Citation for published version (APA):
Multi-professional IAPT CBT training: clinical competence and patient outcomes

Sheena Liness1,*, Sarah Beale1, Susan Lea2, Suzanne Byrne1, Colette R. Hirsch1,§ and David M. Clark3,§

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)

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.

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).
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 (PWPs; 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 PWPs), ‘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.
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 profession\(^1\): 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 PWPs \((27.38\%, n = 69)\) 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 \((n = 13)\).

\(^1\)Core 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](http://www.babcp.com/Accreditation/Core-Profession.aspx)

\(^2\)Trainees who failed a therapy tape were offered additional support constituting three clinical skills tutorials and informal feedback on an additional tape.
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>
<thead>
<tr>
<th>Disorder</th>
<th>Measure</th>
<th>Items $n$ (range)</th>
<th>Caseness threshold</th>
<th>Reliable change index</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAD</td>
<td>PSWQ (Meyer et al., 1990)</td>
<td>16 (0–80)</td>
<td>≥45</td>
<td>≥7</td>
</tr>
<tr>
<td>PTSD</td>
<td>IES-R (Weiss and Marmar, 1996)</td>
<td>22 (0–88)</td>
<td>≥33</td>
<td>≥9</td>
</tr>
<tr>
<td>OCD</td>
<td>OCI (Foa et al., 1998)</td>
<td>42 (0–168)</td>
<td>≥40</td>
<td>≥32</td>
</tr>
<tr>
<td>Health anxiety</td>
<td>s-HAI (Salkovskis et al., 2002)</td>
<td>14 (0–54)</td>
<td>≥18</td>
<td>≥4</td>
</tr>
<tr>
<td>Social anxiety</td>
<td>SPIN (Connor et al., 2000)</td>
<td>17 (0–68)</td>
<td>≥19</td>
<td>≥10</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>PDSS (Shear et al., 2001)</td>
<td>7 (0–28)</td>
<td>≥8</td>
<td>n/a</td>
</tr>
</tbody>
</table>

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.

Table 1. Disorder-specific clinical outcome measures
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: \( F_{\text{total}} (2,198) = 93.78, p < .001, \text{partial } \eta^2 = .49; F_{\text{generic}} (2,198) = 49.03, p < .001, \text{partial } \eta^2 = .33; \) and \( F_{\text{specific}} (2,198) = 102.04, p < .001, \text{partial } \eta^2 = .72. \) 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: \( F_{\text{total}} (2,116) = 57.32, p < .001, \text{partial } \eta^2 = .50; F_{\text{generic}} (2,116) = 32.66, p < .001, \text{partial } \eta^2 = .36; \) and \( F_{\text{specific}} (2,116) = 64.62, p < .001, \text{partial } \eta^2 = .76. \) 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...
Gains in competence were significant for both the generic \[ t (42) = -10.92, p < .001 \] 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

\(^4\)As 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>
<thead>
<tr>
<th></th>
<th>Clinical cases n</th>
<th>Total cases n</th>
<th>Mean\textsubscript{pre} (SD\textsubscript{pre})</th>
<th>Mean\textsubscript{post} (SD\textsubscript{post})</th>
<th>d.f.</th>
<th>t</th>
<th>d\textsuperscript{a}</th>
<th>Recovery n (%)</th>
<th>Reliable improvement n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Depression overall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IAPT criteria\textsuperscript{b}</td>
<td>501</td>
<td>520</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>336 (67.07)</td>
<td>455 (87.67)</td>
</tr>
<tr>
<td>PHQ-9</td>
<td>488</td>
<td>520</td>
<td>17.15 (4.92)</td>
<td>6.60 (5.53)</td>
<td>245</td>
<td>37.03**</td>
<td>1.80</td>
<td>364 (74.59)</td>
<td>423 (81.35)</td>
</tr>
<tr>
<td><strong>Anxiety overall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IAPT criteria\textsuperscript{b}</td>
<td>1134</td>
<td>1406</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>874 (77.07)</td>
<td>1116 (79.37)</td>
</tr>
<tr>
<td>GAD-7</td>
<td>1089</td>
<td>1406</td>
<td>12.19 (5.43)</td>
<td>4.46 (4.42)</td>
<td>245</td>
<td>45.74**</td>
<td>1.38</td>
<td>858 (78.79)</td>
<td>1067 (75.89)</td>
</tr>
<tr>
<td><strong>GAD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSWQ</td>
<td>138</td>
<td>140</td>
<td>67.99 (8.25)</td>
<td>43.11 (13.73)</td>
<td>105</td>
<td>18.57**</td>
<td>1.95</td>
<td>85 (78.79)</td>
<td>131 (93.57)</td>
</tr>
<tr>
<td>GAD-7</td>
<td>129</td>
<td>148</td>
<td>13.76 (4.91)</td>
<td>4.75 (3.96)</td>
<td>110</td>
<td>18.45**</td>
<td>1.95</td>
<td>107 (82.94)</td>
<td>121 (81.76)</td>
</tr>
<tr>
<td><strong>PTSD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IES-R</td>
<td>101</td>
<td>118</td>
<td>52.36 (17.89)</td>
<td>13.23 (17.20)</td>
<td>96</td>
<td>20.46**</td>
<td>2.26</td>
<td>88 (87.13)</td>
<td>111 (94.07)</td>
</tr>
<tr>
<td>GAD-7</td>
<td>106</td>
<td>130</td>
<td>12.59 (5.14)</td>
<td>4.22 (4.92)</td>
<td>103</td>
<td>14.40**</td>
<td>1.60</td>
<td>87 (82.07)</td>
<td>104 (80.00)</td>
</tr>
<tr>
<td><strong>OCD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCI</td>
<td>191</td>
<td>257</td>
<td>61.77 (28.94)</td>
<td>24.60 (24.16)</td>
<td>168</td>
<td>23.97**</td>
<td>1.46</td>
<td>140 (73.30)</td>
<td>136 (52.92)</td>
</tr>
<tr>
<td>GAD-7</td>
<td>244</td>
<td>287</td>
<td>13.28 (4.96)</td>
<td>5.41 (4.80)</td>
<td>183</td>
<td>22.98**</td>
<td>1.58</td>
<td>177 (72.54)</td>
<td>223 (77.70)</td>
</tr>
<tr>
<td><strong>Health anxiety</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s-HAI</td>
<td>115</td>
<td>118</td>
<td>31.45 (8.51)</td>
<td>15.26 (9.10)</td>
<td>99</td>
<td>11.69**</td>
<td>1.08</td>
<td>88 (76.52)</td>
<td>109 (92.37)</td>
</tr>
<tr>
<td>GAD-7</td>
<td>95</td>
<td>126</td>
<td>12.11 (5.37)</td>
<td>3.51 (3.45)</td>
<td>104</td>
<td>16.64**</td>
<td>1.65</td>
<td>80 (84.21)</td>
<td>104 (82.54)</td>
</tr>
<tr>
<td><strong>Social anxiety</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIN</td>
<td>296</td>
<td>303</td>
<td>40.77 (11.79)</td>
<td>16.99 (12.40)</td>
<td>186</td>
<td>28.74**</td>
<td>1.87</td>
<td>186 (62.84)</td>
<td>254 (83.83)</td>
</tr>
<tr>
<td>GAD-7\textsuperscript{c}</td>
<td>233</td>
<td>319</td>
<td>11.21 (5.06)</td>
<td>3.99 (3.76)</td>
<td>194</td>
<td>26.33**</td>
<td>1.62</td>
<td>193 (82.83)</td>
<td>245 (76.80)</td>
</tr>
<tr>
<td><strong>Panic disorder</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDSS</td>
<td>67</td>
<td>74</td>
<td>15.09 (5.48)</td>
<td>4.38 (4.60)</td>
<td>58</td>
<td>13.53**</td>
<td>1.83</td>
<td>51 (76.12)</td>
<td>—</td>
</tr>
<tr>
<td>GAD-7</td>
<td>213</td>
<td>257</td>
<td>13.13 (5.30)</td>
<td>5.01 (5.01)</td>
<td>170</td>
<td>20.88**</td>
<td>1.48</td>
<td>163 (76.53)</td>
<td>196 (76.26)</td>
</tr>
</tbody>
</table>

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.

\*\textit{p} < .001.

\textsuperscript{a}Standard 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).

\textsuperscript{b}IAPT 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).

\textsuperscript{c}Case 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: \(F_{profession} (2,160) = 7.46, p = .001\), partial \(\eta^2 = .08\) and \(F_{time} (1,160) = 396.41, p < .001\), 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\(_{ClinPsy} = 31.26, SD_{ClinPsy} = 4.89\); mean\(_{OtherCore} = 29.18, SD_{OtherCore} = 4.67\); mean\(_{NoCore} = 27.79, SD_{NoCore} = 5.43\)), but significantly increased and attained

---

\(^5\)Regression was performed in Stata version 15 to adjust standard errors for clustering by therapist ID.
competence by end-of-training (mean\text{ClinPsy} = 38.80, SD\text{ClinPsy} = 3.44; mean\text{OtherCore} = 36.62, SD\text{OtherCore} = 4.50; mean\text{NoCore} = 37.40, SD\text{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 PN/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


