ORIGINAL ARTICLE

The Sustained Effects of CBT Training on Therapist Competence and Patient Outcomes

Sheena Liness[1] · Sarah Beale[1] · Susan Lea[2] · Suzanne Byrne[1] · Colette R. Hirsch[1] · David M. Clark[3]

Abstract
The continued effectiveness of in-service psychological therapy training requires evaluation. This study assessed therapist competence and patient clinical outcome during cognitive behaviour therapy training and 12+ months post-training. Trainee competence was assessed using audio-recorded sessions rated on the Cognitive Therapy Scale Revised at the beginning ($n = 33$) and end of training ($n = 45$), and at least 12 months post-training ($n = 45$). Pre-to-posttreatment clinical outcome for trainees' patients during the course ($n = 360$) and post-training ($n = 360$) was evaluated using standardised self-report measures. The relationship between therapist competence and patient outcomes was explored. Trainees achieved competence during training (100%, $n = 45$) and largely maintained competence post-training (84%, $n = 38$). Patients demonstrated pre-to-posttreatment effect sizes between 1.38 and 1.89 and reliable improvement exceeding 80% during and after training. Competence was not significantly associated with patient outcome. Trainees predominantly maintained competence and achieved good clinical outcomes post-training. Structured training and continued use of regular supervision possibly supported retention of competence.

Keywords Cognitive behavioural therapy (CBT) · IAPT · Follow-up · Competence · Patient outcomes · Supervision

Introduction

The delivery of psychological therapies at scale is an important international issue, with many countries developing programmes for the dissemination of evidence-based practice across common mental health problems. If these are to be successful, systems for training large numbers of therapists to competence (benchmarked to a required standard) must be developed and evaluated. Whilst evaluating the acquisition of therapy competence and patient outcomes during training

is important, monitoring the maintenance of competence and patient outcome post-training in the practice setting is essential to realise the public benefit.

Over the last decade, two national training programmes with different training models have substantially increased the availability of evidence-based treatment with promising results. In the UK, the Improving Access to Psychological Therapies (IAPT) initiative, which focuses on evidence-based treatments for depression and anxiety disorders, has reported large pre-post treatment effect sizes for depression and anxiety cases (Clark 2011, 2018a) with large numbers of therapists trained to an agreed level of competence (Branson et al. 2015; Liness et al., under review). In the USA, the Department of Veteran Affairs (VA) focuses on evidence-based treatments for posttraumatic stress disorder and related conditions, with over 10,000 practitioners trained (Rosen et al. 2017) and patient clinical improvement with medium to large effect sizes across disorders (Eftekhari et al. 2013; Karlin et al. 2012). Availability of treatment for common mental health disorders is also being scaled up in low and middle-income countries (Lund et al. 2016; Singla et al. 2017). Promising evidence indicates that non-specialist practitioners can be trained to deliver effective treatment in these

Joint last authors Colette R. Hirsch and David M. Clark due to equal contributions to the manuscript.

✉ Sheena Liness
sheena.liness@kcl.ac.uk

[1] Department of Psychology, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, PO77, Henry Wellcome Building, De Crespigny Park, London SE5 8AF, UK

[2] University of Hull, Hull, UK

[3] Department of Experimental Psychology, University of Oxford, Oxford, UK

settings (Patel et al. 2017; Singla et al. 2017). Whilst a positive picture is emerging for the dissemination of evidence-based therapies, success must be measured in the longer-term by assessing therapists' clinical competence once they are fully embedded within the clinical services. Two inter-related aspects of therapist skill are involved in assessing the standard of therapy: (1) therapy adherence—whether the therapist implemented the relevant procedures (2) therapy competence—whether procedures were implemented in a competent manner (Blackburn et al. 2001). These are commonly seen as requisite for the effective delivery of therapy in routine practice (Barber et al. 2003; Fairburn and Cooper 2011; Muse and McManus 2013). If training is to demonstrate sustained effects, it is important that therapy adherence and competence, and good clinical outcomes are maintained. This study focuses on therapy competence and clinical outcomes.

Information on the sustained use of evidence-based interventions in the workplace after training has been captured in surveys and questionnaires (Karlin et al. 2012; Liness et al. 2017; Rosen et al. 2016), with trainees self-reporting continued use. However, self-report may be an inaccurate reflection of therapeutic practice (Walfish et al. 2012). Psychological therapists may also drift from evidence-based interventions following qualification (Waller and Turner 2016), thus demonstrating the need for on-going evaluation.

The few studies employing independent observer ratings on clinical assessment to investigate long-term maintenance of trainees' competence and patient outcomes have generated mixed results. CBT competence and patient outcome improved pre- to 6 months post-training and were maintained at 12-months follow-up for 12 therapists who received brief training in the treatment of depression in the community (Simons et al. 2010). A comparison of three CBT training methods similarly found that adherence and competence—assessed via structured role play—improved pre-to- post-training, and that therapists in a seminar plus supervision and web condition retained gains following training (Sholomskas et al. 2005). Additionally, a study of low-intensity practitioners delivering self-help CBT in the IAPT programme reported significantly better clinical outcome 12 months after training (Branson et al. 2018). However, novice therapists followed-up at 6-, 12-, and 18-months post-training in computer-assisted CBT for anxiety (Brown et al. 2013) demonstrated significant reductions in adherence and a non-significant trend towards reduction in competence over time; booster-training was proposed for novice therapists to limit potential decline in protocol adherence.

The influence of therapist competence on patient outcome presents as a mixed picture (Webb et al. 2010) with some evidence for a relationship to outcome in CBT for depression (Zarafonitis-Müller et al. 2014) and social anxiety disorder (Ginzburg et al. 2012). Higher CBT competence, but not adherence, was associated with better clinical outcomes in novice therapists (Brown et al. 2013). No general relationship was found on UK IAPT-based training courses (Branson et al. 2015; Jolley et al. 2015) apart from a small relationship in depression cases (Liness et al., under review) and some evidence that patients treated by the most competent therapists are more likely to achieve reliable improvement in symptoms (Branson et al. 2015, 2018). Given the mixed findings, small sample sizes, and varied methods, measures, and focus of these studies, further follow-up studies are essential to assess the long-term benefits of large-scale training initiatives, the role of therapist competence, and the transfer of skills from training to the workplace.

Training, Supervision and Service Integration

A work environment and structure that supports training initiatives and encourages therapist development may enhance the transfer of skills learned and maintenance of competence (Beidas and Kendall 2010; Stirman et al. 2016). Within that structure, the importance of on-going consultation and supervision post-training has been highlighted (Edmunds et al. 2013; Herschell et al. 2010; Miller et al. 2004); however, empirical support for the role of supervision in CBT training is limited. A systematic review examining the effects of supervision on CBT therapist competence across five studies reported potential effects on novice therapists' competence (Alfonsson et al. 2018). Palliative care practitioners randomised to supervision vs no supervision post-brief CBT training differed significantly in skills 6-months post-training, with the no supervision group decreasing and the supervision group increasing (Mannix et al. 2006). Consultation following brief CBT training for adolescent anxiety was deemed essential, as number of consultation hours predicted greater therapist adherence and skill at 3-month follow-up (Beidas et al. 2012).

Informed by emerging evidence, supervision or consultation is a key part of both the IAPT and VA programmes. Weekly ninety-minute group telephone consultation with an expert clinician is built into the VA training model (Karlin et al. 2012). IAPT training incorporates weekly 90-min group supervision for 30 weeks from a CBT therapist accredited by the British Association of Behavioural and Cognitive Psychotherapy (BABCP) as part of the course (Department of Health [DoH] 2018). Both the VA and IAPT supervision models involve case review, feedback on audio-taped sessions and a focus on skill development. IAPT trainees also receive weekly individual supervision from a BABCP-accredited CBT therapist at their work placement site, and evidence 70 h of supervision by the end of training. Supervisors in IAPT services also attend supervision training provided by their trainees' university to enhance collaboration and to encourage comparable supervision methods

during and after training. IAPT graduates who remain in IAPT services continue to receive regular supervision from a BABCP-accredited CBT therapist. Given the specific design of the IAPT initiative, and emerging evidence for the importance of on-going supervision and consultation, exploration of this topic at follow-up is warranted.

IAPT CBT Training

IAPT CBT trainees work in IAPT services whilst being trained in evidence-based CBT for adults with depression and anxiety disorders. They attend university two days per week and work in IAPT clinical-services three days per week, where they maintain a caseload of twelve cases. Training follows a national curriculum (DoH 2011) covering behavioural activation and cognitive therapy models for depression, and evidence-based protocols for anxiety disorders (Roth and Pilling 2008). Trainees are required to treat a range of disorders and are formally assessed on both depression and a range of anxiety cases. Therapists in-training are closely monitored, with regular skills assessments, on-going informal and formal skills feedback, weekly clinical supervision involving live review of therapy clips and whole sessions using the Cognitive Therapy Scale -Revised (CTS-R; Blackburn et al. 2001), and patient outcome measure monitoring. They are encouraged to see training cases, defined as straightforward anxiety or depression presentations with clear main problem descriptors, to enable the trainee to learn National Institute for Health and Care Excellence (NICE) guidance evidence-based protocols for specific difficulties (Roth and Pilling 2008). CBT trainees submit eight cases for formal assessment. IAPT trainees are employed on 1-year training contracts, with an expectation that trainees will continue to work in IAPT services post-training, usually with an increase to full-time work with larger and more diverse caseloads. The intensive training, time, and resources invested in IAPT and other large-scale workforce initiatives, and the lack of studies evaluating the long-term benefits of training, makes follow-up assessment important. This study was set up to investigate therapist competence and patient outcomes during training and post-training when working as qualified CBT therapists.

Aims and Hypotheses

This study aimed to assess if IAPT trainees' CBT competence gained during training was maintained after training. It also set out to evaluate trainees' patient outcomes during and after training. It was predicted that

(1) the level of CBT therapy competence attained during training would be maintained post-training

(2) trainees would achieve good patient outcome in training and at follow-up.

The relationship between therapist competence to patient outcome was also explored, and descriptive information on patient demographics and supervision provision at follow-up was gathered to evaluate continuity of practice in the workplace.

Method

Participants

Therapists were 45 former trainees from seven cohorts (2008–2015) of the IAPT Postgraduate Diploma in CBT at King's College London. Eligibility criteria were course completion by September 2015, and current employment in a service with ethics approval. Exclusion criteria were primary employment in a non-clinical role at follow-up. Therapist were 73% ($n = 33$) female and 27% ($n = 12$) male, and 91% ($n = 41$) were white and 9% ($n = 4$) were from a black and/or minority ethnic background. Therapist professions were clinical psychologist (35%, $n = 16$), psychological wellbeing practitioner (31%, $n = 14$), mental health nurse (11%, $n = 5$), counsellor/psychotherapist (9%, $n = 4$), counselling psychologist (7%, $n = 3$) and other (7%, $n = 3$). Median age at training was 33.00 years (IQR = 9.00).

Former trainees were contacted at least 1-year following course completion regarding study participation. Due to timing of ethics approval, trainees clustered into two groups based on time to follow-up: prospective follow-up after 12 months (42%, $n = 19$) and retrospective follow-up after 12+ months to 6 years (58%, $n = 26$). Of 119 contacted trainees, 45% ($n = 54$) agreed to participate. Of those who declined, 43% ($n = 28$) declined due to workload, 34% ($n = 22$) due to life-events, and 23% ($n = 15$) did not respond. Of those who agreed to participate, 83% ($n = 45$) completed the study. All nine non-completers withdrew due to workload. There were no significant differences in age, gender, profession, pre-training CTS-R scores, mean end-of-training CTS-R scores, or final course result between trainees who completed the study and those who declined to participate or those who withdrew, $p > .05$.

Measures

Therapist competence was evaluated via assessor ratings of audio-recorded therapy sessions with the 12-item Cognitive Therapy Scale-Revised (CTS-R; Blackburn et al. 2001): score range = 0–72; competence threshold set at ≥ 36. Items are rated from 0 to 6 (0 = non-competent, 6 = expert). Items can be separated into two sub-scales: general therapy skills

(Items 1–5) and CBT-specific skills (Items 6–12). The CTS-R has been criticized for inconsistent inter-rater reliability (Gordon 2006; Reichelt et al. 2003; Rozek et al. 2018), a lack of focus on specific mental health problems, and the lack of an empirically validated competence threshold (Fairburn and Cooper 2011; Muse and McManus 2013), although a score of ≥ 36 (item mean ≥ 3) is recommended in the manual. However, the CTS-R is the most widely-used measure of CBT competence on UK training courses and incorporates assessment of adherence as well as competence within item descriptors (James et al. 2001). A more recent study (Kazantzis et al. 2018) reported acceptable internal consistency and inter-rater reliability. In the present study, internal inter-rater reliability was high between pairs of the twenty course markers for a random sample of 25% ($n = 31$) of the 123 submitted tapes: one-way random single measures ICC $(31,30) = 0.85$. Lack of resources mean that in most studies course tutors or supervisors are used to rate tapes; however, in the present study a randomly-selected 18% ($n = 22$) of tapes were second-rated by an external expert who was blind to time (pre-training, post-training, or follow-up). Inter-rater reliability between course and external markers was high: ICC $(22, 21) = 0.80$.

Clinical outcomes were evaluated using patient self-report questionnaires routinely collected at the start and end of therapy. Depression symptoms for all cases were assessed with the Patient Health Questionnaire (PHQ-9; Kroenke et al. 2001). Anxiety symptoms for all cases were assessed with the Generalized Anxiety Disorder-7 (GAD-7; Spitzer et al. 2006). If provided, anxiety disorder-specific symptoms were assessed using one of the following measures: Social Phobia Inventory (SPIN; Connor et al. 2000), Penn State Worry Questionnaire (PSWQ; Meyer et al. 1990), Obsessive Compulsive Inventory (OCI; Foa et al. 1998), Panic Disorder Severity Scale (PDSS; Shear et al. 1997), Mobility Inventory (MI; Chambless et al. 1985), Impact of Events Scale-Revised (IES-R; Weiss and Marmar 1996), Short Health Anxiety Inventory (sHAI; Salkovskis et al. 2002).

Case complexity was benchmarked against course criteria for complexity, namely primary or comorbid psychosis, personality disorder, autism spectrum disorder, substance dependence, severe and/or treatment-relevant physical health conditions, and severe psychosocial difficulties. Two researchers from the present study rated cases as complex or non-complex based on these criteria extracted from case information. Inter-rater reliability was acceptable, kappa = 0.67, $p < .001$ (McHugh 2012). All disagreed cases were discussed and an agreed rating was reached.

Design and Procedure

This study employed an observational longitudinal design, assessing trainees' therapy competence at the start and end

of training and at follow-up, and assessing patient outcomes on the course and at follow-up. Therapists and patients provided informed consent for their audio-recordings and clinical data to be used anonymously and confidentially for research purposes.

Therapy Recordings

Trainees submitted therapy audio-recordings of a recent treatment session at the beginning of the course (Baseline), towards the end of training (End-of-Training), and at follow-up (Follow-Up). At all time points, trainees were requested to submit for assessment a representative session rather than their strongest. Only 33 Baseline recordings were available as the baseline evaluation was introduced in 2009; 45 recordings of sessions were available for end-of-training and at follow-up. All course and follow-up therapy recordings were rated by course staff who were accredited CBT therapists and supervisors with formal training and extensive experience rating the CTS-R; assessors were blind to clinical outcome, subject, and session number. Additional CTS-R rater training was not provided for the study as all markers were course staff who had already been trained to use the CTS-R for assessment. All staff who join the course are provided with induction training in using the CTS-R, which includes blind marking multiple recordings with an experienced marker until high inter-rater agreement is achieved. Following the initial training, on-going reliability monitoring of CTS-R ratings comprises second-ratings on all tapes below the competence threshold (CTS-R < 36) and on a random selection of approximately 20% for each tape submission to check rater drift. As reported above, both internal and external interrater reliability was high for randomly-selected tapes in the present study.

Each recording was scored by a randomly selected course marker. Written reviews and ratings of sessions were provided to trainees within a month of submission to encourage incorporating feedback into clinical practice; markers were thus not blinded to the stage of the study.

Clinical Cases

During the course, trainees identified eight cases for on-going supervision and wrote them up as clinical case reports, including outcome measures, after the patients had completed treatment. Correspondingly, participants at follow-up provided clinical outcome data from their eight most recent cases. At follow-up, trainees also provided case information regarding number of treatment sessions and physical and psychosocial comorbidities, and the frequency and content of their current supervision. Data were available for 720 cases in total, with 360 from during the course and 360 at follow-up. Primary diagnoses of the cases on the course

were anxiety disorders (73%, $n = 264$) and depression (27%, $n = 96$). Primary diagnoses at follow-up were anxiety disorders (58%, $n = 209$), depression (40%, $n = 143$), and other (2%, $n = 8$). Distribution of primary diagnoses differed significantly between course and follow-up cases, exact $X^2 = 23.64$, $p < .001$. There was a significant difference in the proportion of complex cases on the course ($n = 5$, 1%) than at follow-up ($n = 128$, 35%), $X^2 = 139.53$, $p < .001$. Cases from two therapists were excluded due to missing outcome measure data, thus clinical outcomes analyses were conducted based on 43 therapists' data. The Benjamini–Hochberg Procedure (Benjamini and Hochberg 1995) was used to correct for multiple comparisons across all hypothesis tests, with false discovery rate $Q = 0.05$.

Results

Therapist Competence

A repeated-measures analysis of covariance (ANCOVA) found a significant effect of time (Baseline vs End-of-Training vs Follow-Up) on CTS-R scores, controlling for time to follow-up (12 months vs 2–6 years): $F(1.83, 57.02) = 11.46$, $p < .001$, partial $\eta^2 = 0.27$. Post-hoc tests found a significant difference between the estimated marginal means for Baseline ($M = 30.45$, $SD = 4.20$) and End-of-Training ($M = 40.23$, $SD = 2.86$) tapes, $p < .001$, but no significant difference between End-of-Training and Follow-Up ($M = 40.15$, $SD = 3.98$), $p = 0.10$.

Having established an overall effect on total scores, corresponding ANCOVAs for the CTS-R subscales (generic and specific) were conducted. There was a significant effect of time on generic CTS-R subscale scores: $F(2, 62) = 6.17$, $p = 0.004$, partial $\eta^2 = 0.17$. Scores differed significantly between Baseline ($M = 14.04$, $SD = 1.82$) and End-of-Training ($M = 17.10$, $SD = 1.37$) $p < .001$, but not between End-of-Training and Follow-Up ($M = 17.70$, $SD = 1.84$), $p = 0.12$. There was also a significant effect of time on specific CTS-R subscale scores: $F(1.64, 50.98) = 7.42$, $p = 0.003$, partial $\eta^2 = 0.19$. Scores differed significantly between Baseline ($M = 16.41$, $SD = 2.89$) and End-of-Training ($M = 22.26$, $SD = 1.83$) $p < .001$, but not between End-of-Training and Follow-Up ($M = 22.45$, $SD = 2.36$), $p = 0.68$. Figure 1 reports the mean total and subscale CTS-R scores across time; both total and subscale CTS-R scores increased significantly across training and were maintained at follow-up.

Therapist Competence at Follow-Up

At follow-up, 85% ($n = 38$) of therapists remained competent and 15% ($n = 7$) were non-competent. At follow-up, no

Fig. 1 Total and subscale CTS-R scores across time
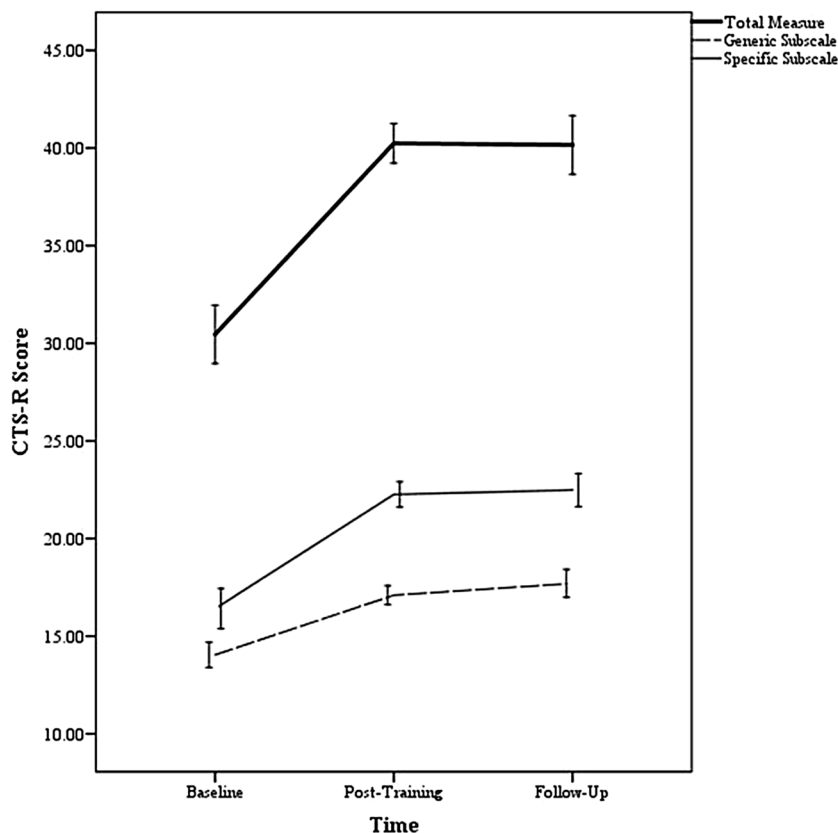
significant ($p > .05$) differences were found between competent and non-competent therapists using χ^2 tests for gender, highest educational qualification, and profession, and Mann–Whitney test for baseline CTS-R total or subscale scores, and end-of-training total or subscale CTS-R scores. Follow-up total and subscale CTS-R scores were significantly greater for competent than non-competent therapists, reported in Table 1.

Clinical Outcomes

Effect sizes on the course and at follow-up were the mean Cohen's *d* for therapists' eight cases based on national standard deviations by diagnosis (depression, anxiety, or other; NHS Digital 2018). Reliable improvement (RI) rates on the course and at follow-up report the percentage of therapists' cases that reduced by the clinically-reliable change index for ≥ 1 measure and did not reliably deteriorate on the other (DoH 2014). Appropriate disorder-specific measures were used in place of the GAD-7 for anxiety disorders by diagnosis (DoH 2014). Table 2 reports repeated-measures

ANCOVAs assessing change in 43 therapists' effect sizes and reliable improvement rates over time (Course vs Follow-Up), controlling for time to follow-up. There were no significant changes in clinical outcomes between the course and follow-up.

Relationship Between CTS-R Scores and Clinical Outcomes

Table 3 reports Spearman correlations between total and subscale CTS-R scores and clinical outcomes for end-of-training and follow-up. There were no significant correlations. There was a positive trend ($p = 0.06$) between the end-of-training total CTS-R and PHQ-9 effect size. At follow-up, the total CTS-R exhibited positive trends ($p = 0.05$–0.09) with RI rate and GAD-7 effect size. There were also trends at follow-up ($p = 0.07$–0.08) between the generic and specific subscales with RI rate, and the specific subscale with GAD-7 effect size.

Mann–Whitney tests found no significant differences between competent and non-competent therapists for clinical

Table 1 Follow-up CTS-R scores for competent and non-competent therapists

	Competent ($n = 38$) Median (*IQR*)	Non-Competent ($n = 7$) Median (*IQR*)	U	p	r
Total	41.75 (3.63)	33.50 (1.50)	0.00	< 0.001	0.62
Generic	18.00 (2.50)	15.50 (2.00)	2.00	< 0.001	0.61
Specific	23.00 (2.63)	19.00 (2.50)	4.50	< 0.001	0.60

CTS-R Cognitive Therapy Scale-Revised (Blackburn et al. 2001)

Table 2 Therapists' clinical outcomes for course and follow-up

	Course (*M, SD*)	Follow-up (*M, SD*)	n	*df*	F	p	Partial η^2
Effect size							
PHQ-9	1.38 (0.50)	1.71 (0.57)	43	1, 41	0.33	0.57	0.008
GAD-7	1.73 (0.63)	1.89 (0.63)	43	1, 41	0.37	0.55	0.009
RI rate	82.56 (16.17)	86.00 (16.86)	43	1, 41	0.03	0.87	0.008

PHQ-9 Patient Health Questionnaire (Kroenke et al. 2001), *GAD-7* Generalized Anxiety Disorder-7 (Spitzer et al. 2006), *RI* reliable improvement

Table 3 Association between CTS-R scores and clinical outcomes

	RI rate ρ (p)	PHQ-9 effect size ρ (p)	GAD-7 effect size ρ (p)
End-of-training total	0.16 (0.31)	0.28 (0.06)	0.24 (0.11)
End-of training generic	0.09 (0.57)	0.00 (0.99)	0.01 (0.95)
End-of-training specific	0.02 (0.88)	0.05 (0.75)	0.11 (0.46)
Follow-up total	0.29 (0.05)	0.18 (0.24)	0.26 (0.09)
Follow-up generic	0.27 (0.07)	0.17 (0.27)	0.23 (0.16)
Follow-up specific	0.28 (0.07)	0.20 (0.19)	0.27 (0.08)

CTS-R Cognitive Therapy Scale-Revised (Blackburn et al. 2001), *RI* reliable improvement, *PHQ-9* Patient Health Questionnaire (Kroenke et al. 2001), *GAD-7* Generalised Anxiety Disorder 7 (Spitzer et al. 2006)

outcomes, reported in Table 4. There was a trend ($p = 0.06$) for greater PHQ-9 effect sizes for competent therapists but overall clinical outcomes were stable from training to follow up in both groups.

Clinical Role and Supervision

During training, therapists received weekly 90-min group supervision from a BABCP-accredited CBT therapist for 30 weeks as part of the course. Supervision content included review of live therapy clips every session and discussion of on-going training cases. Trainees also received weekly individual supervision from an accredited CBT therapist at their work placement site, and were required to evidence 70 h of such supervision by the end of training.

At follow-up, 89% ($n = 40$) of therapists were still employed in IAPT services, and 64% ($n = 29$) of those were employed within the NHS trust in which they had trained, while 11% ($n = 5$) were working in CBT specialist services. Supervision was provided by accredited CBT therapists in line with IAPT procedure. At follow-up, 49% ($n = 22$) of therapists continued to receive weekly and 40% ($n = 18$) fortnightly clinical supervision, with a median duration of 1.5 h ($IQR = 1$ h). Competent therapists did not differ significantly from non-competent therapists in the frequency of supervision received: $U = 124.50$, $p = 0.77$. Unlike course supervision where audio tape therapy review is an integral part of supervision, only 27% ($n = 12$) of therapists reported continued use of live supervision methods, i.e. review of actual therapy sessions at follow-up. While 29% ($n = 11/38$) of the competent therapists at follow-up still reviewed tapes in supervision compared to 14% ($n = 1/7$) of the non-competent therapists, this difference was not significant: exact $X^2 = 0.65$, $p = 0.39$.

Therapists reported offering an adequate course of therapy with a mean of 11.82 ($SD = 3.20$, range $= 8$–19) treatment sessions on the course, and 12.40 ($SD = 5.02$, range 7–23) at follow-up.

Discussion

This study aimed to assess if trainee therapists' gains in competence and clinical outcomes attained during IAPT CBT training were maintained post-qualification in the workplace, and to investigate the relationship between competence and clinical outcomes at both time points. CBT competence was evaluated at the beginning and end of training, and at least 12 months following qualification using the CTS-R. All trainees achieved a recognised minimum level of competence by the end of training, and the majority maintained that level of competence at follow-up. Trainee clinical outcome was also examined, with course and follow-up cases demonstrating significant improvement across all measures despite an increase in complex presentations at follow-up.

CBT Competence

IAPT trainees achieved therapy competence by the end of the course, and the majority maintained competence when working as qualified therapists, notably in line with the expected competence threshold (Blackburn et al. 2001; Branson et al. 2015; Keen and Freeston 2008). While baseline CTS-R scores were lower in the present study ($M = 30.45$) than in some other CBT training studies (35.72–38.94; see Barnfield et al. 2007; Branson et al. 2015; Keen and Freeston 2008), trainees obtained competence in line with previous literature by the end of training. The CTS-R subscales for generic and CBT-specific items also improved and were maintained at follow-up, adding to the evidence for post-training maintenance of both generic and CBT-specific competence (Sholomskas et al. 2005; Simons et al. 2010). The subset who did not maintain competence post-training was small, which makes conclusions difficult. The seven therapists who were non-competent at follow-up nearly reached CTS-R competence (36) with a median of 33.50, perhaps indicating drift rather than major relapse and that additional input may have enabled them to maintain competence.

Table 4 Clinical Outcomes for Competent and Non-Competent Therapists at Follow-Up

	Competent ($n = 38$) Median (IQR)	Non-Competent ($n = 7$) Median (IQR)	U	p	r
RI rate	93.75 (25.00)	87.50 (37.50)	95.50	0.29	0.16
PHQ-9 effect size	1.86 (0.49)	1.48 (1.22)	68.00	0.06	0.29
GAD-7 effect size	1.93 (0.93)	1.82 (1.27)	89.00	0.22	0.19

RI reliable improvement, *PHQ-9* Patient Health Questionnaire (Kroenke et al. 2001), *GAD-7* Generalised Anxiety Disorder 7 (Spitzer et al. 2006)

Patient Outcome

This study adds to the growing literature that trainees obtain excellent patient outcomes during CBT training (Forand et al. 2011, Öst et al. 2012) and on large workforce training programmes (Branson et al. 2015; Eftekhari et al. 2013) by extending it to post-training. It also supports the maintenance of excellent patient outcomes post-training in the face of increased caseload and complexity. Therapists delivered doses of therapy in accordance with NICE (2011) guidance at both time-points, possibly promoting patient gains. High average number of treatment sessions have been found to improve therapy outcomes (Clark et al. 2018b).

Therapist Competence and Patient Outcome

No significant relationship between therapist competence and patient outcome was found at the end of training or at follow up, although positive trends emerged. Previous training studies reported similar results (Branson et al. 2015; Jolley et al. 2015; Liness et al., under review). There was also no significant difference in patient outcome between competent and non-competent therapists at follow up. The absence of significant findings was likely influenced by the small number of non-competent therapists ($n = 7$) at follow-up, a somewhat ironic consequence of so many therapists retaining competence. The lack of an association of therapist competence to patient outcome may be due to the restriction of range in therapist competence (Öst et al. 2012). Other factors that may influence the competence-outcome relationship and patient outcome at any given time including therapeutic alliance (Webb et al. 2010) and patient characteristics such as symptom severity and chronicity, motivation and life events (Clark et al. 2018b; Webb et al. 2010), were not accessible in this study. Investigating the effect of patient diagnosis was not feasible in the current study, but diagnosis may moderate the competence-outcome relationship– with stronger evidence for depression (Kazantzis et al. 2018; Webb et al. 2010; Zarafonitis-Müller et al. 2014) and in a single study using a disorder-specific competence measure for social anxiety (Ginzburg et al. 2012). Additionally, the time-point of the recording in the therapeutic process may influence the competence-outcome relationship (Kazantzis et al. 2018), but was not feasible to manipulate in the present study. Emerging positive trends in this study highlight the need for future investigations aimed at investigating the competence- outcome relationship and moderators therein. Future studies would benefit from larger numbers, varied therapist abilities, diverse patient presentations, varied and disorder-specific competence measures, and evaluation of the competence-outcome link at several time-points in the therapeutic process.

Clinical Role and Supervision

The majority of therapists in this study continued to work in IAPT services post-training and a substantial number remained in the same NHS trust. Patient characteristics at follow-up were similar to training cases, with complexity increased slightly as expected. Trainees continuing to work in similar workplace settings and see similar patient presentations may assist the consolidation of therapist competence, and contrasts with previous large-scale training initiatives where skill maintenance difficulties were partially attributed to a lack of continuity from training to the work environment (Brooker et al. 2003). Therapists' continued use of regular expert supervision may also have supported maintenance of competence (Mannix et al. 2006; Rakovshik et al. 2016). Therapists however reported reduced use of 'live' supervision (i.e. review of therapy tapes). This is a key characteristic of training supervision, and its quick reduction post-qualification is noteworthy as therapists have higher and more complex caseloads. Only 1/7 non-competent therapists reported that supervision included feedback on audio recordings of therapy, possibly indicating the usefulness of ongoing review and feedback of actual therapy sessions after initial training (Bearman et al. 2017; Weck et al. 2017); however, competent and non-competent therapists did not differ significantly in post-training access to live supervision or in frequency of supervision overall. Small numbers and the lack of a control condition make conclusions difficult. Further investigation is needed on a larger sample post-training.

Limitations

Present findings have the constraints of a naturalistic study with no control condition. Therapists self-selected audio-taped sessions that were submitted for assessment and also patients that they brought to weekly supervision and subsequently included in their case reports. The representativeness of both the audio sessions and patients is unknown as submitted cases only partially reflect overall caseloads. However, trainees were instructed to select representative audio tapes, and patient selection occurred before clinical outcome was known. Clinical outcome at follow-up was requested for the most recent eight cases irrespective of outcome. Several therapists were unable to complete the study due to workload, consequently the sample size was smaller than intended. Delay in ethics approval also affected participation and led to varied timings of follow up submissions for the earlier cohorts. Recruitment, ethics approval and attrition are common difficulties in workplace research (Brosan et al. 2007; Stirman et al. 2015). Further evaluation of training effects and maintenance of competence is crucial for the sustainability of workforce training such as the IAPT programme, and perhaps merits

increased resources and funding for services to enable timely research and busy therapists' participation.

Conclusions and Future Directions

This is the first study to our knowledge that evaluates CBT therapist competence and patient outcomes from the IAPT workforce training programme during training and at follow-up. The maintenance of therapists' gains post-training is promising for large-scale evidence-based training worldwide. The highly-structured IAPT set-up that encourages training and service collaboration, and emphasizes the importance of regular clinical supervision both during and after training, possibly has a role to play in the sustainability of competence. The generalizability of this structure to less resource-rich environments is unknown and requires further long-term evaluation and access to the necessary resources to increase participation by busy practitioners. The focus on alternative clinical assessment and supervision methods, the training of non-mental health professionals, and the creation of on-line resources are all encouraging developments for embedding evidence-based practices worldwide (Fairburn and Patel 2014; Kobak et al. 2017). Overall, this study demonstrates that a large-scale training programme can equip therapists with sustainable competence to address the challenge of mental health in the population.

Compliance with Ethical Standards

Conflict of Interest Sheena Liness and Suzanne Byrne run the IAPT CBT training at the IoPPN/KCL course that is the subject of this study. David M. Clark is NHS England's Clinical Advisor for the IAPT programme. Susan Lea, Colette Hirsch and Sarah Beale declare that they have no conflict of interest.

Informed Consent All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Animal Rights No animal studies were carried out by the authors for this article.

References

Alfonsson, S., Parling, T., Spännargård, Å, Andersson, G., & Lundgren, T. (2018). The effects of clinical supervision on supervisees and patients in cognitive behavioral therapy: A systematic review. *Cognitive Behaviour Therapy, 47*(3), 206–228.

Barber, J. P., Liese, B. S., & Abrams, M. J. (2003). Development of the cognitive therapy adherence and competence scale. *Psychotherapy Research, 13*(2), 205–221.

Barnfield, T. V., Mathieson, F. M., & Beaumont, G. R. (2007). Assessing the development of competence during postgraduate cognitive-behavioral therapy training. *Journal of Cognitive Psychotherapy, 21*(2), 140.

Bearman, S. K., Schneiderman, R. L., & Zoloth, E. (2017). Building an evidence base for effective supervision practices: An analogue experiment of supervision to increase EBT fidelity. *Administration and Policy in Mental Health and Mental Health Services Research, 44*(2), 293–307.

Beidas, R. S., Edmunds, J. M., Marcus, S. C., & Kendall, P. C. (2012). Training and consultation to promote implementation of an empirically supported treatment: A randomized trial. *Psychiatric Services, 63*(7), 660–665.

Beidas, R. S., & Kendall, P. C. (2010). Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice, 17*(1), 1–30.

Benjamini, Y., & 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*(1), 289–300.

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*(04), 431–446.

Branson, A., Myles, P., Mahdi, M., & Shafran, R. (2018). The relationship between competence and patient outcome with low-intensity cognitive behavioural interventions. *Behavioural and Cognitive Psychotherapy, 46*(1), 101–114.

Branson, A., Shafran, R., & Myles, P. (2015). Investigating the relationship between competence and patient outcome with CBT. *Behaviour Research and Therapy, 68*, 19–26.

Brooker, C., Saul, C., Robinson, J., King, J., & Dudley, M. (2003). Is training in psychosocial interventions worthwhile? Report of a psychosocial intervention trainee follow-up study. *International Journal of Nursing Studies, 40*(7), 731–747.

Brosan, L., Reynolds, S., & Moore, R. G. (2007). Factors associated with competence in cognitive therapists. *Behavioural and Cognitive Psychotherapy, 35*(2), 179–190.

Brown, L. A., Craske, M. G., Glenn, D. E., Stein, M. B., Sullivan, G., Sherbourne, C., Bystriksy, A., Welch, S., Campbell-Sills, L., Lang, A., Roy-Byrne, P., & Rose, R. (2013). CBT competence in novice therapists improves anxiety outcomes. *Depression and Anxiety, 30*(2), 97–115.

Chambless, D. L., Caputo, G. C., Jasin, S. E., Gracely, E. J., & Williams, C. (1985). The mobility inventory for Agoraphobia. *Behaviour Research and Therapy, 23*(1), 35–44.

Clark, D. M. (2011). Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: The IAPT experience. *International Review of Psychiatry, 23*(4), 318–327.

Clark, D. M. (2018a). 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., & Janecka, M. (2018b). Transparency about the outcomes of mental health services (IAPT approach): An analysis of public data. *The Lancet, 391*(10121), 679–686.

Connor, K. M., Davidson, J. R., Churchill, L. E., Sherwood, A., Weisler, R. H., & Foa, E. (2000). Psychometric properties of the Social Phobia Inventory (SPIN). *The British Journal of Psychiatry, 176*(4), 379–386.

Department of Health. (2011). *National curriculum for high intensity cognitive behavioural therapy courses*. Leeds: NHS Digital.

Department of Health. (2014). *Improving access to psychological therapies: Measuring improvement and recovery—adult Services, Version 2*. Leeds: NHS Digital.

Department of Health. (2018). *The improving access to psychological therapies manual*. Leeds: NHS Digital.

Edmunds, J. M., Beidas, R. S., & Kendall, P. C. (2013). Dissemination and implementation of evidence-based practices: Training and consultation as implementation strategies. *Clinical Psychology: Science and Practice, 20*(2), 152–165.

Eftekhari, A., Ruzek, J. I., Crowley, J. J., Rosen, C. S., Greenbaum, M. A., & Karlin, B. E. (2013). Effectiveness of national implementation of prolonged exposure therapy in veterans affairs care. *JAMA Psychiatry, 70*(9), 949–955.

Fairburn, C. G., & Cooper, Z. (2011). Therapist competence, therapy quality, and therapist training. *Behaviour Research and Therapy, 49*(6–7), 373–378.

Fairburn, C. G., & Patel, V. (2014). The Global Dissemination of Psychological Treatments: A road map for research and practice. *American Journal of Psychiatry, 171*(5), 495–498.

Foa, E. B., Kozak, M. J., Salkovskis, P. M., Coles, M. E., & Amir, N. (1998). The Validation of a New Obsessive-Compulsive Disorder Scale: The Obsessive-Compulsive Inventory. *Psychological Assessment, 10*(3), 206–214.

Forand, N. R., Evans, S., Haglin, D., & Fishman, B. (2011). Cognitive behavioral therapy in practice: Treatment delivered by trainees at an outpatient clinic is clinically effective. *Behavior Therapy, 42*(4), 612–623.

Ginzburg, D. M., Bohn, C., Höfling, V., Weck, F., Clark, D. M., & Stangier, U. (2012). Treatment specific competence predicts outcome in cognitive therapy for social anxiety disorder. *Behaviour Research and Therapy, 50*(12), 747–752.

Gordon, P. K. (2006). A comparison of two versions of the Cognitive Therapy Scale. *Behavioural and Cognitive Psychotherapy, 35*(3), 343–353.

Herschell, A. D., Kolko, D. J., Baumann, B. L., & Davis, A. C. (2010). The role of therapist training in the implementation of psychosocial treatments: A review and critique with recommendations. *Clinical Psychology Review, 30*(4), 448–466.

James, I. A., Blackburn, I.-M., & Reichelt, F. K. (2001). *Manual of the Revised Cognitive Therapy Scale (CTS-R)*. Unpublished Manuscript, Newcastle Cognitive and Behavioural Therapies Centre, Newcastle.

Jolley, S., Onwumere, J., Bissoli, S., Bhayani, P., Singh, G., Kuipers, E., Craig, T., & Garety, P. (2015). A pilot evaluation of therapist training in cognitive therapy for psychosis: Therapy quality and clinical outcomes. *Behavioural and Cognitive Psychotherapy, 43*(4), 478–489.

Karlin, B. E., Brown, G. K., Trockel, M., Cunning, D., Zeiss, A. M., & 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*(5), 707–718.

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*(4), 385–397.

Keen, A. J., & Freeston, M. H. (2008). Assessing competence in cognitive-behavioural therapy. *The British Journal of Psychiatry, 193*(1), 60–64.

Kobak, K. A., Wolitzky-Taylor, K., Craske, M. G., & Rose, R. D. (2017). Therapist Training on cognitive behavior therapy for anxiety disorders using Internet-based technologies. *Cognitive Therapy and Research, 41*(2), 252–265.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9. *Journal of General Internal Medicine, 16*(9), 606–613.

Liness, S., Beale, S., Lee, S., Byrne, S., Hirsch, C. R., & Clark, D. M. (under review). *Multi-professional IAPT CBT training: Clinical competence and patient outcomes*. Manuscript submitted for publication.

Liness, S., Lea, S., Nestler, S., Parker, H., & Clark, D. M. (2017). What IAPT CBT high-intensity trainees do after training. *Behavioural and Cognitive Psychotherapy, 45*(1), 16–30.

Lund, C., Tomlinson, M., & 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*(56), s1–s3.

Mannix, K. A., Blackburn, I. M., Garland, A., Gracie, J., Moorey, S., Reid, B., Standart, S., & Scott, J. (2006). Effectiveness of brief training in cognitive behaviour therapy techniques for palliative care practitioners. *Palliative Medicine, 20*(6), 579–584.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy, 28*(6), 487–495.

Miller, W. R., Yahne, C. E., Moyers, T. B., Martinez, J., & Pirritano, M. (2004). A randomized trial of methods to help clinicians learn motivational interviewing. *Journal of Consulting and Clinical Psychology, 72*(6), 1050–1062.

Muse, K., & McManus, F. (2013). A systematic review of methods for assessing competence in cognitive–behavioural therapy. *Clinical Psychology Review, 33*(3), 484–499.

National Institute of Health and Care Excellence (NICE). (2011). *Common mental health disorders: Identification and pathways to care*. Retrieved from Clinical Guideline CG123. https://www.nice.org.uk/guidance/cg123/evidence/full-guideline-181771741.

NHS Digital. (2018). *Psychological therapies: Annual report on the use of IAPT services, England 2016–2017*. Leeds: Health and Social Care Information Centre.

Öst, L. G., Karlstedt, A., & Widén, S. (2012). The effects of cognitive behavior therapy delivered by students in a psychologist training program: An effectiveness study. *Behavior Therapy, 43*(1), 160–173.

Patel, V., Weobong, B., Weiss, H. A., Anand, A., Bhat, B., Katti, B., Dimidjian, S., Araya, R., Hollon, S. D., King, M., & Vijayakumar, L. (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*(10065), 176–185.

Rakovshik, S. G., McManus, F., Vazquez-Montes, M., Muse, K., & Ougrin, D. (2016). Is supervision necessary? Examining the

effects of internet-based CBT training with and without super-vision. *Journal of Consulting and Clinical Psychology, 84*(3), 191–199.

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*(2), 87–99.

Rosen, C. S., Matthieu, M. M., Stirman, S. W., Cook, J. M., Landes, S., Bernardy, N. C., Chard, K. M., Crowley, J., Eftekhari, A., Finley, E. P., Hamblen, J. L., Harik, J. M., Kehle- Forbes, S. M., Meis, L. A., Osei-Bonsu, P. E., Rodriguez, A. L., Ruggiero, K. J., Ruzek, J. I., Smith, B. N., Trent, L., & 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*(6), 957–977.

Rosen, R. C., Ruzek, J. I., & 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.

Roth, A. D., & Pilling, S. (2008). Using an evidence-based methodology to identify the competences required to deliver effective cognitive and behavioural therapy for depression and anxiety disorders. *Behavioural and Cognitive Psychotherapy, 36*(02), 129–147.

Rozek, D. C., Serrano, J. L., Marriott, B. R., Scott, K. S., Hickman, L. B., Brothers, B. M., Lewis, C. C., & Simons, A. D. (2018). Cognitive behavioural therapy competency: Pilot data from a comparison of multiple perspectives. *Behavioural and Cognitive Psychotherapy, 46*(2), 244–250.

Salkovskis, P. M., Rimes, K. A., Warwick, H. M. C., & Clark, D. M. (2002). The Health Anxiety Inventory: Development and validation of scales for the measurement of health anxiety and hypochondriasis. *Psychological Medicine, 32*(05), 843–853.

Shear, M. K., Brown, T. A., Barlow, D. H., Money, R., Sholomskas, D. E., Woods, S. W., Gorman, J. M., & Papp, L. A. (1997). Multicenter Collaborative Panic Disorder Severity Scale. *American Journal of Psychiatry, 154*(11), 1571–1575.

Sholomskas, D. E., Syracuse-Siewert, G., Rounsaville, B. J., Ball, S. A., Nuro, K. F., & Carroll, K. M. (2005). We don't train in vain: A dissemination trial of three strategies of training clinicians in cognitive-behavioral therapy. *Journal of Consulting and Clinical Psychology, 73*(1), 106–115.

Simons, A. D., Padesky, C. A., Montemarano, J., Lewis, C. C., Murakami, J., Lamb, K., DeVinney, S., Reid, M., Smith, D. A., & 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*(5), 751–756.

Singla, D. R., Kohrt, B. A., Murray, L. K., Anand, A., Chorpita, B. F., & 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., & Löwe, B. (2006). A brief measure for assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092–1097.

Stirman, S. W., Gutner, C., Crits-Christoph, P., Edmunds, J., Evans, A. C., & Beidas, R. S. (2015). Relationships between clinician-level attributes and fidelity-consistent and fidelity-inconsistent modifications to an evidence-based psychotherapy. *Implementation Science, 10*(1), 115.

Stirman, S. W., Gutner, C. A., Langdon, K., & Graham, J. R. (2016). Bridging the gap between research and practice in mental health service settings: An overview of developments in implementation theory and research. *Behavior Therapy, 47*(6), 920–936.

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports, 110*(2), 639–644.

Waller, G., & Turner, H. (2016). Therapist drift redux: Why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. *Behaviour Research and Therapy, 77*, 129–137.

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*(2), 200–211.

Weck, F., Kaufmann, Y. M., & Höfling, V. (2017). Competence feedback improves CBT competence in trainee therapists: A randomized controlled pilot study. *Psychotherapy Research, 27*(4), 501–509.

Weiss, D. S., & Marmar, C. R. (1996). The Impact of Event Scale—Revised. In J. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 399–411). New York: Guilford.

Zarafonitis-Müller, S., Kuhr, K., & Bechdolf, 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*(9), 502–510.