webb *et al.*, 2010), whilst a modest significant overall relationship between competence and clinical outcome was demonstrated in a meta-analysis of CBT studies (Zarafonitis-Müller *et al.*, 2014). Larger effects were found for depression across both meta-analyses. Few studies have examined the relationship between therapy competence and clinical outcome in a trainee population. Evaluations of therapist training in cognitive therapy for psychosis (Jolley *et al.*, 2015) and in a sample of IAPT trainees (Branson *et al.*, 2015) found no relationship between therapist competence and patient outcome. Branson *et al.* (2015) reported a trend for reliable improvement in anxiety symptoms for patients of the most competent therapists, and reliable deterioration for patients of the least competent therapists, although these numbers were very small ($n = 4$). Larger student numbers in the present study enabled us to investigate this relationship further. The influence of professional background on CBT workforce training also merits further exploration.

On a CBT training course, preliminary evidence indicated that clinical psychologists performed better than other professional groups on academic assignments and therapeutic skills (McManus *et al.*, 2010). Compared with graduate mental health workers (a similar role to PWP), 'other' professions on a brief CBT training course demonstrated greater gains in clinical skills; however, this finding was not statistically significant, and the sample size was small (Westbrook *et al.*, 2008). National training programmes have reported the successful training of people from a wide range of mental health and non-professional backgrounds (Clark, 2018; Rosen *et al.*, 2017), and the scaling up of training of lay or community health workers has been fundamental for improving access to psychological treatments in low- and middle-income countries (Patel *et al.*, 2017; Van Ginneken *et al.*, 2013). IAPT trainees' diverse professional backgrounds and the inclusion of the non-core profession PWP enabled us to investigate this further in the current study.

The current study

This study aimed to evaluate the effectiveness of an IAPT CBT training course by assessing competence in CBT for depression and anxiety disorders at the beginning, middle and end of training, and reviewing trainees' patient outcomes on training cases submitted. It also set out to evaluate the relationship between trainee competence and patient outcome, and to investigate the impact of different professional backgrounds in training. We hoped to add to an emerging training research literature by including baseline competence, reporting generic and disorder-specific patient outcome measures, and focusing on a large and professionally diverse sample. It was hypothesized that (1) IAPT CBT trainees' competence would improve during training and meet the agreed standard of competence (2) IAPT trainees' patients would improve on clinical outcome measures (3) trainee therapists' competence would predict patient clinical outcome, with a larger effect in depression, and (4) differences in professional backgrounds would emerge.

Method

This is an observational naturalistic study with a longitudinal design. Consequently, inclusion of a no-training control group was not feasible as the research was embedded in the training programme.

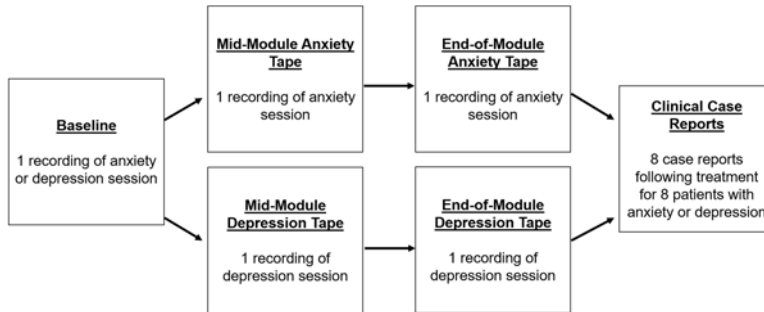


Figure 1. Assignment submissions across the course.

Participants

Participants were 252 CBT trainees from seven course years who trained between 2008 and 2015; all trainees who started the course were included in the current study. The sample consisted of 79.76% ($n = 201$) females and 20.24% ($n = 51$) males, while 82.94% ($n = 209$) were white and 17.06% ($n = 43$) were black and minority ethnic. Average age was 34.23 years ($SD = 6.78$ years). Final awards were Distinction (.79%, $n = 2$), Merit (30.16%, $n = 76$), Pass (63.89%, $n = 161$), and Fail/Withdrawn (5.16%, $n = 13$). Trainees who failed/withdrew were not demographically different from those who completed the course. Overall, 66.67% ($n = 168$) of trainees had a core profession¹: clinical psychologists (33.73%, $n = 85$), counselling psychologists (14.68%, $n = 37$), mental health nurses (10.32%, $n = 26$), occupational therapists (3.57%, $n = 9$), accredited counsellors (2.78%, $n = 7$), forensic psychologists (.79%, $n = 2$), psychiatrists (.40%, $n = 1$), and social workers (.40%, $n = 1$). The remaining 33.33% ($n = 84$) of trainees had no core profession, being PWP's (27.38%, $n = 69$) and non-accredited counsellors/psychotherapists (5.95%, $n = 15$).

Procedure/course submissions

Therapist competence and patient outcome data were extracted from coursework submissions. Trainees submitted tapes of mid-therapy sessions for each of the anxiety and depression modules for assessment: an informal 'mid-module' tape to provide skills-focused feedback, and a formally examined 'end-of-module' tape. From Year 2 of IAPT training (2009) the course introduced a baseline tape of either an anxiety or depression session to gauge clinical skills and frontload support in the first month of training; this submission became mandatory in Year 3. Efforts to blind markers to trainee identity included removal of identifying information and allocation to markers with limited familiarity with individual trainees. Markers were not blind to stage of training. Failed assignments were allowed one resubmission following additional support.² Fig. 1 details the course submission process.

Each trainee treated and submitted reports for eight supervised cases, including self-report clinical outcome measures completed by patients. Inclusion criteria for course cases were treatment as an adult outpatient in a local IAPT service with primary diagnosis of depression or an anxiety disorder; exclusion criteria were primary personality disorder, psychosis, substance dependence, and/or severe psychosocial impairment. Case information was available for 1927 patients; the remaining were missing due to withdrawal/failure ($n = 73$) and human error ($n = 13$).

¹Core and non-core profession status was categorized in accordance with British Association for Behavioural and Cognitive Psychotherapies (BABCP) guidelines. Please refer to the following for further information: <http://www.babcp.com/Accreditation/Core-Profession.aspx>

²Trainees who failed a therapy tape were offered additional support constituting three clinical skills tutorials and informal feedback on an additional tape.

Table 1. Disorder-specific clinical outcome measures

Disorder	Measure	Items <i>n</i> (range)	Caseness threshold	Reliable change index
GAD	PSWQ (Meyer <i>et al.</i> , 1990)	16 (0–80)	≥45	≥7
PTSD	IES-R (Weiss and Marmar, 1996)	22 (0–88)	≥33	≥9
OCD	OCI (Foa <i>et al.</i> , 1998)	42 (0–168)	≥40	≥32
Health anxiety	s-HAI (Salkovskis <i>et al.</i> , 2002)	14 (0–54)	≥18	≥4
Social anxiety	SPIN (Connor <i>et al.</i> , 2000)	17 (0–68)	≥19	≥10
Panic disorder	PDSS (Shear <i>et al.</i> , 2001)	7 (0–28)	≥8	n/a

GAD, generalized anxiety disorder; PTSD, post-traumatic stress disorder; OCD, obsessive compulsive disorder; PSWQ, Penn State Worry Questionnaire; IES-R, Impact of Events Scale – Revised; OCI, Obsessive Compulsive Inventory; s-HAI, Short Health Anxiety Inventory; SPIN, Social Phobia Inventory; PDSS, Panic Disorder Severity Scale.

Primary diagnoses were: depression ($n = 520$), social phobia ($n = 320$), obsessive compulsive disorder ($n = 287$), panic disorder ($n = 257$), generalized anxiety disorder ($n = 148$), specific phobia ($n = 127$), post-traumatic stress disorder ($n = 130$), health anxiety disorder ($n = 126$), body dysmorphic disorder ($n = 11$), and anxiety disorder not otherwise specified ($n = 1$).

Measures

Trainee competence

Trainee competence for anxiety and depression tapes was rated with the UK 12-item Cognitive Therapy Scale – Revised (CTS-R; Blackburn *et al.*, 2001; range 0–72; pass mark/competence threshold ≥ 36) which is widely used in IAPT training. Items 1–5 assess general interpersonal/therapeutic competence (Generic Subscale) and items 6–12 assess CBT-specific skills (Specific Subscale). The CTS-R demonstrates good internal consistency and acceptable inter-rater reliability on the total scale (Blackburn *et al.*, 2001; James *et al.*, 2001; Kazantzis *et al.*, 2018) and on generic and CBT specific subscales (Kazantzis *et al.*, 2018). Expert raters trained in scoring together demonstrate higher reliability (Gordon, 2006; Reichelt *et al.*, 2003). All raters in this study were CBT accredited practitioners and had worked and trained together long-term. Internal inter-rater reliability between course staff was high based on a random sample of 9% ($n = 114$) of the 1225 submitted tapes: one-way random single measures ICC (114, 113) = .95. For this study, 4% ($n = 43$) of tapes were marked by an external expert blind to trainee identity and time. Inter-rater reliability between internal and external markers was good: two-way random single measures ICC (42, 41) = .68. This ICC is comparable to similar studies using the CTS-R (Branson *et al.*, 2015; Gordon, 2006).

Clinical outcomes

Two validated self-report measures extracted from trainee case reports assessed anxiety and depression symptoms for all patients. The 7-item Generalized Anxiety Disorder-7 (GAD-7; Spitzer *et al.*, 2006) assessed anxiety symptoms (range 0–21, caseness threshold ≥ 8 , reliable change index ≥ 4). The 9-item Patient Health Questionnaire (PHQ-9; Kroenke *et al.*, 2001) assessed depression symptoms for all cases (range 0–27, caseness threshold ≥ 10 , reliable change index ≥ 6). The PHQ-9 also assessed disorder-specific clinical outcomes for depression cases. Table 1 details the IAPT-recommended (Department of Health, 2014) self-report measures used by trainees to assess anxiety disorder-specific clinical outcomes.

Results

The Benjamini–Hochberg Procedure (Benjamini and Hochberg, 1995) was applied to all hypothesis tests to correct for multiplicity, with false discovery rate $Q = .05$.³

³All analysis was conducted in SPSS version 25, except where stated.

Table 2. Summary of trainees' CTS-R outcomes by module

CTS-R by tape	<i>n</i>	Mean	<i>SD</i>	<36	≥36
Baseline ^a	169	29.09	5.32	150 (88.76%)	19 (11.24%)
Mid-module anxiety ^b	245	35.38	6.47	131 (53.47%)	114 (46.53%)
End-of-module anxiety	246	38.19	5.10	45 (18.29%)	201 (81.70%)
Resubmitted anxiety	43	39.76	4.45	3 (6.98%)	40 (93.02%)
Mid-module depression	232	34.66	5.59	127 (54.74%)	105 (45.26%)
End-of-module depression	246	37.88	4.76	50 (20.33%)	196 (79.67%)
Resubmitted depression	44	39.31	2.79	(0.00%)	44 (100.00%)

^aBaseline tapes were only introduced to the course as of Year 2, therefore there are fewer total baseline tapes than other course tapes. Baseline tapes consisted of 103 anxiety tapes, 64 depression tapes, and two tapes of other presentations.

^bCases missing either the anxiety and/or depression tape were excluded from relevant mean calculations. CTS-R, Cognitive Therapy Scale – Revised.

Missing data

Only 169 out of 252 trainees (67%) submitted baseline tapes. Missing tapes were mainly from Years 1 and 2 when submission of a baseline case was not a course requirement. Eight of the 252 trainees withdrew, and five trainees failed the course. CTS-R scores were missing for seven mid-module anxiety tapes, twenty mid-module depression tapes, six end-of-module anxiety and six end-of-module depression tapes due to fails/withdrawals and human error. Due to the observational retrospective design, missing data were excluded and reported on a case-by-case basis. Consequently, reported *n* values for therapist variables vary. Seventy-three case reports were missing due to fails/withdrawals, while 16 were missing due to human error during submission.

Trainees' clinical competence

Table 2 gives mean CTS-R scores and frequency data for trainees' competence (CTS-R total ≥ 36) across the anxiety and depression modules. The majority of participants were below competence at baseline but attained competence by the end-of-module tapes. The majority of trainees who failed and required resubmissions also later attained competence.

Change in competence from baseline

Change in total CTS-R scores and in generic and specific subscale scores across training are reported in Fig. 2a for the anxiety module and Fig. 2b for the depression module. For the 100 trainees with a baseline, mid-module, and end-of-module anxiety tape, repeated measures analyses of variance (ANOVAs) indicated that time had a significant effect on therapist competence for the total measure and the generic and specific subscales, respectively: $F_{\text{total}}(2,198) = 93.78$, $p < .001$, partial $\eta^2 = .49$; $F_{\text{generic}}(2,198) = 49.03$, $p < .001$, partial $\eta^2 = .33$; and $F_{\text{specific}}(2,198) = 102.04$, $p < .001$, partial $\eta^2 = .72$. All *post-hoc* tests demonstrated significantly increased competence between each tape, $p < .001$.

For the 59 trainees with a baseline, mid-module, and end-of-module depression tape, a repeated measures ANOVA found that time exhibited a significant effect on competence for the total measure, and generic and specific subscales, respectively: $F_{\text{total}}(2,116) = 57.32$, $p < .001$, partial $\eta^2 = .50$; $F_{\text{generic}}(2,116) = 32.66$, $p < .001$, partial $\eta^2 = .36$; and $F_{\text{specific}}(2,116) = 64.62$, $p < .001$, partial $\eta^2 = .76$. As above, all *post-hoc* tests demonstrated significantly increased competence between each tape, $p < .001$.

Resubmissions

For both modules, trainees who failed their end-of-module tape significantly improved in competence on the resubmitted tape. In the anxiety module, competence increased from 30.98

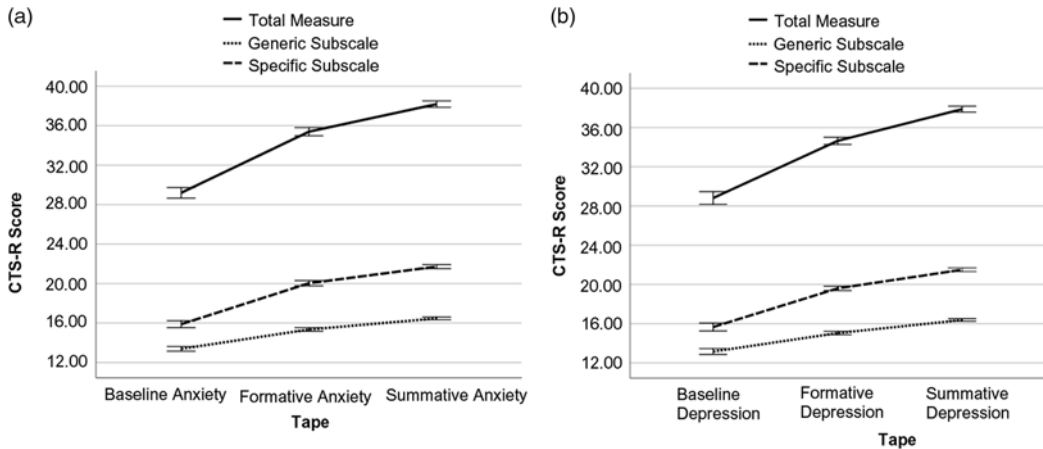


Figure 2. Change in mean competence scores and standard errors across anxiety and depression modules.

($SD = 3.44$) to 39.76 ($SD = 4.45$) on the CTS-R: $t(42) = -10.92, p < .001$. Gains in competence were significant for both the generic [$t(42) = -8.63, p < .001$] and specific, [$t(42) = -12.84, p < .001$] CTS-R subscales.

In the depression module, competence improved on average from 30.72 ($SD = 3.22$) to 39.31 ($SD = 2.79$): $t(43) = -12.44, p < .001$. Gains in competence were significant for both the generic [$t(43) = -10.67, p < .001$] and specific [$t(43) = -11.72, p < .001$] CTS-R subscales.

Reliability of change in competence

Based on Branson *et al.*'s (2015) reliable change index for competence of ± 4.5 points on the CTS-R, 77.00% ($n = 77$) of trainees with a baseline and end-of-module anxiety tape ($n = 100$) showed reliable improvement, 23.00% ($n = 23$) showed no reliable change, and no trainees showed reliable deterioration between the baseline and end-of-module tapes. Of trainees with a baseline and end-of-module depression tape ($n = 61$), 72.13% ($n = 44$) showed reliable improvement, 26.23% ($n = 16$) showed no reliable change, and 1.64% ($n = 1$) of trainees showed reliable deterioration between the baseline and end-of-module tapes.

In sum, trainees demonstrated significant increases in competence both between tapes and across the course overall, with the majority of trainees demonstrating reliable improvement.

Trainees' patient outcomes

Table 3 gives t -tests of mean pre- to post-difference scores on clinical outcome measures, and frequencies of recovery and reliable improvement.⁴ Available disorder-specific measure outcomes are reported. Recovery in IAPT refers to pre-treatment clinical cases on either the PHQ-9 and/or the GAD-7 that were below caseness for both measures post-treatment (Department of Health, 2014). Reliable improvement in IAPT refers to the reduction by the clinically reliable change index on at least one generic measure and no reliable deterioration on the other (Department of Health, 2014; Jacobson and Truax, 1991). Patients with depression received a mean of 13 sessions ($SD = 4$) and those with anxiety a mean of 11 sessions ($SD = 3$). Data for the Panic Disorder

⁴As the distribution of diagnoses treated by each trainee differed, one-sample t -tests of pre- to post-treatment difference scores were performed in Stata version 15 using an intercept-only regression beta test to adjust standard errors for clustering by trainee ID (Rogers, 1994).

Table 3. Pre- and post-treatment clinical outcomes for trainees' patients

	Clinical cases <i>n</i>	Total cases <i>n</i>	Mean _{pre} (<i>SD</i> _{pre})	Mean _{post} (<i>SD</i> _{post})	d.f.	<i>t</i>	<i>d</i> ^a	Recovery <i>n</i> (%)	Reliable improvement <i>n</i> (%)
Depression overall									
IAPT criteria ^b	501	520	—	—	—	—	—	336 (67.07)	455 (87.67)
PHQ-9	488	520	17.15 (4.92)	6.60 (5.53)	245	37.03**	1.80	364 (74.59)	423 (81.35)
Anxiety overall									
IAPT criteria ^b	1134	1406	—	—	—	—	—	874 (77.07)	1116 (79.37)
GAD-7	1089	1406	12.19 (5.43)	4.46 (4.42)	245	45.74**	1.38	858 (78.79)	1067 (75.89)
GAD									
PSWQ	138	140	67.99 (8.25)	43.11 (13.73)	105	18.57**	1.95	75 (54.35)	131 (93.57)
GAD-7	129	148	13.76 (4.91)	4.78 (3.96)	110	18.45**	1.95	107 (82.94)	121 (81.76)
PTSD									
IES-R	101	118	52.36 (17.89)	13.23 (17.20)	96	20.46**	2.26	88 (87.13)	111 (94.07)
GAD-7	106	130	12.59 (5.14)	4.22 (4.92)	103	14.40**	1.60	87 (82.07)	104 (80.00)
OCD									
OCI	191	257	61.77 (28.94)	24.60 (24.16)	168	23.97**	1.46	140 (73.30)	136 (52.92)
GAD-7	244	287	13.28 (4.96)	5.41 (4.80)	183	22.98**	1.58	177 (72.54)	223 (77.70)
Health anxiety									
s-HAI	115	118	31.45 (8.51)	15.26 (9.10)	99	11.69**	1.08	88 (76.52)	109 (92.37)
GAD-7	95	126	12.11 (5.37)	3.51 (3.45)	104	16.64**	1.65	80 (84.21)	104 (82.54)
Social anxiety									
SPIN	296	303	40.77 (11.79)	16.99 (12.40)	186	28.74**	1.87	186 (62.84)	254 (83.83)
GAD-7 ^c	233	319	11.21 (5.06)	3.99 (3.76)	194	26.33**	1.62	193 (82.83)	245 (76.80)
Panic disorder									
PDSS	67	74	15.09 (5.48)	4.38 (4.60)	58	13.53**	1.83	51 (76.12)	—
GAD-7	213	257	13.13 (5.30)	5.01 (5.01)	170	20.88**	1.48	163 (76.53)	196 (76.26)

GAD, generalized anxiety disorder; PTSD, post-traumatic stress disorder; OCD, obsessive compulsive disorder; PHQ-9, Patient Health Questionnaire-9; GAD-7, Generalized Anxiety Disorder-7; PSWQ, Penn State Worry Questionnaire; IES-R, Impact of Events Scale – Revised; OCI, Obsessive Compulsive Inventory; S-HAI, Short Health Anxiety Inventory; SPIN, Social Phobia Inventory; PDSS, Panic Disorder Severity Scale. ***p* < .001.

^aStandard deviations based on trainees' scores and robust standard errors corrected for clustering (Rogers, 1994); Cohen's *d* corrected for mean dependence (Morris and DeShon, 2002).

^bIAPT criteria recovery: clinical cases on the PHQ-9 and/or GAD-7 that reduce below caseness on both measures post-treatment (Department of Health, 2014). IAPT reliable improvement: reduction by clinically reliable change index on at least one generic measure and no reliable deterioration (Department of Health, 2014; Jacobson and Truax, 1991).

^cCase report for one patient included only the SPIN and was missing the GAD-7; therefore, total number of social phobia cases overall was *n* = 320.

Severity Scale were limited to 77 cases, as this measure was only implemented in 2012–2013. Mean scores on all measures significantly improved for all disorders. Based on IAPT criteria, the majority of depression and anxiety patients recovered (67.07 and 77.01%, respectively) and improved reliably (87.67 and 79.40%, respectively). Large pre- to post-effect sizes of 1.08–2.26 were evident across disorders.

Relationship between trainee therapy competence and clinical outcome

Simple linear regressions assessed whether end-of-module CTS-R scores on the depression or anxiety tape predicted percentage change in the PHQ-9 scores for depression cases and the GAD-7 scores for anxiety cases.⁵

The depression tape significantly predicted percentage change in PHQ-9 scores for depression cases, $F(1,243) = 4.57$, $p = .03$, and explained 1.3% of the variance in case outcomes ($R^2 = .013$). A 1-point increase on the CTS-R was associated with a .76% increase in PHQ-9 percentage improvement ($\beta = .76$, $SE = .35$). These findings indicate that CTS-R scores predicted a small amount of improvement for depression cases. The anxiety tape did not significantly predict percentage change in GAD-7 scores for anxiety cases, $F(1,244) = 1.79$, $p = .18$.

Professional background

Given the variety of professions trained in IAPT, an investigation of the different professional backgrounds focused on three groups: clinical psychologists ($n = 85$), other trainees with a core profession ($n = 83$), and trainees with no core profession ($n = 84$).

Academic final grade

A one-way ANOVA found a significant difference in mean final grade for academic assignments between professional groups: $F(2,240) = 26.95$, $p < .001$, partial $\eta^2 = .18$. Clinical psychologists had a higher mean final grade (mean = 62.68, $SD = 6.08$) than trainees with another core profession (mean = 57.14, $SD = 5.35$), and trainees with no core profession (mean = 57.81, $SD = 5.18$), both $p < .001$. Mean final grade did not significantly differ between trainees with another core profession and trainees with no core profession, $p = .99$.

Additional support

A Kruskal–Wallis test found that additional support requirements differed significantly between professional groups: $H = 39.48$, $p < .001$. Clinical psychologists required less support than other professional groups, both $p < .001$. Trainees with another core profession required less support than those with no core profession, $p = .01$.

CTS-R competence ratings

A mixed ANOVA found significant main effects of profession and time on CTS-R scores: $F_{\text{profession}}(2,160) = 7.46$, $p = .001$, partial $\eta^2 = .08$ and $F_{\text{time}}(1,160) = 396.41$, $p < .001$, partial $\eta^2 = .71$. Clinical psychologists demonstrated higher CTS-R scores than other professional groups at baseline and end-of-training, all $p < .04$. Trainees with another core profession and those with no core profession did not significantly differ, all $p > .05$. All professional groups were below competence at baseline (mean_{ClinPsy} = 31.26, $SD_{\text{ClinPsy}} = 4.89$; mean_{OtherCore} = 29.18, $SD_{\text{OtherCore}} = 4.67$; mean_{NoCore} = 27.79, $SD_{\text{NoCore}} = 5.43$), but significantly increased and attained

⁵Regression was performed in Stata version 15 to adjust standard errors for clustering by therapist ID.

competence by end-of-training (mean_{ClinPsy} = 38.80, *SD*_{ClinPsy} = 3.44; mean_{OtherCore} = 36.62, *SD*_{OtherCore} = 4.50; mean_{NoCore} = 37.40, *SD*_{NoCore} = 3.67), all $p < .001$.

Main effects were qualified by a significant interaction: $F(2,160) = 3.17, p = .05, \eta^2 = .04$. A *post-hoc* one-way ANOVA found significant differences in CTS-R change scores: $F(2,160) = 3.17, p = .04, \text{partial } \eta^2 = .04$. Trainees with no core profession demonstrated greater change than both other professional groups, all $p < .04$; clinical psychologists and other core professionals did not significantly differ, $p = .91$.

Overall, clinical psychologists demonstrated stronger academic performance and higher CTS-R scores than other professional groups. All professional groups attained competence. Trainees without a core profession evidenced a steeper increase in competence attainment but required more additional support.

Discussion

This study focused on the training of 252 IAPT CBT therapists and their 1927 submitted training cases, and predicted that trainees' clinical competence would improve during training and that they would evidence good patient outcome. It also set out to explore the relationship between therapist competence and clinical outcome, and to investigate differences in clinical and academic performance across professional and non-professional backgrounds.

CBT competence

Trainee CBT clinical competence, assessed on the CTS-R (Blackburn *et al.*, 2001), improved significantly from baseline to end-of-course, passing the competence threshold and demonstrating reliable improvement for treating people with depression and with anxiety disorders. These results support evidence that CBT training improves competence (Branson *et al.*, 2015; Karlin *et al.*, 2012; McManus *et al.*, 2010). Assessment of the CTS-R subscales showed improvement across both generic and CBT-specific skills in training (Simons *et al.*, 2010). Notably, baseline competence was considerably lower in the present study (29.09) than in previous comparable CBT training studies using the CTS-R (35.72–38.94; see Barnfield *et al.*, 2007; Branson *et al.*, 2015; Keen and Freeston, 2008). This is possibly accounted for by the inclusion of trainees without a core profession. It is encouraging for national workforce training initiatives that all professional groups achieved competence by the end of training.

Clinical outcome for submitted cases

Patient clinical outcomes on submitted cases were positive, with large effect sizes across all measures. Patients' recovery compared favourably to London and national IAPT services, where the government target is for 50% of eligible referrals to recover (Department of Health, 2017). Trainees achieved overall recovery rates of 77% for anxiety cases and 67% for depression cases. Overall reliable improvement rates were also encouraging at 79% of trainee anxiety cases and 88% of trainee depression cases. These positive patient outcomes were consistent with previous CBT training studies (Jolley *et al.*, 2015; Öst *et al.*, 2012), possibly due to factors found to influence therapy fidelity and good outcome: the training is extensive (Rakovshik and McManus, 2010) and includes regular supervision and close monitoring of training cases via audio recordings with comprehensive individual feedback for skills development (Alfonsson *et al.*, 2017; Weck *et al.*, 2017). Trainees also offered patients the NICE-recommended number of sessions, providing an average of 12 therapy sessions per case, compared with an IAPT national average of seven (Department of Health, 2016a), and were encouraged to adhere to NICE guidance protocols (NICE, 2011). Both number of treatment sessions (Clark *et al.*, 2017) and adherence to protocols (Gyani *et al.*, 2013) are associated with positive clinical outcomes.

Relationship of therapy competence and clinical outcome

A significant but small relationship between trainee competence and clinical outcomes for depression but not anxiety cases was found. These findings demonstrate some affinity with the lack of or small relationship found in the CBT training literature (Branson *et al.*, 2015; Jolley *et al.*, 2015) and may reflect a genuinely small relationship between therapist competence and patient outcome. It is also possible, however, that the restriction of range in training cases and outcome measures which is determined by training course requirements for straightforward training cases could be causing an under-estimation of the true relationship. Other important influences on the therapy outcome of these cases such as patient suitability, co-morbidity, duration of problem, life circumstances and social deprivation (Clark *et al.*, 2017; Fairburn and Cooper, 2011; Webb *et al.*, 2010), were unable to be accounted for as patient characteristics were not consistently presented in case studies. The lack of significance for anxiety cases may also be due to use of the generic CTS-R (Blackburn *et al.*, 2001) assessment scale, a scale adapted from the original Cognitive Therapy Scale (Young and Beck, 1980). Although these are the two most widely used competence assessment scales in CBT training (Rakovshik and McManus, 2010), a recent examination of both scales (Kazantzis *et al.*, 2018) found that whilst both the CTS and CTS-R demonstrated acceptable internal consistency and inter-rater reliability and predicted reductions in depressive symptoms at early treatment, neither scale demonstrated significant competence-outcome relationships for late treatment sessions. Further research with a focus on specific therapist competencies could incorporate assessing anxiety cases with anxiety disorder-specific competence assessment tools where stronger associations have been reported (Ginzburg *et al.*, 2012).

Professional background

Differences emerged across professions. Clinical psychologists had significantly higher baseline and end-of-course CBT competence, and higher final grades than other trainees. Similar professional disparities between clinical psychologists and other professions have been reported previously (McManus *et al.*, 2010), possibly indicating that clinical psychologists require less intensive training to attain CBT competence post-core training. Other core professions had varied mental health backgrounds including mental health nurses and occupational therapists. While core training for these roles is robust, it is diverse and less CBT-focused. It therefore follows that these trainees attain competence requiring some additional support. Trainees without a core profession also attained competence but required a greater amount of additional support. With the scaling-up of evidence-based therapy worldwide, and subsequent increase in training of non-professionals on specific interventions (Patel *et al.*, 2017; Rosen *et al.*, 2017), it is reassuring to find that non-core professional practitioners can be trained to a recognized level of therapy competence on a CBT training course which covers evidence-based interventions across depression and a range of anxiety disorders. Identifying training needs across professions early in training, in part via a baseline therapy tape, may help trainees achieve competence by tailoring supervision and additional support to those who need it.

Limitations

This study is an evaluation of a specific training course, which limits the generalizability of findings. In the absence of randomization to training or no training, it is impossible to determine how much of the observed improvements are a specific consequence of the training as opposed to trainees' further clinical experience during the training period. However, the large and professionally diverse sample, inclusion of baseline therapy assessment and clinical outcomes addressed limitations of previous studies. Internal markers were blind to trainee identity but not stage of training, and trainees self-selecting therapy sessions for assessment and the use of self-report patient outcome measures are all standard practices on training courses, but the representativeness of these therapy

sessions and patient cases to general therapy practice is uncertain. Self-selection of cases may create a bias that inflates estimates of competence; however, cases were selected in supervision prior to knowledge of outcome (i.e. at the start of treatment) and all eight cases were analysed. Self-selection could also be seen to support learning by fostering engagement in considerable self-reflection (Branson *et al.*, 2015). Alternatives are assessment of more sessions (Keen and Freeston, 2008) or random selection of sessions (Ginzburg *et al.*, 2012). Both alternatives demand increased training course resources.

Despite these limitations, this study adds to the limited research on CBT workforce training with naturalistic real-world training data. On course completion, the majority of IAPT trainees are competent CBT practitioners whose training cases improve significantly. A small relationship between trainee therapy competence and clinical outcome for depression cases is reported. Participating in a CBT training course that focuses on evidence-based interventions, provides regular skills feedback and close supervision, and has the resources to conduct baseline competence assessment and increase support for varied professional backgrounds produces competent CBT therapists and excellent clinical outcomes.

Acknowledgements. With many thanks to all the IAPT CBT trainees and course staff who contributed to this study, and to Hannah Parker and Steffen Nestler for their help with the data.

Conflicts of interest. Sheena Liness and Suzanne Byrne run the IAPT CBT training at the IoPPN/King's College London course that is the subject of this study and David M. Clark is NHS England's Clinical Advisor for the IAPT programme. Sarah Beale, Susan Lea and Colette Hirsch have no conflicts of interest with respect to this publication.

Ethics statement. All contributors abided by the Ethical Principles of Psychologists and Code of Conduct as set out by the APA. This study was approved by the King's College London Psychiatry, Nursing, and Midwifery Research Ethics Committee: reference number PNM/12/13-50.

Financial support. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Colette Hirsch receives salary support from the National Institute for Health Research (NIHR), Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed in this article are those of the author(s) and not necessarily those of King's College London, the NIHR, or the Department of Health.

References

- Alfonsson, S., Parling, T., Spännargård, Å., Andersson, G. and Lundgren, T. (2017). The effects of clinical supervision on supervisees and patients in cognitive behavioral therapy: a systematic review. *Cognitive Behaviour Therapy*, 47, 206–228.
- Baguley, C., Farrand, P., Hope, R., Leibowitz, J., Lovell, K., Lucock, M., *et al.* (2010). Good practice guidance on the use of self-help materials within IAPT services. Technical Report: IAPT. Retrieved from: <http://eprints.hud.ac.uk/9017/>
- Barnfield, T. V., Mathieson, F. M. and Beaumont, G. R. (2007). Assessing the development of competence during post-graduate cognitive-behavioral therapy training. *Journal of Cognitive Psychotherapy*, 21, 140–147.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300.
- Blackburn, I. M., James, I. A., Milne, D. L., Baker, C., Standart, S., Garland, A., *et al.* (2001). The revised cognitive therapy scale (CTS-R): psychometric properties. *Behavioural and Cognitive Psychotherapy*, 29, 431–446.
- Branson, A., Shafraan, R. and Myles, P. (2015). Investigating the relationship between competence and patient outcome with CBT. *Behaviour Research and Therapy*, 68, 19–26.
- Clark, D. M. (2018). Realizing the mass public benefit of evidence-based psychological therapies: the IAPT program. *Annual Review of Clinical Psychology*, 14, 159–183.
- Clark, D.M., Canvin, L., Green, J., Layard, R., Pilling, S. and Janecka, M. (2017). Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. *The Lancet*, 391, 679–686.
- Connor, K. M., Davidson, J. R., Churchill, L. E., Sherwood, A., Weisler, R. H. and Foa, E. (2000). Psychometric properties of the Social Phobia Inventory (SPIN). *The British Journal of Psychiatry*, 176, 379–386.
- Department of Health (2011). National Curriculum for High Intensity Cognitive Behavioural Therapy Courses. UK: NHS Digital.
- Department of Health (2014). *Improving Access to Psychological Therapies: Measuring Improvement and Recovery – Adult Services. Version 2.* UK: NHS Digital.

- Department of Health** (2016a). *Psychological Therapies: Annual Report on the Use of IAPT Services, England, 2015–16*. UK: NHS Digital.
- Department of Health** (2016b). *2015 Adult IAPT Workforce Census Report*. UK: NHS England and Health Education England.
- Department of Health** (2017). *Improving Access to Psychological Therapies (IAPT) Executive Summary (March 2017)*. UK: NHS Digital.
- Eftekhari, A., Ruzek, J. I., Crowley, J. J., Rosen, C. S., Greenbaum, M. A. and Karlin, B. E.** (2013). Effectiveness of national implementation of prolonged exposure therapy in Veterans affairs care. *JAMA Psychiatry*, 70, 949–955.
- Fairburn, C. G. and Cooper, Z.** (2011). Therapist competence, therapy quality, and therapist training. *Behaviour Research and Therapy*, 49, 373–378.
- Fairburn, C. G. and Patel, V.** (2014). The global dissemination of psychological treatments: a road map for research and practice. *American Journal of Psychiatry*, 171, 495–498.
- Foa, E. B., Kozak, M. J., Salkovskis, P. M., Coles, M. E. and Amir, N.** (1998). The validation of a new obsessive-compulsive disorder scale: the Obsessive-Compulsive Inventory. *Psychological Assessment*, 10, 206.
- Forand, N. R., Evans, S., Haglin, D. and Fishman, B.** (2011). Cognitive behavioral therapy in practice: treatment delivered by trainees at an outpatient clinic is clinically effective. *Behavior Therapy*, 42, 612–623.
- Ginzburg, D. M., Bohn, C., Höfling, V., Weck, F., Clark, D. M. and Stangier, U.** (2012). Treatment specific competence predicts outcome in cognitive therapy for social anxiety disorder. *Behaviour Research and Therapy*, 50, 747–752.
- Gordon, P. K.** (2006). A comparison of two versions of the Cognitive Therapy Scale. *Behavioural and Cognitive Psychotherapy*, 35, 343–353.
- Grey, N., Salkovskis, P., Kuigley, A., Clark, D. M. and Ehlers, A.** (2008). Dissemination of cognitive therapy for panic disorder in primary care. *Behavioural and Cognitive Psychotherapy*, 36, 509–520.
- Gyani, A., Shafran, R., Layard, R. and Clark, D. M.** (2013). Enhancing recovery rates: lessons from year one of IAPT. *Behaviour Research and Therapy*, 51, 597–606.
- Jacobson, N. and Truax, P.** (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12.
- James, I. A., Blackburn, I. M., Milne, D. L. and Reichelt, F. K.** (2001). Moderators of trainee therapists' competence in cognitive therapy. *British Journal of Clinical Psychology*, 40, 131–141.
- Jolley, S., Onwumere, J., Bissoli, S., Bhayani, P., Singh, G., Kuipers, E., *et al.*** (2015). A pilot evaluation of therapist training in cognitive therapy for psychosis: therapy quality and clinical outcomes. *Behavioural and Cognitive Psychotherapy*, 43, 478–489.
- Karlin, B. E., Brown, G. K., Trockel, M., Cuning, D., Zeiss, A. M. and Taylor, C. B.** (2012). National dissemination of cognitive behavioral therapy for depression in the Department of Veterans Affairs Health Care System: therapist and patient-level outcomes. *Journal of Consulting and Clinical Psychology*, 80, 707.
- Karlin, B. E. and Cross, G.** (2014). From the laboratory to the therapy room: national dissemination and implementation of evidence-based psychotherapies in the US Department of Veterans Affairs Health Care System. *American Psychologist*, 69, 19.
- Kazantzis, N., Clayton, X., Cronin, T. J., Farchione, D., Limburg, K. and 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.
- Keen, A. J. and Freeston, M. H.** (2008). Assessing competence in cognitive-behavioural therapy. *British Journal of Psychiatry*, 193, 60–64.
- Kroenke, K., Spitzer, R. L. and Williams, J. B.** (2001). The PHQ-9. *Journal of General Internal Medicine*, 16, 606–613.
- Lund, C., Tomlinson, M. and Patel, V.** (2016). Integration of mental health into primary care in low-and middle-income countries: the PRIME mental healthcare plans. *British Journal of Psychiatry*, 208, s1–3.
- Mason, L., Grey, N. and Veale, D.** (2016). My therapist is a student? The impact of the therapist experience and client severity on cognitive behavioural therapy outcomes for people with anxiety disorders. *Behavioural and Cognitive Psychotherapy*, 44, 193–202.
- McManus, F., Westbrook, D., Vazquez-Montes, M., Fennell, M. and Kennerley, H.** (2010). An evaluation of the effectiveness of diploma-level training in cognitive behaviour therapy. *Behaviour Research and Therapy*, 48, 1123–1132.
- Meyer, T. J., Miller, M. L., Metzger, R. L. and Borkovec, T. D.** (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, 28, 487–495.
- Morris, S. B. and DeShon, R. P.** (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7, 105.
- National Institute of Health and Care Excellence (NICE)** (2011). *Common Mental Health Disorders: Identification and Pathways to Care*. Clinical Guideline CG123. Retrieved from: <https://www.nice.org.uk/guidance/cg123/evidence/full-guideline-181771741>
- NHS England** (2016). *Achieving Better Access to Mental Health Services by 2020*. London: NHS England. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/361648/mental-health-access.pdf

- Öst, L. G., Karlstedt, A. and Widén, S. (2012). The effects of cognitive behavior therapy delivered by students in a psychologist training program: an effectiveness study. *Behavior Therapy*, 43, 160–173.
- Patel, V., Weobong, B., Weiss, H. A., Anand, A., Bhat, B., Katti, B., *et al.* (2017). The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: a randomised controlled trial. *The Lancet*, 389, 176–185.
- Rakovshik, S. G. and McManus, F. (2010). Establishing evidence-based training in cognitive behavioral therapy: a review of current empirical findings and theoretical guidance. *Clinical Psychology Review*, 30, 496–516.
- Reichelt, F. K., James, I. A. and Blackburn, I. M. (2003). Impact of training on rating competence in cognitive therapy. *Journal of Behavior Therapy and Experimental Psychiatry*, 34, 87–99.
- Rogers, W. (1994). Regression standard errors in clustered samples. *Stata Technical Bulletin*, 3(13), 19–23.
- Rosen, R. C., Ruzek, J. I. and Karlin, B. E. (2017). Evidence-based training in the era of evidence-based practice: challenges and opportunities for training of PTSD providers. *Behaviour Research and Therapy*, 88, 37–48.
- Salkovskis, P. M., Rimes, K. A., Warwick, H. M. C. and Clark, D. M. (2002). The Health Anxiety Inventory: development and validation of scales for the measurement of health anxiety and hypochondriasis. *Psychological Medicine*, 32, 843–853.
- Shear, M. K., Rucci, P., Williams, J., Frank, E., Grochocinski, V., Vander Bilt, J., *et al.* (2001). Reliability and validity of the Panic Disorder Severity Scale: replication and extension. *Journal of Psychiatric Research*, 35, 293–296.
- Simons, A. D., Padesky, C. A., Montemaranano, J., Lewis, C. C., Murakami, J., Lamb, K., *et al.* (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.
- Singla, D. R., Kohrt, B. A., Murray, L. K., Anand, A., Chorpita, B. F. and Patel, V. (2017). Psychological treatments for the world: lessons from low- and middle-income countries. *Annual Review of Clinical Psychology*, 13, 149–181.
- Spitzer, R. L., Kroenke, K., Williams, J. B. and Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166, 1092–1097.
- Van Ginneken, N., Tharyan, P., Lewin, S., Rao, G. N., Meera, S. M., Pian, J., *et al.* (2013). Non-specialist health worker interventions for the care of mental, neurological and substance-abuse disorders in low-and middle-income countries. *Cochrane Database of Systemic Reviews*, 11, CD009149.
- Webb, C. A., DeRubeis, R. J. and Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: a meta-analytic review. *Journal of Consulting and Clinical Psychology*, 78, 200–211.
- Weck, F., Kaufmann, Y. M. and Höfling, V. (2017). Competence feedback improves CBT competence in trainee therapists: a randomized controlled pilot study. *Psychotherapy Research*, 27, 501–509.
- Weiss, D. S. and Marmar, C. R. (1996). The Impact of Event Scale– Revised. In J. Wilson and T. M. Keane (eds), *Assessing Psychological Trauma and PTSD* (pp. 399–411). New York: Guilford.
- Westbrook, D., Sedgwick-Taylor, A., Bennett-Levy, J., Butler, G. and 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.
- Young, J. E. and Beck, A. T. (1980). *Cognitive Therapy Scale: Rating manual*. Unpublished manuscript. Philadelphia, PA, USA: University of Pennsylvania.
- Zarafonitis-Müller, S., Kuhr, K. and Bechdorf, A. (2014). The relationship between therapist's competence and adherence to outcome in cognitive-behavioural therapy– results of a meta-analysis. *Fortschritte der Neurologie Psychiatrie*, 82, 502–510.