

MAIN

Development of competence in cognitive behavioural therapy and the role of metacognition among clinical psychology and psychotherapy students

Hillevi Bergvall^{1*} , Ata Ghaderi² , Joakim Andersson², Tobias Lundgren¹ ,
Gerhard Andersson^{1,3,4}  and Benjamin Bohman¹ 

¹Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, and Stockholm Healthcare Services, Region Stockholm, Stockholm, Sweden, ²Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden, ³Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden and ⁴Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden
*Corresponding author. Email: hillevi.bergvall@ki.se

(Received 16 July 2022; revised 25 October 2022; accepted 7 December 2022; first published online 24 January 2023)

Abstract

Background: There is a paucity of research on therapist competence development following extensive training in cognitive behavioural therapy (CBT). In addition, metacognitive ability (the knowledge and regulation of one's cognitive processes) has been associated with learning in various domains but its role in learning CBT is unknown.

Aims: To investigate to what extent psychology and psychotherapy students acquired competence in CBT following extensive training, and the role of metacognition.

Method: CBT competence and metacognitive activity were assessed in 73 psychology and psychotherapy students before and after 1.5 years of CBT training, using role-plays with a standardised patient.

Results: Using linear mixed modelling, we found large improvements of CBT competence from pre- to post-assessment. At post-assessment, 72% performed above the competence threshold (36 points on the Cognitive Therapy Scale-Revised). Higher competence was correlated with lower accuracy in self-assessment, a measure of metacognitive ability. The more competent therapists tended to under-estimate their performance, while less competent therapists made more accurate self-assessments. Metacognitive activity did not predict CBT competence development. Participant characteristics (e.g. age, clinical experience) did not moderate competence development.

Conclusions: Competence improved over time and most students performed over the threshold post-assessment. The more competent therapists tended to under-rate their competence. In contrast to what has been found in other learning domains, metacognitive ability was not associated with competence development in our study. Hence, metacognition and competence may be unrelated in CBT or perhaps other methods are required to measure metacognition.

Keywords: cognitive behaviour therapy; education; professional competence; metacognition

Introduction

While cognitive behavioural therapy (CBT) has substantial research support, training of CBT therapists has received limited attention in research (Becker and Stirman, 2011; Fairburn and Cooper, 2011; Rakovshik and McManus, 2010; Shafran *et al.*, 2009), especially regarding the effects of extensive training programmes (i.e. programmes in higher education, spanning at

© The Author(s), 2023. Published by Cambridge University Press on behalf of British Association for Behavioural and Cognitive Psychotherapies. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

least a year and including theory, practice and supervision). For example, when mental health care professionals received CBT training for a year (two days per week) within the British Improving Access to Psychological Therapies programme (Clark, 2018) the training was associated to CBT competence improvement among the 278 trainees (Cohen's $d = 1.25$; McManus *et al.*, 2010). In addition, a majority (80–82%) scored above the competence threshold after training (i.e. a total score on the Cognitive Therapy Scale-Revised of ≥ 36 points; Liness *et al.*, 2019). Far more studies have focused on less extensive training programmes (≤ 15 days) in different formats, showing improved CBT competence according to a systematic review of 76 studies (Frank *et al.*, 2020).

These studies mainly targeted practising therapists and clinicians, while the more extensive training programmes tend to target students without any prior clinical experiences. Considering the resources allocated to train future psychologists and psychotherapists, there is a paucity of studies evaluating to what extent more extensive training programmes promote CBT competence development in students. Results from an effectiveness study suggested that psychology students in extensive CBT training, achieved an adequate ability to deliver CBT, as patient outcomes were on a par with those of experienced, licensed therapists (Öst *et al.*, 2012). To our knowledge, no studies have examined the effects of extensive CBT training on competence among students without any prior CBT experience.

Universities have a long tradition of assuring student knowledge based on written tests, essays and assignments (Gonsalvez and Crowe, 2014). These methods are cost-effective, easily administered, and have documented that training improves CBT knowledge (Cooper *et al.*, 2017a; Fairburn *et al.*, 2017; Harned *et al.*, 2011; Harned *et al.*, 2013). However, knowledge does not necessarily transfer to clinical skills (Fairburn and Cooper, 2011; Rakovshik and McManus, 2010). To bridge the gap between examining knowledge and implementation, written exams with hypothetical cases have been used in CBT training (Cooper *et al.*, 2017b) and situational judgement tests in medical training (Patterson *et al.*, 2016). Relatedly, studies show that CBT trainees associated competence development with more experiential, practical learning strategies (i.e. supervision, role-play, trainer interaction, reflective practice), rather than theoretical strategies (i.e. tests, essays, literature, lectures; Bennett-Levy *et al.*, 2009; Rakovshik and McManus, 2013).

Psychotherapeutic competence assessments often rely on supervisor observations, which are prone to biases due to the nature of the supervisory relationship and setting (Gonsalvez and Crowe, 2014). As an alternative, Fairburn and Cooper (2011) recommended role-plays with standardised patients, that is, an actor playing a patient with a specific set of presenting problems. Role-play with standardised patients is already well-established for assessing clinical competence in medical training, often in the form of objective structured clinical examinations, with good evidence of their reliability and validity in psychiatry according to a review (Hodges *et al.*, 2014). Although resource consuming, role-plays seem to be a feasible method to assess clinical competence in psychology students (Goodie *et al.*, 2021) as well as CBT trainees (Edwards *et al.*, 2016). With proper training of standardised patients, script adherence and character fidelity was achieved and maintained over time (Edwards *et al.*, 2016). A few studies have used role-plays and found significant improvements in competence following CBT training (Cooper *et al.*, 2017b; Harned *et al.*, 2013; Kobak *et al.*, 2017; Puspitasari *et al.*, 2017). Also, role-play performance has been reported to resemble clinical performance by therapists (Cooper *et al.*, 2017b), indicating ecological validity.

Flavell (1979) coined the term *metacognition* (MC), a higher order thinking which simply put is 'cognition about cognition' or 'thinking about thinking'. Metacognition is commonly classified into MC knowledge and MC regulation (Flavell, 1979; Pintrich, 2002; Schraw, 1998; Schraw and Moshman, 1995). MC knowledge refers to knowledge about cognitive processes, how to carry out a cognitive task, and the effectiveness of strategies. MC regulation refers to the active control of one's thinking or learning. Various regulatory skills have been described in the literature, of which the most common classification consists of planning, monitoring and

evaluation (Schraw, 1998; Schraw and Moshman, 1995). CBT therapists continuously plan, monitor and evaluate cognitive processes as well as their own performance, for example in self-reflection (Bennett-Levy, 2006). Unfortunately, self-assessments have been prone to biases. Kruger and Dunning (1999) suggested that incompetence is not only related to poor skills, but an inability to self-assess those skills, resulting in overconfidence among poor performers. The Dunning-Kruger effect has been observed in various areas (e.g. Ehrlinger *et al.*, 2008). Previous findings are inconclusive regarding therapists' ability to self-evaluate accurately. Brosan *et al.* (2008) found that cognitive therapists, especially those less competent, over-rated their competence compared with observers, in line with the Dunning-Kruger effect. In contrast, McManus *et al.* (2012) found that more competent therapists under-rated their competence compared with supervisors, a tendency referred to as undue modesty (Dunning *et al.*, 2003).

Metacognition is associated with learning and performance in various domains (e.g. mathematics, science and reading; Ohtani and Hisasaka, 2018; Perry *et al.*, 2019). Training in psychological treatments may be improved by targeting therapists' MC skills, also suggested by Fauth *et al.* (2007). Yet, to our knowledge, no studies have explored the role of metacognition in learning psychological treatments, or for that matter, treatment of any kind.

The purpose of the present study was to investigate to what extent clinical psychology and psychotherapy students acquire competence in CBT following an extensive training programme, and what role metacognition plays in the development of CBT skills.

Method

Design and setting

Using a longitudinal observational design, participants were assessed for CBT competence and MC activity before receiving CBT training and after their third semester of CBT training ($M = 1.39$ years between assessments). The CBT training corresponds to a total of 38.5–52.5 European credits, provided within master-level programmes in psychology or psychotherapy. The psychology programme in Sweden spans five years full-time (300 credits) where CBT training is introduced after 3.5 years. The psychotherapy programme spans three years half-time (90 credits), where all students have a previous master-level degree in a health care profession, at least two years of clinical experience, and are required to simultaneously work clinically (i.e. providing psychotherapy), at least half-time. The CBT training of both programmes consists of lectures, tests, a thesis, workshops, and clinical practice under close supervision. Two universities took part in the study, Karolinska Institutet (KI) and Stockholm University (SU) in Stockholm, Sweden.

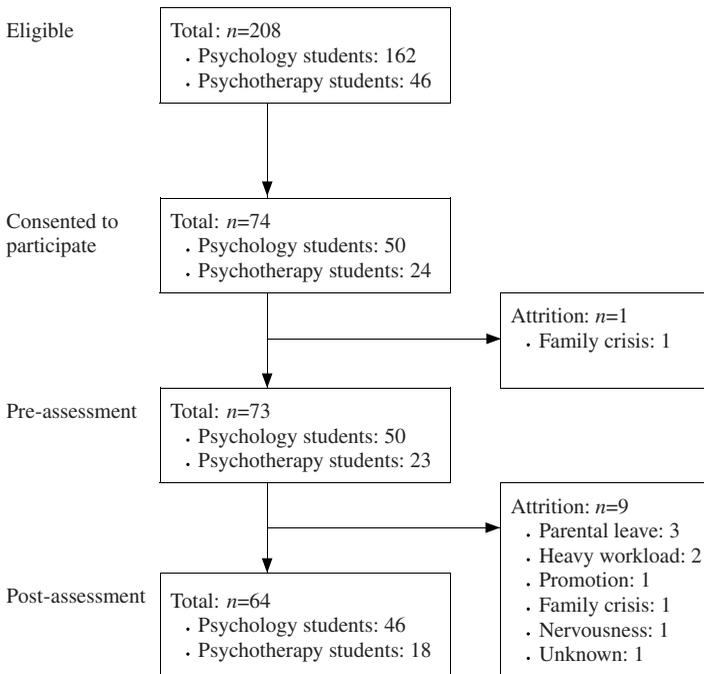
Students received verbal and written information and written informed consent was obtained. Participation was voluntary and could be retracted at any time without the need to state a reason. All data were pseudonymised and stored securely, to which only the first and last authors had access. The study was conducted in accordance with the World Medical Association Declaration of Helsinki.

Participants

Participants were students at the clinical psychology ($n = 50$) and psychotherapy ($n = 23$) programmes at KI and SU, starting their CBT training from August 2016 to September 2017. Among 208 eligible students, 74 enrolled and provided written informed consent. No exclusion criteria were applied. One participant dropped out due to a family crisis, before providing any data and was therefore not included in the analysis. Among the 73 participants who completed pre-assessment, 64 completed post-assessment. See Fig. 1 for participant flow

Table 1. Participant characteristics

| | Psychology students | | Psychotherapy students | | Total | |
|-----------------------------|---------------------|-------|------------------------|-------|---------------|-------|
| | <i>M (SD)</i> | Range | <i>M (SD)</i> | Range | <i>M (SD)</i> | Range |
| Age (years) | 29.6 (6.2) | 22–52 | 40.5 (7.8) | 27–55 | 33.0 (8.4) | 22–55 |
| Clinical experience (years) | 0 | 0 | 5.4 (3.5) | 0–14 | 1.7 (3.2) | 0–14 |
| CBT sessions weekly pre | 0 | 0 | 11.0 (7.7) | 0–25 | 3.5 (6.7) | 0–25 |
| CBT sessions weekly post | 1.5 (2.9) | 0–20 | 11.7 (4.0) | 4–20 | 4.4 (5.6) | 0–20 |

**Figure 1.** Participant flow through the study.

through the study, including reasons for attrition. The mean age of participants was 33 years pre-assessment, 68% were female (33 of 50 psychology students and 17 of 23 psychotherapy students) and 76% studied at KI (35 psychology students and 20 psychotherapy students). See Table 1 for participant characteristics.

Assessment and procedure

At pre- and post-assessment, participants engaged in a CBT role-play session, immediately followed by an MC think-aloud session. The sessions included a standardised patient, that is, an actor playing a patient with a specific set of presenting problems, in this case, indicative of a principal diagnosis of social anxiety disorder with co-morbid depression. The patient was a 32-year-old female called 'Anna', with mild social anxiety (e.g. being shy and self-critical), which had escalated during her parental leave from a high-achieving position. She was now on long-term sick leave, socially isolated, and in a strained marriage. Actors were psychology students in their first or second year, who had not met the participants previously. All assessments were video recorded. However, two MC recordings were lost, and another two were partially lost, due to technical malfunction of recording during the pre-assessment.

Role-play of a cognitive behavioural therapy session

Before the role-play, participants were provided with written background information on the patient and the treatment up to this point, including the home assignment for the session. The participant then conducted a complete 45-minute session of CBT, designed to constitute the ninth treatment session. Standardised patients had previously received three hours of training and the first author monitored adherence during their first role-plays. The actors followed a detailed written instruction, including scripted responses to various potential therapist behaviours. To ensure stability of actors' behaviours, they were to re-read the manuscript before each role-play and regularly monitored throughout the study.

Competence in cognitive behavioural therapy

CBT competence was assessed using the Cognitive Therapy Scale-Revised (CTS-R; James *et al.*, 2001), which is a standard tool for the assessment of CBT competence (Muse and McManus, 2013). It measures skills general to psychotherapy (e.g. feedback, collaboration and pacing) and specific to CBT (e.g. eliciting key cognitions, guided discovery, and homework setting). The CTS-R contains 12 items rated on a scale from 0 to 6 points which correspond to six levels of competence (incompetent, novice, advanced beginner, competent, knowledgeable, and expert). The total score ranges from 0 to 72 points. Commonly, a score of 36 points is used as a cut-off for competence (James *et al.*, 2001), giving an average score of 3 points per item. CTS-R has demonstrated excellent internal consistency (Cronbach's $\alpha = .92-.97$) and adequate inter-rater reliability (intra-class correlation coefficient [ICC] = .86 for pairs of raters; Blackburn *et al.*, 2001).

As part of the present study and prior to the pre-assessment, two CBT experts participated in a two-day workshop on the use of the CTS-R, followed by repeated calibration sessions during the two-year rating period to minimise rater drift. Regular inter-rater reliability checks, six times during the two-year rating period, showed an adequate to excellent inter-rater reliability throughout the study (ICC = .64-.95). One rater was a psychologist, the other a nurse, and both were licensed psychotherapists and CBT supervisors, with extensive experience in CBT training and clinical practice. Raters were independent of the study, and blinded to participants, training programmes, and assessment points.

Metacognitive self-assessment of performance

MC ability to monitor and self-evaluate one's performance was assessed using a participant survey administered after the role-play. Participants rated their CBT competence in the role-play (i.e. their perceived performance as CBT therapists) on a visual analogue scale. These self-assessment scores were made to range from 0 to 72 points, allowing for comparison with observer-ratings of CBT competence based on CTS-R total scores.

Metacognitive task

Metacognition was assessed using a think-aloud methodology, by which the participant is asked to say whatever comes to mind, that is, think aloud, while performing a task. The expressed chain of thoughts is then transcribed and coded as a source of information on MC activity. The standard procedure was employed (Ericsson and Simon, 1993). An independent research assistant gave the participant instructions on how to think aloud, followed by a brief exercise to make sure the participant understood. The MC task included an enactment of six common clinical situations, designed to subject the participant to a clinical challenge and presumably mobilise MC activity (e.g. implied suicide-risk, therapy doubt, questioning of in-session exercises). The procedure was as follows: a research assistant handed over a written instruction, such as 'Find out how Anna is doing'. The participant acted accordingly, for example, by asking 'How are

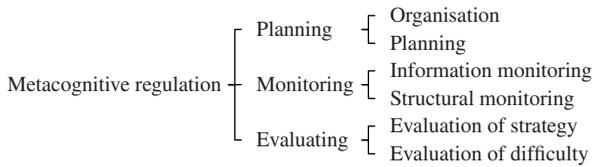


Figure 2. Metacognitive taxonomy.

you today, Anna?’ to which the patient (‘Anna’) gave a standardised reply, covering various information; ‘Not well. The kids have been ill. I just want to disappear. What if therapy doesn’t work?’. The patient then turned silent, allowing the research assistant, if necessary, to prompt the participant to ‘Please think aloud’. The participant would then say whatever came to mind, reflecting on what the patient had just said. The verbal reports of the MC task were recorded, transcribed verbatim, and coded according to an MC taxonomy.

Metacognitive taxonomy

Because we could not identify any measures of metacognition within a psychotherapeutic or other clinical setting, we created a taxonomy for MC regulation within CBT, inspired by a validated taxonomy for coding of MC in various non-clinical learning domains (e.g. science, history) by Meijer *et al.* (2006). As suggested by Flavell (1979) and Meijer *et al.* (2006), we included three types of MC regulation, that is, planning, monitoring and evaluating, each including several categories. A pilot version of the role-play was tested with three CBT therapists, resulting in minor adjustments to ensure both role-play and taxonomy usability. The final version of the MC taxonomy included six categories to be coded, with no hierarchy of importance or advancement between them. See Fig. 2 for an overview of the taxonomy.

In Fig. 2, *Organisation* refers to the handling of available information, e.g. to summarise or interpret information, such as ‘Anna might have suicidal thoughts’. *Planning* entails reflecting on possible future activities and choosing between strategies, e.g. ‘First, a suicide risk assessment’. *Information monitoring* means checking whether one has information needed in the situation, e.g. ‘I’m confused, I don’t know if...’. *Structural monitoring* means checking the therapeutic framing regarding time or content, e.g. ‘We’re getting off track’. *Evaluation of strategy* refers to assessing the results of one’s activities, such as ‘She did not like my suggestion, I should have...’. *Evaluation of difficulty* entails the assessment of difficulty, such as ‘This is a tough case’.

A coding manual was created, with CBT relevant examples of each category. Consistent with Meijer *et al.* (2006), all verbal reports were coded as manifestations of underlying MC regulatory activity. The reports were divided into units, which were coded. A unit was defined to start when a thought was introduced and end when this thought was fully expressed, or when a new thought was introduced. Coding was conducted by the first author who was blinded to participants, training programmes and assessment points. Results indicated a strong (McHugh, 2012) intra-rater reliability (Cohen’s kappa = .84, $p < .001$) when 20% ($n = 27$) of the reports were re-coded three weeks later.

Statistical analysis

The SPSS (version 26, SPSS Inc., Chicago, IL, USA) was used for the analyses. Frequencies of MC regulation categories were calculated. Linear mixed models (LMM) were used to estimate the effect of time on CBT competence and investigate whether effects were moderated by group (psychology or psychotherapy students) and metacognition (frequency of MC categories). Repeated measures of the outcome (i.e. CTS-R total scores at pre- and post-assessments) were nested within individuals. We chose LMM as it is recommended for nested data with repeated measures and for handling missing data appropriately (e.g. Gueorguieva and Krystal, 2004).

The maximum likelihood method was used to estimate model parameters. We started with a basic model including a fixed intercept. Then we successively added random parameters (intercept and slope), and finally a time by condition interaction term to the model. Each model's fit to the observed data was evaluated using the likelihood ratio test, with significance set at .05. A model with a significantly better fit than a previous model was retained. The standardised effect size for between- and within-group effects at post-assessment was calculated as Cohen's *d* for LMM based on the formula recommended by Feingold (2015; Equation 1), using the pre-assessment pooled standard deviation for the entire sample, and for each subsample, respectively. For model-based *d*, 95% confidence intervals (CI) were calculated using the formulas provided in Feingold (2015; Equations 7 and 8).

Accuracy of self-assessed CBT competence as a measure of MC ability was calculated as the difference between self-assessed competence relative to observer-assessed competence using the CTS-R. A Pearson correlation test was used to examine if CBT competence (i.e. CTS-R total scores) was associated with self-assessment accuracy (i.e. self-assessed CBT competence relative to CTS-R total scores). In addition, to compare self-assessment accuracy in groups with different levels of CBT competence (i.e. bottom, second, third and top quartiles of CTS-R total scores), we used the non-parametric Kruskal Wallis test, because the assumption of normality was not met for the quartiles.

Results

Development of competence in cognitive behavioural therapy

Competence in CBT at pre- and post-assessment is presented in Table 2. Overall, CBT competence (i.e. CTS-R total score ≥ 36 points) was achieved by 12 participants (18.8%) at pre-assessment and by 46 participants (71.9%) at post-assessment. None of the initially competent participants deteriorated below the cut-off, which means that 34 participants (53.1%) improved above the competence threshold. Among the psychology students, CBT competence was achieved by two participants (4.3%) at pre-assessment and 31 participants (67.4%) at post-assessment. Among psychotherapy students, CBT competence was achieved by 10 participants (55.6%) at pre-assessment and 15 participants (83.3%) at post-assessment.

A model including a random intercept, fixed effect of time, fixed effect of group (psychology or psychotherapy students), and a time by group interaction term provided the best fit. There was a statistically significant main effect of time, $F_{1,63.98} = 95.64$, $p < .001$, $d = 1.94$, 95% CI [1.64, 2.25], indicating large improvements in CBT competence for the whole sample from pre- to post-assessment. There was also a significant main effect of group, $F_{1,122.38} = 60.63$, $p < .001$, $d = 1.76$, 95% CI [1.31, 2.21], indicating large differences in competence between psychology and psychotherapy students with the latter group of students being more competent at both pre- and post-assessment. There was a significant time by group interaction, $F_{1,63.98} = 16.27$, $p < .001$, $d = -1.13$, 95% CI [-1.70, -0.57], suggesting a larger CBT competence improvement in psychology students, who initially had lower competence scores and therefore more room for change. Based on models separated by groups, the main effect of time on CBT competence for psychology students was $F_{1,42.37} = 162.51$, $p < .001$, $d = 1.99$, 95% CI [1.68, 2.31] and for psychotherapy students $F_{1,19.59} = 11.91$, $p = .003$, $d = 0.77$, 95% CI [0.30, 1.23]. Separate models including different participant characteristics did not improve model fit. Specifically, CBT competence development was not predicted by participant age, university, clinical experience, or weekly CBT sessions.

Accuracy in self-assessment of CBT competence

We assessed the relationship between observer-assessed CBT competence (i.e. CTS-R total scores) and accuracy of self-assessed CBT competence (i.e. a difference score). A Pearson correlation test

Table 2. Means of observer-assessed CBT competence, self-assessed CBT competence and metacognitive activity

| | Pre-assessment | | Post-assessment | |
|------------------------------|------------------------|-----------|------------------------|-----------|
| | <i>M</i> (<i>SD</i>) | Range | <i>M</i> (<i>SD</i>) | Range |
| Total students | | | | |
| CBT competence (CTS-R) | 27.99 (10.24) | 8.0–49.5 | 40.41 (6.47) | 25.5–56.5 |
| CBT competence self-assessed | 24.53 (12.36) | 3.6–55.4 | 33.51 (10.52) | 5.0–52.6 |
| Metacognitive activity | 56.79 (22.09) | 23–130 | 57.81 (21.14) | 22–117 |
| Orientation | 17.24 (9.12) | 4–46 | 16.33 (7.43) | 3–47 |
| Planning | 22.40 (9.35) | 6–50 | 23.38 (9.10) | 10–46 |
| Information monitoring | 7.21 (5.28) | 0–22 | 7.30 (4.97) | 0–18 |
| Structural monitoring | 4.37 (2.76) | 0–12 | 6.11 (3.69) | 0–17 |
| Evaluation of strategies | 4.10 (2.37) | 0–10 | 3.20 (2.40) | 0–10 |
| Evaluation of difficulty | 1.40 (1.60) | 0–7 | 1.50 (1.84) | 0–10 |
| Psychology students | | | | |
| CBT competence (CTS-R) | 23.59 (7.72) | 8.0–39.0 | 39.24 (6.51) | 25.5–56.5 |
| CBT competence self-assessed | 21.15 (12.21) | 3.6–55.4 | 33.26 (10.96) | 5.0–52.6 |
| Metacognitive activity | 58.14 (24.91) | 23–130 | 62.13 (21.62) | 30–117 |
| Orientation | 17.53 (9.65) | 4–46 | 17.63 (7.79) | 3–47 |
| Planning | 22.49 (10.37) | 6–50 | 24.74 (9.55) | 10–46 |
| Information monitoring | 7.73 (5.68) | 0–22 | 8.28 (5.14) | 0–18 |
| Structural monitoring | 4.35 (2.91) | 0–12 | 6.46 (3.91) | 1–17 |
| Evaluation of strategies | 4.18 (2.46) | 1–10 | 3.30 (2.37) | 0–10 |
| Evaluation of difficulty | 1.78 (1.72) | 0–7 | 1.72 (2.02) | 0–10 |
| Psychotherapy students | | | | |
| CBT competence (CTS-R) | 37.57 (8.39) | 23.0–49.5 | 43.42 (5.42) | 34.0–53.0 |
| CBT competence self-assessed | 31.87 (9.25) | 15.8–49.0 | 34.16 (9.57) | 15.8–49.0 |
| Metacognitive activity | 53.62 (13.37) | 29–77 | 46.78 (15.45) | 22–88 |
| Orientation | 16.57 (7.90) | 4–36 | 13.00 (5.25) | 4–24 |
| Planning | 22.19 (6.59) | 8–32 | 19.89 (6.90) | 11–30 |
| Information monitoring | 6.00 (4.06) | 1–20 | 4.78 (3.47) | 1–13 |
| Structural monitoring | 4.43 (2.46) | 0–9 | 5.22 (2.96) | 1–10 |
| Evaluation of strategies | 3.90 (2.19) | 0–10 | 2.94 (2.53) | 0–9 |
| Evaluation of difficulty | 0.52 (0.75) | 0–2 | 0.94 (1.16) | 0–4 |

CTS-R, Cognitive Therapy Scale-Revised.

showed a negative correlation between the variables at both pre-assessment, $r_{71} = -.47$, $p < .001$, and post-assessment, $r_{62} = -0.50$, $p < .001$. This correlation was also significant among psychology and psychotherapy students respectively, r between $-.60$ and $-.49$, $p < .05$ across assessment points. Thus, higher CBT competence was correlated to lower accuracy in self-assessment.

To compare quartiles of CBT competence, a Kruskal Wallis test was conducted. There was a significant difference in self-assessment accuracy between quartiles at pre-assessment, $H_3 = 15.80$, $p = 0.001$, and post-assessment, $H_3 = 12.86$, $p = 0.005$. *Post hoc* pairwise comparisons showed significant differences between Q1–Q4, Q2–Q4 and Q1–Q3 at both time points, and at pre-assessment Q2–Q3 ($p = .001-.047$). On average, the 25% most competent students underestimated their competence with 11.22 points ($SD = 10.10$) at pre-assessment and 13.43 points ($SD = 9.67$) at post-assessment; see Fig. 3 for observer- and self-assessed CBT competence. Meanwhile the 25% least competent were rather accurate in their self-assessments and only slightly overrated their competence at pre-assessment, $M = 2.41$, $SD = 11.70$, and at post-assessment, $M = 0.81$, $SD = 13.11$.

The role of metacognition in CBT competence development

MC activity, both total and its categories, in participants at pre- and post-assessment is presented in Table 2. To investigate the role of MC in CBT competence development, we added MC total

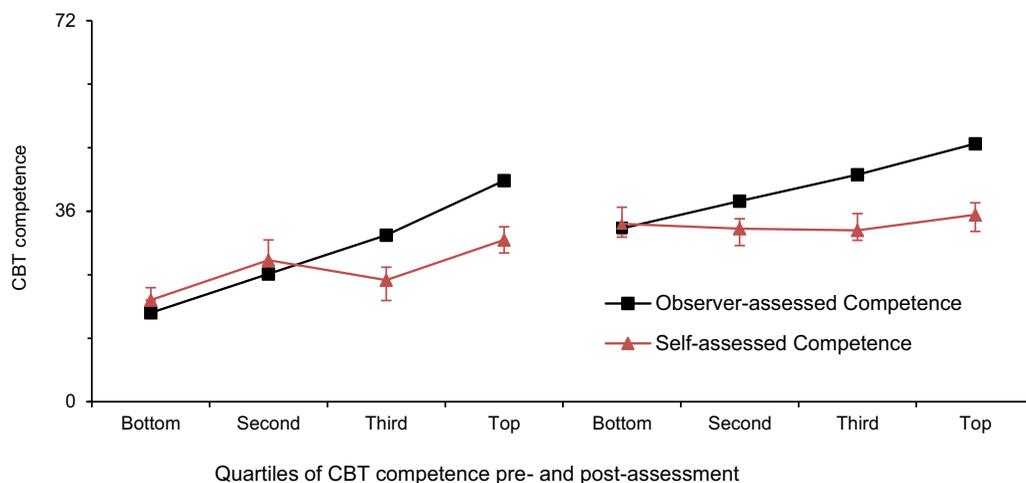


Figure 3. Observer- and self-assessed CBT competence at pre- and post-assessment. Boxes and triangles are means, and vertical lines are standard error bars. Participants are divided into quartiles of observer-rated CBT competence (i.e. CTS-R total scores) at pre- and post-training.

score to the previous model (i.e. an LMM with a random intercept, fixed effects of time, group, MC total score, and interaction terms of time by group, time by MC total score, and time by group by MC total score). Model fit significantly improved, but neither main nor interaction effects for MC total score were statistically significant.

For exploratory purposes, we conducted the same analysis separately for each of the six MC categories (i.e. replaced MC total score with an MC category score). Again, each MC variable significantly improved model fit, but neither main nor interaction effects were significant, except for a main effect of evaluation of difficulty, $F_{1,133.15} = 6.57$, $p = .011$, and an interaction effect of evaluation of difficulty by group, $F_{1,133.15} = 6.17$, $p = .014$. However, MC evaluation of difficulty did not predict competence development over time. Thus, MC activity did not predict CBT competence development in the participants.

Discussion

The purpose of the present study was to investigate to what extent psychology and psychotherapy students acquire competence in CBT following an extensive training programme and the role of metacognition for competence development.

Most students had achieved CBT competence post-assessment. CBT competence was achieved by 83.3% of the psychotherapy students at post-assessment, which is on par with previous findings where 80–82 % of practising therapists were found competent after extensive CBT training (Liness *et al.*, 2019). The finding that three psychotherapy students were below the threshold can be considered a cause of concern; however, it should be noted that they only needed an additional 0.5 to 2 points to reach the competence level and still have 1.5 years of CBT training as part of their programme.

CBT competence was achieved by 67.4% of the psychology students. Students without clinical experience are targeted in most extensive training programmes, yet we have found no studies evaluating their competence development. No psychology students had any previous experience of CBT or clinical practice, most treated only 1–3 patients during training, and the programme is followed by a year of supervised clinical practice. Thus, there is room for

improvement, but the results are encouraging, and it is likely that more students will reach the competence threshold later.

Overall, we found CBT competence had improved for psychology and psychotherapy students at post-assessment. We also found a significant time by group interaction, suggesting larger improvement for psychology students, who had lower initial competence ratings and thus, more room for change. As can be expected, psychotherapy students were older, with more clinical experience and weekly CBT practice. However, therapist variables such as age, previous clinical experience, or weekly CBT practice, did not moderate competence development, unlike results in some previous studies (e.g. McManus *et al.*, 2010).

Explorative analyses showed that more competent therapists were less accurate in their self-assessment of CBT competence. There were significant differences in self-assessment accuracy between groups of high and low CBT competence at both pre- and post-assessment, where the top 25% under-rated their competence, while the bottom 25% were accurate or slightly over-rated their competence. Our results provide support for the Dunning-Kruger effect regarding undue modesty among top performers, which has been observed in other areas (Dunning *et al.*, 2003), but not regarding the less competent. So far, the findings concerning CBT have been inconclusive; for example, Brosan *et al.* (2008) found that especially less competent therapists over-rated their competence, whereas McManus *et al.* (2012) found more competent trainees to under-estimate their competence.

While MC ability has been related to learning in other fields (e.g. Ohtani and Hisasaka, 2018; Perry *et al.*, 2019), it was not related to competence development in our study. If metacognition plays a role in CBT competence development, we were not able to detect it. Psychotherapy may differ from other fields of learning, in involving social interaction and being less straightforward, without one clinically 'correct' response for each situation. Moreover, students of CBT are already trained in (self-)reflection and higher order thinking, which is why they may already possess a high degree of MC ability, in contrast to the often younger, more novice participants in other studies of metacognition and learning. Thus, students of CBT may differ in the quantity, as well as the quality of MC activity. Another possible explanation is that we developed an MC taxonomy for CBT, based on previous research in other areas, which has yet to be validated.

Limitations of the study

First, although it is likely that increased competence was the result of training, as the present study did not have an experimental design (e.g. a randomised controlled design), no inferences about causality can be drawn. Thus, improvements from pre- to post-assessment can be due to other factors than training, such as the passage of time or raters' expectations of improvement. However, expectations were controlled for as raters were blinded to participants, programme and assessment points. While an experimental design could demonstrate causality, it was not feasible in the context of the present study.

Second, one may question the validity and reliability of an artificial simulation to assess CBT competence (Muse and McManus, 2013) and competence has been suggested to vary over sessions (Webb *et al.*, 2010). Therefore, we made efforts to ensure that the role-play resembled a session with an actual patient and that actor's performance was stable across participants and time points, thorough training of standardised patients, pilot testing, and providing background information. Standardised patients can be designed to promote a range of skills in a single session, be replicated pre- and post-assessment, and have other practical advantages over real patient sessions (Muse and McManus, 2013). Moreover, role-plays with standardised patients seem to be a feasible, valid and reliable method to assess clinical competence (Edwards *et al.*, 2016; Goodie *et al.*, 2021; Hodges *et al.*, 2014). One concern has been the perceived authenticity of the standardised patients (Edwards *et al.*, 2016; Hodges *et al.*, 2014), which we did not monitor and may impact external validity. Yet, in another study therapists reported their role-play performance

to resemble their clinical performance (Cooper *et al.*, 2017b) and standardised role-plays have been used to measure competence development following CBT training successfully (Cooper *et al.*, 2017b; Harned *et al.*, 2013; Kobak *et al.*, 2017; Puspitasari *et al.*, 2017).

Third, CTS-R is considered the gold standard in assessing CBT competence, but expert ratings are resource-consuming and inter-rater reliability remains an issue (Muse and McManus, 2013), which in the present study was adequate to excellent ($ICC = .64-.95$). The competence threshold has not been validated for the CTS-R, only for the original version of the scale (Muse and McManus, 2013), hence any conclusions on participants passing the threshold should be interpreted with caution.

Fourth, we used think-aloud methodology which relies on the participant's ability to verbalise their MC activities. This seems to have a minimal impact on the participant's MC activity (Ericsson and Simon, 1993); however, a recent review reported that some studies observed that some prompts may have a positive impact (Double and Birney, 2019), which we tried to avoid by keeping prompts general (i.e. 'please think aloud') and to a minimum. Instead, retrospective self-reports could be used, but are less valid and poorly associated to observations of performance on MC tasks (Craig *et al.*, 2020), why observational methods, such as a think-aloud protocol, are recommended (Veenman *et al.*, 2006).

Finally, we wanted to investigate metacognition as originally defined by Flavell (1979), and were inspired by a taxonomy validated in other fields (Meijer *et al.*, 2006). Perhaps certain MC abilities are more relevant to CBT competence, than this broader conceptual framework. Furthermore, our participants frequently engaged in organisation and planning, but less in monitoring and evaluating categories (see Table 2). While this may reflect a true tendency, it is plausible the MC task failed to mobilise certain MC skills, or the participants found these particularly difficult to verbalise. Perhaps a more specified taxonomy is needed, targeting more clinically relevant MC activities.

Conclusions

Properly trained and qualified health care professionals are essential in the dissemination of evidence-based psychological treatments. Our study found that CBT competence improved and that most students had achieved competence after three terms of CBT training. Thus, competence improved among clinicians as well as novice students, without prior clinical experience and with limited CBT training practice. However, some students had not achieved CBT competence at post-assessment, which demonstrates the need for their upcoming supervised clinical practice and subsequent assessments. Routine assessments with standardised instruments may be integrated into educational programmes, and results used to improve training, e.g. by targeting specific skill areas.

Higher CBT competence was correlated to lower accuracy in self-assessment, where the more competent therapists under-estimated their competence. These results, along with large within-group variation, indicate that self-assessments of CBT competence are not reliable and prone to bias, as previously suggested.

To our knowledge, this is the first study to investigate MC in the context of learning CBT, and, indeed, treatment of any kind. We did not find that MC activity predicted CBT competence, and hence it may not need additional attention in the learning of CBT. However, we think it is plausible that we were just unable to detect it and suggest that our taxonomy may need further revision, to sufficiently differentiate the higher quality of MC ability among psychology and psychotherapy students.

Data availability statement. The data that support the findings of this study are available on request from the last author, B.B. The data are not publicly available due to ethical/privacy restrictions.

Acknowledgements. We would like to thank the participants and the students portraying the standardised patient.

Author contributions. Hillevi Bergvall: Data curation (equal), Formal analysis (lead), Funding acquisition (supporting), Investigation (equal), Methodology (supporting), Project administration (equal), Writing – original draft (lead), Writing – review & editing (equal); Ata Ghaderi: Conceptualization (supporting), Methodology (supporting), Writing – review & editing (supporting); Joakim Andersson: Methodology (supporting), Writing – review & editing (supporting); Tobias Lundgren: Supervision (supporting), Writing – review & editing (supporting); Gerhard Andersson: Supervision (supporting), Writing – review & editing (supporting); Benjamin Bohman: Conceptualization (lead), Data curation (equal), Formal analysis (supporting), Funding acquisition (lead), Investigation (equal), Methodology (lead), Project administration (equal), Supervision (lead), Writing – original draft (supporting), Writing – review & editing (equal).

Financial support. This work was supported by the Karolinska Institutet (FoUI-952486).

Conflict of interest. The authors declare none.

Ethical standards. The Regional Ethical Review Board in Stockholm assessed the study and decided it was not subjected to the Swedish Ethical Review Act, presumably because it did not involve collection of 'sensitive' data (e.g. health, ethnicity). However, the Board stated that it 'had no ethical objections to the study' (2016/1108-31).

References

- Becker, K. D., & Stirman, S. W. (2011). The science of training in evidence-based treatments in the context of implementation programs: current status and prospects for the future. *Administration and Policy in Mental Health*, 38, 217–222. <https://doi.org/10.1007/s10488-011-0361-0>
- Bennett-Levy, J. (2006). Therapist skills: a cognitive model of their acquisition and refinement. *Behavioural and Cognitive Psychotherapy*, 34, 57–78. <https://doi.org/10.1017/S1352465805002420>
- Bennett-Levy, J., McManus, F., Westling, B. E., & Fennell, M. (2009). Acquiring and refining CBT skills and competencies: which training methods are perceived to be most effective? *Behavioural and Cognitive Psychotherapy*, 37, 571–583. <https://doi.org/10.1017/s1352465809990270>
- 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. <https://doi.org/10.1017/S1352465801004040>
- 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. <https://doi.org/10.1017/S1352465808004438>
- Clark, D. M. (2018). Realizing the mass public benefit of evidence-based psychological therapies: the IAPT Program. *Annual Review of Clinical Psychology*, 14, 159–183. <https://doi.org/10.1146/annurev-clinpsy-050817-084833>
- Cooper, Z., Bailey-Straebl, S., Morgan, K. E., O'Connor, M. E., Caddy, C., Hamadi, L., & Fairburn, C. G. (2017a). Using the internet to train therapists: randomized comparison of two scalable methods. *Journal of Medical Internet Research*, 19, e355. <https://doi.org/10.2196/jmir.8336>
- Cooper, Z., Doll, H., Bailey-Straebl, S., Bohn, K., de Vries, D., Murphy, R., O'Connor, M. E., & Fairburn, C. G. (2017b). Assessing therapist competence: development of a performance-based measure and its comparison with a web-based measure. *JMIR Mental Health*, 4, e51. <https://doi.org/10.2196/mental.7704>
- Craig, K., Hale, D., Grainger, C., & Stewart, M. E. (2020). Evaluating metacognitive self-reports: systematic reviews of the value of self-report in metacognitive research. *Metacognition and Learning*, 15, 155–213. <https://doi.org/10.1007/s11409-020-09222-y>
- Double, K. S., & Birney, D. P. (2019). Reactivity to measures of metacognition. *Frontiers in Psychology*, 10, 2755. <https://doi.org/10.3389/fpsyg.2019.02755>
- 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. <https://doi.org/10.1111/1467-8721.01235>
- Edwards, K. S., Parish, S. J., Rosen, R. C., Garvert, D. W., Spangler, S. L., & Ruzek, J. I. (2016). A standardized patient methodology to assess cognitive-behavioral therapy (CBT) skills performance: development and testing in a randomized controlled trial of web-based training. *Training and Education in Professional Psychology*, 10, 149–156. <https://doi.org/10.1037/tep0000119>
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, 105, 98–121. <https://doi.org/https://doi.org/10.1016/j.obhdp.2007.05.002>
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. The MIT Press. <https://doi.org/10.7551/mitpress/5657.001.0001>
- Fairburn, C. G., Allen, E., Bailey-Straebl, S., O'Connor, M. E., & Cooper, Z. (2017). Scaling up psychological treatments: a countrywide test of the online training of therapists. *Journal of Medical Internet Research*, 19, e214. <https://doi.org/10.2196/jmir.7864>

- Fairburn, C. G., & Cooper, Z. (2011). Therapist competence, therapy quality, and therapist training. *Behaviour Research and Therapy*, 49, 373–378. <https://doi.org/10.1016/j.brat.2011.03.005>
- Fauth, J., Gates, S., Vinca, M. A., Boles, S., & Hayes, J. A. (2007). Big ideas for psychotherapy training. *Psychotherapy*, 44, 384–391. <https://doi.org/10.1037/0033-3204.44.4.384>
- Feingold, A. (2015). Confidence interval estimation for standardized effect sizes in multilevel and latent growth modeling. *Journal of Consulting and Clinical Psychology*, 83, 157–168. <https://doi.org/10.1037/a0037721>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Frank, H. E., Becker-Haimes, E. M., & Kendall, P. C. (2020). Therapist training in evidence-based interventions for mental health: a systematic review of training approaches and outcomes. *Clinical Psychology*, 27, e12330. <https://doi.org/10.1111/cpsp.12330>
- Gonsalvez, C. J., & Crowe, T. P. (2014). Evaluation of psychology practitioner competence in clinical supervision. *American Journal of Psychotherapy*, 68, 177–193. <https://doi.org/10.1176/appi.psychotherapy.2014.68.2.177>
- Goodie, J. L., Bennion, L. D., Schvey, N. A., Riggs, D. S., Montgomery, M., & Dorsey, R. M. (2021). Development and implementation of an objective structured clinical examination for evaluating clinical psychology graduate students. *Training and Education in Professional Psychology*. <https://doi.org/10.1037/tep0000356>
- Georgieva, R., & Krystal, J. H. (2004). Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Archives of General Psychiatry*, 61, 310–317. <https://doi.org/10.1001/archpsyc.61.3.310>
- Harned, M. S., Dimeff, L. A., Woodcock, E. A., & Contreras, I. (2013). Predicting adoption of exposure therapy in a randomized controlled dissemination trial. *Journal of Anxiety Disorders*, 27, 754–762. <https://doi.org/10.1016/j.janxdis.2013.02.006>
- Harned, M. S., Dimeff, L. A., Woodcock, E. A., & Skutch, J. M. (2011). Overcoming barriers to disseminating exposure therapies for anxiety disorders: a pilot randomized controlled trial of training methods. *Journal of Anxiety Disorders*, 25, 155–163. <https://doi.org/10.1016/j.janxdis.2010.08.015>
- Hodges, B. D., Hollenberg, E., McNaughton, N., Hanson, M. D., & Regehr, G. (2014). The Psychiatry OSCE: a 20-year retrospective. *Academic Psychiatry*, 38, 26–34. <https://doi.org/10.1007/s40596-013-0012-8>
- James, I. A., Blackburn, I.-M., & Reichelt, F. K. (2001). *Manual of the Revised Cognitive Therapy Scale (CTS-R)*. University of Newcastle.
- 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, 252–265. <https://doi.org/10.1007/s10608-016-9819-4>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77, 1121–1134. <https://doi.org/doi.proxy.kib.ki.se/10.1037/0022-3514.77.6.1121>
- Liness, S., Beale, S., Lea, S., Byrne, S., Hirsch, C. R., & Clark, D. M. (2019). Multi-professional IAPT CBT training: clinical competence and patient outcomes. *Behavioural and Cognitive Psychotherapy*, 47, 672–685. <https://doi.org/10.1017/s1352465819000201>
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22, 276–282. <https://doi.org/10.11613/BM.2012.031>
- McManus, F., Rakovshik, S. G., 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. <https://doi.org/10.1111/j.2044-8260.2011.02028.x>
- McManus, F., Westbrook, D., Vazquez-Montes, M., Fennell, M., & Kennerley, H. (2010). An evaluation of the effectiveness of diploma-level training in cognitive behaviour therapy. *Behaviour Research and Therapy*, 48, 1123–1132. <https://doi.org/10.1016/j.brat.2010.08.002>
- Meijer, J., Veenman, M. V. J., & van Hout-Wolters, B. H. A. M. (2006). Metacognitive activities in text-studying and problem-solving: development of a taxonomy. *Educational Research and Evaluation*, 12, 209–237. <https://doi.org/10.1080/13803610500479991>
- Muse, K., & McManus, F. (2013). A systematic review of methods for assessing competence in cognitive-behavioural therapy. *Clinical Psychology Review*, 33, 484–499. <https://doi.org/10.1016/j.cpr.2013.01.010>
- Ohtani, K., & Hisasaka, T. (2018). Beyond intelligence: a meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning*, 13, 179–212. <https://doi.org/10.1007/s11409-018-9183-8>
- 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, 160–173. <https://doi.org/10.1016/j.beth.2011.05.001>
- Patterson, F., Zibarras, L., & Ashworth, V. (2016). Situational judgement tests in medical education and training: research, theory and practice: AMEE Guide No. 100. *Medical Teacher*, 38, 3–17. <https://doi.org/10.3109/0142159X.2015.1072619>

- Perry, J., Lundie, D., & Golder, G. (2019). Metacognition in schools: what does the literature suggest about the effectiveness of teaching metacognition in schools? *Educational Review*, 71, 483–500. <https://doi.org/10.1080/00131911.2018.1441127>
- Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory Into Practice*, 41, 219–225. https://doi.org/10.1207/s15430421tip4104_3
- Puspitasari, A. J., Kanter, J. W., Busch, A. M., Leonard, R., Dunsiger, S., Cahill, S., Martell, C., & Koerner, K. (2017). A randomized controlled trial of an online, modular, active learning training program for behavioral activation for depression. *Journal of Consulting and Clinical Psychology*, 85, 814–825. <https://doi.org/10.1037/ccp0000223>
- Rakovshik, S. G., & McManus, F. (2010). Establishing evidence-based training in cognitive behavioral therapy: a review of current empirical findings and theoretical guidance. *Clinical Psychology Review*, 30, 496–516. <https://doi.org/10.1016/j.cpr.2010.03.004>
- Rakovshik, S. G., & McManus, F. (2013). An anatomy of CBT training: trainees' endorsements of elements, sources and modalities of learning during a postgraduate CBT training course. *the Cognitive Behaviour Therapist*, 6. <https://doi.org/10.1017/S1754470X13000160>
- Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, 26, 113–125. <https://doi.org/10.1023/A:1003044231033>
- Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7, 351–371. <https://doi.org/1040-726X/95/1200-0351507.50/0>
- Shafraan, R., Clark, D. M., Fairburn, C. G., Arntz, A., Barlow, D. H., Ehlers, A., Freeston, M., Garety, P. A., Hollon, S. D., Ost, L. G., Salkovskis, P. M., Williams, J. M., & Wilson, G. T. (2009). Mind the gap: improving the dissemination of CBT. *Behaviour Research and Therapy*, 47, 902–909. <https://doi.org/10.1016/j.brat.2009.07.003>
- Veenman, M. V. J., van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: conceptual and methodological considerations. *Metacognition Learning*, 1, 3–14. <https://doi.org/10.1007/s11409-006-6893-0>
- 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, 200–211. <https://doi.org/10.1037/a0018912>

Cite this article: Bergvall H, Ghaderi A, Andersson J, Lundgren T, Andersson G, and Bohman B (2023). Development of competence in cognitive behavioural therapy and the role of metacognition among clinical psychology and psychotherapy students. *Behavioural and Cognitive Psychotherapy* 51, 200–213. <https://doi.org/10.1017/S1352465822000686>