*The Cognitive Behaviour Therapist* (2020), vol. 13, e1, page 1 of 12
doi:10.1017/S1754470X19000357

BABCP
www.babcp.com

ORIGINAL RESEARCH

# Trainee self-assessment of cognitive behaviour therapy competence during and after training

Sarah Beale*[,1], Sheena Liness[1,2] and Colette R. Hirsch[1,2]

[1]Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK and [2]Centre for Anxiety Disorders and Trauma, South London and Maudsley NHS Foundation Trust, London, UK
*Corresponding author. Email: sarah.beale@kcl.ac.uk

(Received 9 May 2019; revised 3 September 2019; accepted 7 October 2019)

## Abstract

Large-scale cognitive behavioural therapy (CBT) training and implementation programmes, such as the pioneering Improving Access to Psychological Therapies (IAPT) initiative in the UK, aim to develop a workforce of competent therapists who can deliver evidence-based interventions skilfully. Self-awareness of competence enables CBT therapists to accurately evaluate their clinical practice and determine professional development needs. The accuracy of self-assessed competence, however, remains unclear when compared with assessments conducted by markers with expertise in CBT practice and evaluation. This study investigated the relationship between self- and expert-rated competence – assessed via therapy recordings rated on the Cognitive Therapy Scale Revised (CTS-R) scale – for a large sample of IAPT CBT trainees during training and, for the first time, at post-training follow-up. CBT trainees ($n = 150$) submitted therapy recordings at baseline, mid-training and end-of-training. At 12+ month follow-up, a subset of former trainees ($n = 30$) submitted recordings from clinical practice. There were positive relationships ($r = .27$ to $.56$) between self and expert CTS-R scores at all time points. The proportion of tapes demonstrating significant agreement between self and expert ratings (CTS-R difference <5 points) increased significantly across training and remained stable at follow-up. Findings indicate that accurate self-awareness of competence can be developed during structured CBT training and retained in the workplace. These outcomes are encouraging given the importance of self-awareness to CBT practice and accreditation. Future investigation into the development and maintenance of accurate self-awareness of competence is warranted.

### Key learning aims

(1) What is the relationship between self-ratings and expert ratings of CBT competence during training and at post-training follow-up?
(2) Does agreement between self and expert competence ratings improve with CBT training?
(3) How does agreement between self and expert ratings change across training for more- and less-competent trainees?
(4) Can accurate self-awareness of competence be retained post-training in the workplace?

**Keywords:** cognitive behavioural therapy (CBT); competence; IAPT; self-assessment; training

## Introduction

The Improving Access to Psychological Therapies (IAPT) initiative in England aims to train over 10,500 cognitive behavioural therapy (CBT) therapists by 2025 to address 25% of the national prevalence of anxiety disorders and depression (NHS England, 2016). To meet these ambitious goals and support the expansion of CBT training and provision, effective methods to assess

© British Association for Behavioural and Cognitive Psychotherapies 2020.

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press

therapists' competence, or skilfulness at delivering CBT, are required. Expert objective ratings of therapy recordings are the gold-standard competence assessment, but are resource-intensive (Rozek *et al.*, 2018). Interest in varied assessment methods, including therapist self-assessment of competence, has consequently grown (Rozek *et al.*, 2018; Singla *et al.*, 2014). Valid and feasible methods to assess treatment integrity and therapy quality – including competence – are integral for the large-scale implementation of evidence-based therapies such as CBT (Shoenwald *et al.*, 2011). Self-assessment is considered an 'indirect' alternative to direct objective assessment, with potential feasibility and cost benefits. Consequently, it may be a resource-efficient method for monitoring therapy fidelity in CBT research and practice and/or a useful tool for quality assurance in routine care. Furthermore, different competence assessment methods – including self-report – may be warranted across the wide variety of contexts in which CBT is studied and implemented and thus should be investigated (McLeod *et al.*, 2013).

Self-assessment is also an important element of reflective practice that enables therapists to determine their own strengths and areas for improvement, and therefore to engage effectively with supervision and professional development (Fouad *et al.*, 2009; Kaslow *et al.*, 2009). Consequently, it is considered a core skill for professional accreditation that should be developed during training (BABCP, 2012; Hool, 2010). The acquisition and persistence of accurate CBT self-assessment skills is therefore an important topic for evaluation.

### Self-assessment of competence in experienced therapists

Experienced therapists tend to over-estimate their own ability (Chevron and Rounsaville, 1983; Parker and Waller, 2015; Perlesz *et al.*, 1990; Walfish *et al.*, 2012), particularly the least competent therapists with the lowest objectively assessed skill (Brosan *et al.*, 2008). The 'better than average' bias (Brown, 1986), the tendency to self-rate one's abilities as greater than those of the average person, may lead to therapists over-estimating their competence (Waltman *et al.*, 2016). Therapist drift, the conscious or unconscious failure of experienced clinicians to deliver evidence-based treatment despite available resources (Waller, 2009), may also negatively impact the accuracy of self-assessments (Waltman *et al.*, 2016). The least competent practitioners may demonstrate the greatest over-estimation as they lack awareness of their own lack of skill (Dunning *et al.*, 2003). Over-estimation of competence, particularly by less objectively competent practitioners, is an important issue as therapists with inaccurately high self-perceptions may feel less motivated to seek out continuing professional development, engage in supervision, or implement feedback to improve their practice (Brosan *et al.*, 2008). Consequently, quality of therapy may drop below acceptable standards and yield potentially negative impacts on clients and service outcomes. Interestingly, more competent therapists conversely exhibit lower self-ratings and even under-estimation of competence (McManus *et al.*, 2012), attributed in part to greater self-reflection and awareness of areas requiring further skill development. Professional self-doubt, which may be reflected in lower self-reported CBT competence, has been correspondingly linked to stronger clinical outcomes in psychotherapy than over-confidence (Meichenbaum and Lilienfeld, 2018; Nissen-Lie *et al.*, 2017).

While poor self-awareness in experienced clinicians is concerning and may indicate that self-report is an inadequate method of assessing competence in routine practice, many studies are cross-sectional and have not accounted for the quality and duration of prior training or ongoing supervision. Training that involves regular comparison between self and expert ratings of competence – an established element of IAPT CBT training – may support accurate self-awareness (Kaslow *et al.*, 2009; Waltman *et al.*, 2016). Therefore, investigation of self-assessments in CBT trainee cohorts during and after training is indicated.

### Self-assessment of competence in CBT trainees

Training courses play an important role in developing therapists' self-awareness of competence. Learning theory suggests that therapists' self-representations of their own skill change with

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press

increasing levels of feedback and experience (Lyon *et al.*, 2011). Self-representations are further suggested to become more accurate through comparison with expert evaluations (Kaslow *et al.*, 2009; Waltman *et al.*, 2016). Relevant studies, however, have not reached consensus regarding the relationship between CBT trainees' self- versus expert evaluations.

Comparison of trainees' self-ratings and expert ratings of competence for a single session during training (Rozek *et al.*, 2018) and towards the beginning and end of training (Mathieson *et al.*, 2009) indicated trends towards agreement, with varying significance. Limited longitudinal evidence indicates that while there is considerable variability in the strength of the relationship between self and expert ratings, agreement tends to improve across training (Loades and Myles, 2016; McManus *et al.*, 2012). A small study found evidence of trainee over-estimation compared with objective markers (Rozek *et al.*, 2018), while larger studies indicated under-estimation – particularly by more-competent trainees (McManus *et al.*, 2012), or switching over time (Loades and Myles, 2016). Methodological issues, including relatively small sample sizes ($n = 5$–$64$) and the use of different measures for self and expert ratings and/or of unvalidated competency rating scales may have contributed to inconsistency in findings. The relationship between self and expert ratings of CBT competence across training requires methodologically sound clarification.

### Self-assessment of competence beyond training

Self-awareness of competence during the transition from training into the post-training workplace has never been directly investigated, and is important given the proposed role of self-awareness in promoting evidence-based practice. IAPT CBT trainees appear to maintain expert-rated competence 12+ months after training (Liness *et al.*, 2018). If consensus between self- and expert-rated competence is attained by the end of training, this may persist into the workplace. Alternatively, previous reports of therapist drift (Waller, 2009) and over-estimation of competence in experienced therapists may indicate that agreement decreases over time. Given the inconsistent findings during training and absence of research into the post-training transition, longitudinal evaluation of the relationship between self-assessed and expert-assessed competence is warranted.

### Aims and hypotheses

This study aimed to clarify the relationship between self and expert assessments of CBT competence during training using a large longitudinal sample of IAPT trainees, and to directly assess agreement following transition into the workplace. Self and expert competence ratings for audiotaped sessions were compared at baseline, mid-training, end-of-training, and at 12+ month follow-up. Trainees who did not obtain competence by the end-of-training tape were provided with additional support targeting therapeutic skills and self-reflection, allowing for preliminary investigation into whether this group at risk of inaccurate self-assessments could achieve self-awareness of competence with further input. Hypotheses were:

(1) Self and expert competence ratings would demonstrate significant positive associations at all time points;
(2) The proportion of tapes on which self and expert competence ratings agree would increase significantly across training to reach good agreement by final submissions;
(3) Good agreement between self and expert competence ratings would be retained at follow-up.

Exploratory analyses were conducted to assess change in agreement for competent and non-competent trainees (those who attained versus those that did not attain competence by the end of training) across the course.

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press

## Method

### Participants

Participants were 150 therapists who completed a one-year postgraduate diploma in CBT between 2009 and 2014 and submitted a baseline therapy recording for training. Missing self or marker rating data were excluded for 12 of the 162 eligible trainees. Median age was 32.00 years (IQR = 7). The trainee sample was 77% female ($n = 116$) and 23% male ($n = 34$). Ethnicities were 83% white ($n = 125$), 8% Asian ($n = 12$), 6% black ($n = 9$) and 3% mixed ethnicity ($n = 4$). Professions were 33% Psychological Wellbeing Practitioner ($n = 50$), 33% clinical psychologist ($n = 49$), 12% counselling psychologist ($n = 18$), 12% counsellor/psychotherapist ($n = 17$), 5% mental health nurse ($n = 8$), 3% occupational therapist ($n = 5$), and 2% other ($n = 3$).

### Measures

Therapeutic competence was assessed with the UK Cognitive Therapy Scale – Revised (CTS-R; Blackburn *et al.*, 2001). The CTS-R consists of 12 items assessing general therapeutic and CBT-specific competencies, with a total score range of 0–72 and competence threshold ≥36. Internal consistency is high (α range = .75 to .97; Blackburn *et al.*, 2001). In the present sample, inter-rater reliability was excellent for expert markers: ICC (114, 113) = .95 (Liness *et al.*, 2019). Expert markers were course staff with extensive experience practising and supervising CBT and scoring the competence measure. They were trained in rating the CTS-R when they started the course via workshops and co-rating therapy tapes with other course staff. Subsequently, they were subject to regular accuracy monitoring via second marking of randomly selected tapes.

### Procedure

Data were collected during training for coursework submissions and were included in a larger observational study evaluating IAPT CBT training (Liness *et al.*, 2019). A one-day workshop in rating the CTS-R was provided to trainees prior to the baseline submission. Weekly supervision including live review of clips of trainees' therapy was also an integral part of training. Trainees submitted audio-recorded sessions with CTS-R self-ratings for the course at the beginning of training (baseline), after 5–6 months of training (mid-training), and after 9–10 months of training (end-of-training). Expert markers scored recordings on the CTS-R upon submission by students. All students receive detailed feedback for skill development on all submissions. Recordings consisted of an adult anxiety or depression mid-treatment session that was requested to be representative of trainees' usual practice. All trainees submitted recordings of both anxiety and depression treatment as competence was required across disorders. For further details about the training course, please refer to Liness *et al.* (2019).

Trainees were expected to obtain competence by the end of the course. Those who had not achieved competence by this time point ('non-competent') received additional support consisting of three clinical skills tutorials and practice-based feedback. These trainees then provided a resubmission tape and CTS-R rating.

Former trainees who were still working in trusts that had granted ethics approval for research were contacted at least 12 months post-training regarding participation in a larger follow-up study (Liness *et al.*, 2018). Self-rating data at follow-up were available for 30 therapists who were included in the present study. Therapists were requested to submit a recording and CTS-R self-rating of a mid-treatment CBT session representative of their usual practice, which was then expert-scored on the CTS-R. There were no significant differences between therapists who participated at follow-up and those who did not in self- or expert-assessed baseline CTS-R or self-assessed end-of-training CTS-R ($p < .05$). A significant difference emerged between mean expert-assessed end-of-training CTS-R for follow-up participants

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press

(mean = 39.47, $SD = 3.13$) and non-participants (mean = 37.37, $SD = 3.79$) [$t$ (148) = –2.81, $p = .006$], indicating possible self-selection bias for high-performing former trainees providing follow-up ratings and tapes.

## Results[1]

### Descriptive statistics for self and expert competence ratings

Table 1 reports trainees' mean self and expert CTS-R ratings at each time point. Mean self and expert ratings were both below the CTS-R competence threshold (≥36) at baseline and mid-training, crossed the competence threshold by the end of training, and remained above the threshold at follow-up.

### Association between self and expert ratings

#### Overall findings during training

Pearson's correlations revealed significant positive relationships between self and expert CTS-R ratings for trainees' baseline [$r$ (148) = .29, $p < .001$], mid-training [$r$ (148) = .50, $p < .001$] and final submissions [$r$ (148) = .27, $p < .001$]. To give a true picture of the relationship between self and expert ratings upon course completion, non-competent trainees' resubmission scores were entered as their final submissions.

#### Competent and non-competent trainees

Of 150 trainees, 113 obtained competence by end-of-training ('competent') and 37 did not ('non-competent'). Competent and non-competent trainees were not found to differ in distribution of professions (Fisher's exact test = 9.50, $p = .11$), therefore comparisons of self and expert ratings were conducted on the basis of competence only and not professional group.

Competent trainees' self-ratings positively correlated with expert ratings at baseline [$r$ (111) = .39, $p < .001$], mid- [$r$ (111) = .60, $p < .001$], and end-of-training [$r$ (111) = .27, $p = .004$]. Conversely, non-competent trainees' self-ratings did not correlate with corresponding expert ratings at baseline [$r$ (35) = .10, $p = .56$], mid- [$r$ (35) = .25, $p = .13$], or end-of-training [$r$ (35) = –.24, $p = .14$], but correlated positively for resubmission tapes [$r$ (35) = .40, $p = .005$].

#### Follow-up

Follow-up participants' self and expert ratings were positively associated [$r$ (28) = .56, $p < .001$]. Follow-up analyses were not conducted separately for competent and non-competent trainees, as only three previously non-competent trainees returned for follow-up.

### Significant agreement between self and expert ratings

Following Loades and Myles' (2016) criterion to assess CTS-R agreement between self and expert ratings, a difference of >5 points in either direction indicated significant trainee under- or over-estimation, while a difference within 5 points indicated significant agreement. This criterion was derived based on the psychometric properties of the CTS-R described by Gordon (2007). Figure 1 presents patterns of agreement for the whole sample at each time point.

[1]The Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995) was used to correct for multiple comparisons across all hypothesis tests, with false discovery rate $Q = .05$.

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press

**Table 1.** Self and marker CTS-R ratings across time points

| | *n* | Mean (*SD*) Self | Mean (*SD*) Marker |
|---|---|---|---|
| Baseline | 150 | 29.12 (6.77) | 29.27 (5.14) |
| Mid-training | 150 | 33.03 (4.57) | 34.69 (4.22) |
| End of training | 150 | 38.43 (2.84) | 37.79 (3.75) |
| Follow-up | 30 | 39.52 (4.71) | 39.95 (4.07) |

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

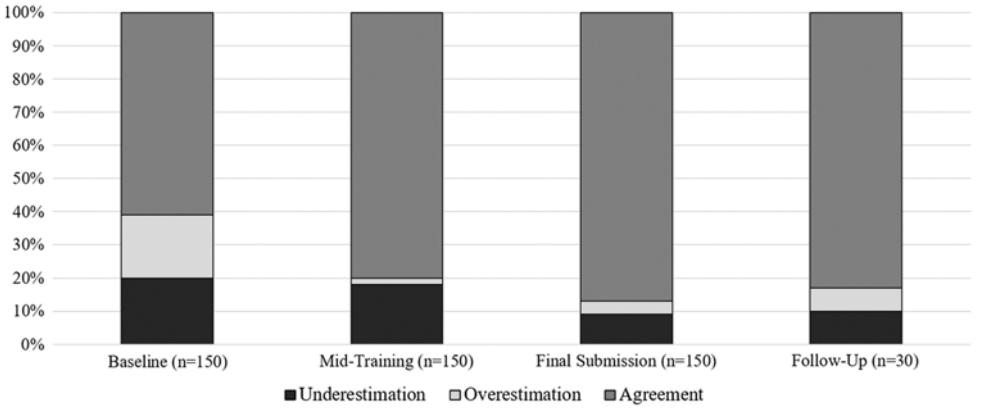



**Figure 1.** Patterns of agreement between self and expert CTS-R scores.

### *Overall findings during training*

As shown in Fig. 1, the majority of tapes at all time points demonstrated agreement between self and expert CTS-R scores. McNemar's tests (agreement vs disagreement) found that agreement increased significantly between baseline (61%) and mid-training (80%) [$\chi^2$ (1) = .84, $p < .001$], and slightly but not significantly between mid-training (80%) and final submissions (87%) [$\chi^2$ (1) = .36, $p = .15$]. Visual inspection of the patterns of disagreement indicate relatively even distribution between under- and over-estimation at baseline (20 vs 19%, respectively), marked tendency towards under-estimation mid-training (18 vs 2%), and some tendency towards under-estimation on final submissions (10 vs 4%).

### *Competent and non-competent trainees*

Figure 2 presents patterns of agreement across training for competent and non-competent trainees. The majority of submissions demonstrated agreement between self and expert CTS-R scores across all time points for both competent and non-competent trainees, although patterns of agreement varied considerably for these groups.

Agreement increased significantly for competent trainees between baseline (63%) and mid-training (80%) [$\chi^2$ (1) = 1.93, $p = .003$], and slightly but not significantly between mid-training (80%) and end-of-training (87%) [$\chi^2$ (1) = .003, $p = .28$]. At all time points, competent trainees demonstrated greater frequencies of under-estimation than over-estimation. Notably, there was no over-estimation at mid-training.

There was a non-significant increase from baseline (54%) to mid-training (78%) for non-competent trainees [$\chi^2$ (1) = .29, $p = .06$], and a non-significant decrease by end-of-training (60%) [$\chi^2$ (1) = .38, $p = .12$]. Agreement increased significantly by the resubmission tape (86%) [$\chi^2$ (1) = .001, $p = .02$]. Patterns of disagreement indicated a substantial tendency towards

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press

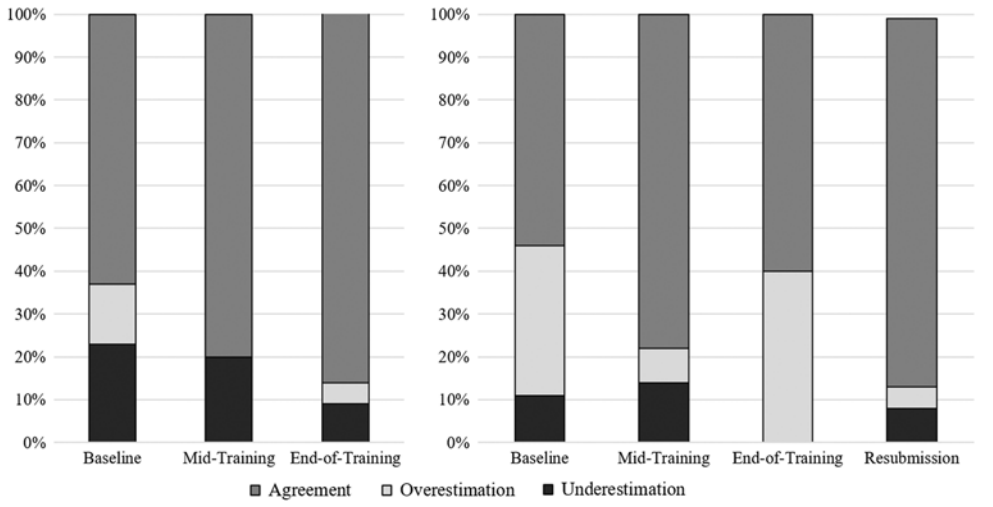



**Figure 2.** Patterns of agreement for competent and non-competent trainees.

over-estimation at baseline, followed by a small tendency towards under-estimation mid-training. At end-of-training, all disagreements indicated over-estimation; however, a small tendency towards under-estimation was found on the resubmission tape.

### *Follow-up*

A large majority (83%) of tapes demonstrated agreement between self and expert CTS-R scores at follow-up. McNemar's test found no significant change in agreement between final submissions and follow-up for the 30 follow-up participants [$\chi^2$ (1) = 1.20, $p = .10$]. The distribution of agreement was identical for final submissions and follow-up: 83% agreement, 10% under-estimation, and 7% over-estimation. These findings provide preliminary evidence that agreement between self and expert ratings is retained at follow-up.

## Discussion

This study aimed to clarify the relationship between self and expert assessments of CBT competence across training and following transition into the workplace. Recordings of CBT treatment sessions were scored by IAPT trainees and expert markers on the validated CTS-R (Blackburn *et al.*, 2001) at the beginning, middle and end of the course, and at post-training follow-up. As predicted, self and expert ratings demonstrated positive relationships and the proportion of tapes achieving agreement increased across training. Patterns of agreement between self and expert ratings emerged in the expected directions for trainees who were competent versus non-competent at the end of training, although there was unexpected variability across training for non-competent trainees. At post-training follow-up, former trainees' self- and expert-rated competence agreed as anticipated.

### *Agreement across training for the overall cohort*

Self and expert CTS-R scores were positively associated across the course, providing evidence of agreement during training. However, the strength of associations varied, which may have been in part due to competent and non-competent trainees having different profiles of agreement between self and expert ratings of their tapes (see discussion below). At all time points, the majority of tapes

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press

(61–87%) demonstrated agreement between self and expert ratings. The proportion of significant agreement (CTS-R difference <5; Loades and Myles, 2016) was lowest at baseline, which may be unsurprising as this is when trainees are learning both about the rating system and about CBT itself and thus there is greater opportunity to misunderstand both. Agreement increased significantly between baseline and mid-training, and non-significantly between mid-training and final submissions. These findings indicate that as trainees' CBT experience and use of rating scales increased, trainee self-ratings of competence began to converge more with expert raters who were highly experienced in both CBT and in the rating system. Improved agreement between self-rated and expert-rated competence is encouraging given the importance of accurate self-assessments for informing clinical practice and calibrating professional development in the post-training workplace.

These findings have important implications regarding the use of self-evaluation as a resource-efficient CBT competence assessment method – particularly in training. The changing proportion of agreement across training supports previous assertions that self-evaluation is not an appropriate stand-alone competence assessment for trainees (Rozek *et al.*, 2018). Rather, self-assessment appears to be more appropriate as a self-reflective learning tool (Waltman *et al.*, 2016), with trainees' understanding of their own skill becoming more accurate with repeated comparison to expert ratings and with increasing experience. The results also provide further support for the regular use of live therapy clips in supervision, as trainee self-report may not be sufficiently accurate to inform supervision during training.

Furthermore, these findings indicate that CTS-R rater training – which was delivered to trainees at the beginning of the course – is insufficient for the development of accurate self-assessment in the absence of further CBT training and experience. Whilst CTS-R rater training appears important for reliability (Gordon, 2007; Reichelt *et al.*, 2003), it may be that CBT experience itself is necessary for raters to understand and evaluate the underlying construct of how competent CBT is delivered. This has important potential implications beyond self-assessment, as novice raters or those with limited CBT training/experience may be unsuitable to conduct objective assessor ratings using the CTS-R. These findings converge with preliminary evidence of poorer reliability in novices' judgements of competency on other measures compared with experts' despite rater training (Schmidt *et al.*, 2018), and is an important topic for further enquiry.

Patterns of variability for the minority of tapes that disagreed – namely trainees' tendency to under-estimate rather than over-estimate their competence from the mid-training point – were consistent with those found in smaller longitudinal training studies (Loades and Myles, 2016; McManus *et al.*, 2012). Mid-training under-estimation of self-rated competence was consistent with Lyon *et al.*'s (2011) learning theory account of skill development. This suggests that skill and self-representations decrease as the implementation of new skills and standards interferes with skill perception, but subsequently increase with awareness of surpassing baseline competence after further training (Lyon *et al.*, 2011). This overall tendency towards under-estimation rather than trainee over-confidence for tapes that disagreed may be encouraging in light of evidence suggesting that professional self-doubt is linked with stronger clinical outcome (Meichenbaum and Lilienfeld, 2018; Nissen-Lie *et al.*, 2017). Although investigating the link between self-assessed competence on the CTS-R and clinical outcome was not within the scope of this study, this is a relevant area for future research given these findings from the literature.

### *Agreement across training for competent and non-competent trainees*

Competent trainees' ($n = 113$) self-ratings demonstrated positive relationships with expert ratings at all time points, with the proportion of agreement increasing across training. Additionally, tapes that disagreed tended to demonstrate under-estimation – consistent with McManus *et al.* (2012). Competent trainees may have used defensive pessimism to manage anxiety, a strategy often

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press

attributed to high performers (Norem and Cantor, 1986), and been more aware of areas requiring improvement (McManus *et al.*, 2012). This self-under-estimation in competent trainees may reflect professional self-doubt, a construct that has been linked with good clinical outcome in other therapeutic modalities (Meichenbaum and Lilienfeld, 2018; Nissen-Lie *et al.*, 2017).

Non-competent trainees ($n = 37$) demonstrated greater variability in agreement. Whilst non-competent trainees tended towards over-estimation at baseline and end-of-training, there was a drop in confidence mid-training that mirrored the pattern demonstrated by the overall cohort. Agreement only improved significantly by the resubmission tape, following additional support by course staff, and there was a significant positive relationship between self and expert ratings at this time point only. Additional support to develop competence may have promoted appropriate self-reflection about therapy skills; however, it was not feasible to evaluate whether this support was responsible for improved agreement beyond the effects of greater practice. Further investigation into effective extra support for non-competent trainees is recommended given this preliminary indication that it may be useful in fostering accurate self-awareness. In particular, future research should investigate whether non-competent trainees struggled on specific items of the CTS-R in order to inform targets for additional support.

### *Agreement at post-training follow-up*

This was the first study that directly assessed agreement between former trainees' self-ratings and expert ratings of competence following transition into the workplace. Self and expert ratings were positively associated, with the majority demonstrating agreement (83%). Furthermore, there was no significant change in the proportion of significant agreement between end-of-training and follow-up for therapists that participated. Former trainees' maintenance of accurate self-ratings is encouraging given the international scaling up of CBT training and provision, and given the relevance of self-appraisals to professional accreditation and therapy delivery.

The structure of IAPT CBT training and clinical services may have assisted maintenance of accurate self-ratings of competence. Trainees work 3 days in IAPT services during the course and the majority remain in IAPT services upon course completion (Liness *et al.*, 2019); therefore, strong ties are established between the course and services. While the amount of supervision decreases post-training, the majority of IAPT CBT therapists report regular individual supervision (Liness *et al.*, 2017). Continuity between training and clinical practice and ongoing supervision are important for therapists to maintain clinical skills (Brooker *et al.*, 2003; Mannix *et al.*, 2006; Rakovshik *et al.*, 2016) and may similarly assist therapists to retain accurate self-assessment of competency in CBT. However, results at follow-up may have been affected by self-selection bias; therapists who participated in the follow-up demonstrated greater expert-assessed end-of-training competence and relatively few previously non-competent trainees contributed tapes or self-assessment ratings. Similar difficulties with recruitment for follow-up studies are noted across the training and workplace literature (Brosan *et al.*, 2007; Miller *et al.*, 2004; Stirman *et al.*, 2015). Furthermore, the relatively brief 1-year follow-up period may have been insufficient for the development of therapist drift in experienced clinicians (Waller, 2009) and future research should investigate longer follow-up. Given that accurate self-appraisals of competence are proposed to play an important role in post-training CBT clinical practice (Fouad *et al.*, 2009) and may provide a resource-efficient adjunctive method to assess therapy quality (Shoenwald *et al.*, 2011), future studies with larger follow-up samples and longer follow-up periods are recommended to investigate factors that contribute to ongoing accuracy.

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press

## Limitations

This study has several limitations linked to its naturalistic, observational design. Sessions were self-selected by trainees and markers were not blinded to time point. It was not feasible to include a control group for training overall or for additional support. Future studies would benefit from a larger follow-up sample to investigate the relationship between self-assessed competence post-training and other practice-related factors and clinical outcomes. Also, while the CTS-R provided a validated and consistent method to compare self and expert competence ratings, estimates of its inter-rater reliability vary considerably across the literature (Blackburn *et al.*, 2001; Gordon, 2007; James *et al.*, 2001; Kazantzis *et al.*, 2018; Reichelt *et al.*, 2003). Trained raters, however, appear to achieve stronger reliability (Gordon, 2007; Reichelt *et al.*, 2003) and excellent reliability was obtained by the trained and experienced raters in the present sample (Liness *et al.*, 2019). The CTS-R also provides a specific rating of competence for a single session at a given time point, and the addition of global self and supervisor ratings of therapist competence across a full caseload may also be informative in future studies.

## Conclusions

Training courses are responsible for developing an increasingly large workforce of therapists able to effectively deliver interventions and monitor their own abilities and professional needs. The present findings suggest that self-awareness of competence can be developed during training – even for struggling trainees. Agreement appears to be maintained at post-training follow-up. These findings add to evidence (Liness *et al.*, 2018) that structured training with strong links to clinical services, as provided by the IAPT training and service delivery model, is beneficial for clinicians' long-term development and outcomes.

**Acknowledgements.** With many thanks to all the therapists and patients who participated in this study, and the IAPT services who granted access to information. We also thank research assistants Hannah Parker and Steffen Nessler for their help with data collection and management.

**Financial support.** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Dr 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 are those of the authors and not necessarily those of NHS, the National Institute for Health Research (NIHR), or the Department of Health.

**Conflicts of interest.** Sheena Liness runs the IoPPN/KCL IAPT CBT training course that is the subject of this study. Sarah Beale and Colette Hirsch declare that they have no conflicts of interest with respect to this publication.

**Ethics statements.** All contributors abided by the Ethical Principles of Psychologists and Code of Conduct as set out by the APA. This study was approved by the King's College London Psychiatry, Nursing, and Midwifery Research Ethics Committee: reference number PNM/12/13-50. Approval on behalf of NHS services was granted by the NHS West Midlands-South Birmingham Research Ethics Committee: reference number 13/WM/0475.

**Key practice points**

(1) Accurate self-awareness of competence can be developed across long-term CBT training.
(2) Self-evaluations are not recommended as a formal assessment method for CBT trainees. They may, however, be a helpful learning tool.
(3) Additional support for less-competent trainees may help to develop accurate self-awareness of competence in this group.
(4) Good agreement between self and expert judgements of competence can be maintained in the post-training workplace. Continuity between training and clinical services provided by the IAPT programme may have been helpful.

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press

## Further reading

**Loades, M. E., & Myles, P. J.** (2016). Does a therapist's reflective ability predict the accuracy of their self-evaluation of competence in cognitive behavioural therapy? *The Cognitive Behaviour Therapist*, *9*.

**Waltman, S. H., Frankel, S. A., & Williston, M. A.** (2016). Improving clinician self-awareness and increasing accurate representation of clinical competencies. *Practice Innovations*, *1*, 178.

## References

**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)*, 289–300.

**British Association for Behavioural and Cognitive Psychotherapies (BABCP)** (2012). *BABCP Minimum Training Standards for the Practice of Cognitive Behavioural Therapy*. Available at: https://www.babcp.com/files/Accreditation/General/Minimum-Training-Standards-V7-0215.pdf

**Blackburn, I. M., James, I. A., Milne, D. L., Baker, C., Standart, S., Garland, A., & Reichelt, F. K.** (2001). The Revised Cognitive Therapy Scale (CTS-R): psychometric properties. *Behavioural and Cognitive Psychotherapy*, *29*, 431–446.

**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*, 731–747.

**Brown, J. D.** (1986). Evaluations of self and others: self-enhancement biases in social judgments. *Social Cognition*, *4*, 353–376.

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

**Brosan, L., Reynolds, S., & Moore, R. G.** (2008). Self-evaluation of cognitive therapy performance: do therapists know how competent they are? *Behavioural and Cognitive Psychotherapy*, *36*, 581–587.

**Chevron, E. S., & Rounsaville, B. J.** (1983). Evaluating the clinical skills of psychotherapists: a comparison of techniques. *Archives of General Psychiatry*, *40*, 1129–1132.

**Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J.** (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*, 83–87.

**Fouad, N. A., Grus, C. L., Hatcher, R. L., Kaslow, N. J., Hutchings, P. S., Madson, M. B., Collins, F.L., & Crossman, R. E.** (2009). Competency benchmarks: a model for understanding and measuring competence in professional psychology across training levels. *Training and Education in Professional Psychology*, *3*(4S), S5.

**Gordon, P. K.** (2007). A comparison of two versions of the Cognitive Therapy Scale. *Behavioural and Cognitive Psychotherapy*, *35*, 343–353.

**Hool, N.** (2010). BABCP Core Curriculum Reference Document. British Association for Behavioural and Cognitive Psychotherapies. Available at: https://www.babcp.com/documents

**James, I. A., Blackburn, I. M., Milne, D. L., & Reichelt, F. K.** (2001). Moderators of trainee therapists' competence in cognitive therapy. *British Journal of Clinical Psychology*, *40*, 131–141.

**Kaslow, N. J., Grus, C. L., Campbell, L. F., Fouad, N. A., Hatcher, R. L., & Rodolfa, E. R.** (2009). Competency assessment toolkit for professional psychology. *Training and Education in Professional Psychology*, *3*(4S), S27.

**Kazantzis, N., Clayton, X., Cronin, T. J., Farchione, D., Limburg, K., & Dobson, K. S.** (2018). The Cognitive Therapy Scale and Cognitive Therapy Scale-Revised as measures of therapist competence in cognitive behavior therapy for depression: relations with short and long term outcome. *Cognitive Therapy and Research*, *42*, 385–397.

**Liness, S., Beale, S., Lea, S., Byrne, S., Hirsch, C. R., & Clark, D. M.** (2018). The sustained effects of CBT training on therapist competence and patient outcomes. *Cognitive Therapy and Research*, *43*, 631–641.

**Liness, S., Beale, S., Lea, S., Byrne, S., Hirsch, C. R., & Clark, D. M.** (2019). Multi-professional IAPT CBT training: clinical competence and patient outcomes. *Behavioural and Cognitive Psychotherapy*, *47*, 1–14.

**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*, 16–30.

**Loades, M. E., & Myles, P. J.** (2016). Does a therapist's reflective ability predict the accuracy of their self-evaluation of competence in cognitive behavioural therapy? *The Cognitive Behaviour Therapist*, *9*.

**Lyon, A. R., Stirman, S. W., Kerns, S. E., & Bruns, E. J.** (2011). Developing the mental health workforce: review and application of training approaches from multiple disciplines. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*, 238–253.

**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 behavior therapy techniques for palliative care practitioners. *Palliative Medicine*, *20*, 579–584.

**Mathieson, F. M., Barnfield, T., & Beaumont, G.** (2009). Are we as good as we think we are? Self-assessment versus other forms of assessment of competence in psychotherapy. *The Cognitive Behaviour Therapist*, *2*, 43–50.

**McLeod, B. D., Southam-Gerow, M. A., Tully, C. B., Rodriguez, A., & Smith, M. M.** (2013). Making a case for treatment integrity as a psychosocial treatment quality indicator for youth mental health care. *Clinical Psychology: Science and Practice*, *20*(1), 14–32.

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press

**McManus, F., Rakovshik, S., Kennerley, H., Fennell, M., & Westbrook, D.** (2012). An investigation of the accuracy of therapists' self-assessment of cognitive-behaviour therapy skills. *British Journal of Clinical Psychology*, *51*, 292–306.

**Meichenbaum, D., & Lilienfeld, S. O.** (2018). How to spot hype in the field of psychotherapy: a 19-item checklist. *Professional Psychology: Research and Practice*, *49*, 22.

**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*, 1050–1062.

**NHS England** (2016). The Five Year Forward View for Mental Health. London: NHS England. Available at: https://www.england.nhs.uk/wp-content/uploads/2016/02/Mental-Health-Taskforce-FYFV-final.pdf

**Nissen-Lie, H. A., Rønnestad, M. H., Høglend, P. A., Havik, O. E., Solbakken, O. A., Stiles, T. C., & Monsen, J. T.** (2017). Love yourself as a person, doubt yourself as a therapist? *Clinical Psychology & Psychotherapy*, *24*, 48–60.

**Norem, J. K., & Cantor, N.** (1986). Defensive pessimism: harnessing anxiety as motivation. *Journal of Personality and Social Psychology*, *51*, 1208.

**Parker, Z. J., & Waller, G.** (2015). Factors related to psychotherapists' self-assessment when treating anxiety and other disorders. *Behaviour Research and Therapy*, *66*, 1–7.

**Perlesz, A. J., Stolk, Y., & Firestone, A. F.** (1990). Patterns of learning in family therapy training. *Family Process*, *29*, 29–44.

**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 supervision. *Journal of Consulting and Clinical Psychology*, *84*, 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*, 87–99.

**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*, 244–250.

**Schoenwald, S. K., Garland, A. F., Chapman, J. E., Frazier, S. L., Sheidow, A. J., & Southam-Gerow, M. A.** (2011). Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*, 32–43.

**Schmidt, I. D., Strunk, D. R., DeRubeis, R. J., Conklin, L. R., & Braun, J. D.** (2018). Revisiting how we assess therapist competence in cognitive therapy. *Cognitive Therapy and Research*, *42*, 369–384.

**Singla, D. R., Weobong, B., Nadkarni, A., Chowdhary, N., Shinde, S., Anand, A., Fairburn, C. G., Dimijdane, S., Velleman, R., Weiss, H., & Patel, V.** (2014). Improving the scalability of psychological treatments in developing countries: an evaluation of peer-led therapy quality assessment in Goa, India. *Behaviour Research and Therapy*, *60*, 53–59.

**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*, 115.

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

**Waller, G.** (2009). Evidence-based treatment and therapist drift. *Behaviour Research and Therapy*, *47*, 119–127.

**Waltman, S. H., Frankel, S. A., & Williston, M. A.** (2016). Improving clinician self-awareness and increasing accurate representation of clinical competencies. *Practice Innovations*, *1*, 178.

**Cite this article:** Beale S, Liness S, and Hirsch CR. Trainee self-assessment of cognitive behaviour therapy competence during and after training. *The Cognitive Behaviour Therapist*. https://doi.org/10.1017/S1754470X19000357

https://doi.org/10.1017/S1754470X19000357 Published online by Cambridge University Press