

ORIGINAL RESEARCH

Practitioner competence and clinical outcomes during cognitive behavioural and cognitive-analytic guided self-help for anxiety

Niall Power^{1,2} , Melanie Simmonds-Buckley^{1,3}, Jaime Delgado^{1,3}, Jessica Smithies⁴, Charlotte Bee⁴ and Stephen Kellett^{1,3}

¹Clinical and Applied Psychology Unit, Department of Psychology, University of Sheffield, Sheffield, UK, ²South West Yorkshire Partnership NHS Foundation Trust, UK, ³Rotherham Doncaster and South Humber NHS Foundation Trust, UK and ⁴Pennine Care NHS Foundation Trust, UK

Corresponding author: Niall Power; Email: niall.power@swyt.nhs.uk

(Received 10 March 2024; revised 30 July 2024; accepted 28 September 2024)

Abstract

Despite the ubiquity of guided self-help (GSH) interventions in Primary Care psychological services, there have been no previous studies of the relationship between the competence of qualified practitioners and treatment outcomes. This study compared competence-outcome associations in two types of GSH. Competence and clinical outcome measures were drawn from a clinical trial comparing the efficacy of two types of GSH for anxiety disorders, based on cognitive behaviour therapy (CBT-GSH) or cognitive analytic therapy (CAT-GSH). These interventions were delivered over the telephone by qualified and supervised practitioners. Audio-recordings of GSH sessions ($n=94$) were rated using a validated competence measure. Clinical outcomes were anxiety, depression and functioning. Secondary outcomes were attendance and need for further intervention after GSH. Competence ratings were highly reliable. No significant associations were found between competence and clinical outcomes, treatment engagement or need for further intervention. In this clinical trial, GSH competence ratings were not associated with clinical outcomes. Directions for future competence-outcome research are provided for GSH interventions.

Key learning aims

- (1) Become familiar with the current empirical literature on therapist competence and associations with clinical outcome.
- (2) Raise awareness of a recently developed alternative form of guided self-help based on the theory and principles of cognitive analytic therapy.
- (3) Enhance understanding of the relationship between practitioner competence and clinical outcome in guided self-help for anxiety disorders.

Keywords: adherence; competence; guided self-help; process-outcome

Introduction

Guided self-help (GSH) interventions have emerged in response to the pressing need to improve access to psychological support, given the high prevalence of common mental health problems around the world (Singla *et al.*, 2017). GSH interventions have great potential to be delivered at a larger scale and at lower cost, shortening waiting times for support due to their brevity (Ho *et al.*, 2016). GSH is also known to be more acceptable to treatment-seeking people compared with

self-directed bibliotherapy or computerised self-help, as the availability of a human facilitator enhances engagement with these interventions (Gellatly *et al.*, 2007). Shafran *et al.* (2021) defined GSH as interventions that utilise psychoeducational materials, over 6 hours or less of practitioner contact time, sessions being typically <30 minutes, and being facilitated by qualified practitioners rather than psychotherapists. Meta-analyses of randomised controlled trials show that GSH is efficacious when treating depression and anxiety disorders, both when compared with waitlist controls (Gellatly *et al.*, 2007; Ho *et al.*, 2020) and evidence-based psychotherapies (Cuijpers *et al.*, 2010).

Commonly available GSH interventions are based on the theory and principles of cognitive behavioural therapy (CBT), but translated into a psychoeducational format (Baguley *et al.*, 2010). Evidence from meta-analyses of clinical trials (e.g. Gellatly *et al.*, 2007) and practice-based studies (Wakefield *et al.*, 2021) indicates that CBT-based GSH is clinically effective for the alleviation of depression and anxiety symptoms. However, there is also evidence that this model of treatment has low acceptability to some patients (Chan and Adams, 2014) and high relapse rates (Ali *et al.*, 2017). Hence, there is an empirical and theoretical justification to develop and test the acceptability and efficacy of GSH interventions based on other theoretical models. One example of an alternative modality of GSH is based on principles of cognitive analytic therapy (CAT). CAT is an integrative psychotherapy that is supported by meta-analytic evidence of acceptability (Simmonds-Buckley *et al.*, 2022) and effectiveness (Hallam *et al.*, 2021). CAT-GSH has fidelity to GSH principles (Meadows and Kellett, 2017), and uses a three-phase format (i.e. reformulation, recognition, revision) to change the relational patterns – developmentally anchored in childhood – which are conceptualised to be maintaining factors for current anxiety symptoms. CAT-GSH is therefore distinctive from CBT-GSH due to its past-present focus (Meadows and Kellett, 2017), whilst being highly acceptable to practitioners (Wray *et al.*, 2022) and equally as effective as CBT-GSH (Kellett *et al.*, 2023).

Despite the reliance on psychoeducational materials, there is evidence of considerable variability in treatment outcomes achieved by different practitioners when delivering individual (Firth *et al.*, 2015) and group-based GSH (Delgado *et al.*, 2016). A possible explanation is that more effective practitioners are more clinically competent and more adherent to the psychoeducational approach. Sharpless and Barber (2009) defined ‘adherence’ as the extent to which a therapist delivers theory-specific techniques/interventions, whereas ‘competence’ refers to the skilfulness with which interventions are implemented. Fidelity/integrity are terms used interchangeably to refer to a composite measure of therapist adherence, competence and treatment differentiation (Perepletchikova *et al.*, 2007). In a recent systematic review of 62 studies, Power *et al.* (2022) conducted a meta-analysis to examine associations between clinical outcomes and measures of therapist adherence, competence and integrity/fidelity. There was evidence of a small but statistically significant correlation between higher competence and better outcomes, and a small-to-moderate and statistically significant correlation between higher integrity and better outcomes. Current evidence supports the notion that competence and integrity are important aspects of effective clinical practice, although this literature is primarily drawn from traditional psychotherapy studies and so the generalisability to GSH is unclear. No studies have examined competence-outcome associations during GSH delivered by qualified practitioners and the findings using trainee samples are mixed and contradictory (e.g. see Branson *et al.*, 2018; Delgado *et al.*, 2020).

The current study therefore aimed to contribute to the evidence base for GSH by investigating competence-outcome associations in a sample of patients who received one of two types of GSH for an identified anxiety disorder and were treated by qualified practitioners. The study examined the relationship between competence and rate of sessional clinical change, attendance, drop-out and the need for further interventions. The following hypotheses were tested: (1) Higher PWP competence would be significantly associated with improved patient outcomes (anxiety, depression and functioning), although there would be no differences between CBT-GSH and

CAT-GSH groups. (2) Higher PWP competence would be significantly associated with lower patient drop-out and increased treatment session attendance across the overall sample, although there would be no differences between CBT-GSH and CAT-GSH groups. (3) PWP competence level would be significantly associated with step-up rates within and across treatment conditions although there would be no differences between CBT-GSH and CAT-GSH groups (exploratory). (4) PWP competence would predict patient rate of change in sessional anxiety outcomes across the overall sample (exploratory).

Method

Ethics and design

This study was a secondary analysis of data from a pragmatic and partially randomised patient preference trial conducted in a Primary Care psychological therapy service in England (Kellett *et al.*, 2023). The trial was ethically approved (IRAS ref: 240751), registered in a database for clinical trials (NCT03730532), and compared the efficacy of CBT-GSH and CAT-GSH. The consent form and ethical approval process for the clinical trial encompassed the analyses reported in this follow-up study. As part of the trial, participants either chose (based on standardised treatment description information) or were randomised when they had no preference to either CBT-GSH or CAT-GSH. Anxiety disorders were identified using the Mini International Neuropsychiatric Interview (MINI; Sheehan *et al.*, 1997). Each participant had one GSH session randomly selected for audio recording and rating. The primary outcome in the trial was the Beck Anxiety Inventory (BAI; Beck *et al.*, 1988) at 8- and 24-week follow-up. Secondary outcomes included sessional outcomes on the Generalised Anxiety Disorder-7 (GAD-7; Spitzer *et al.*, 2006), Patient Health Questionnaire-9 (PHQ-9; Kroenke *et al.*, 2001) and Work and Social Adjustment Scale (WSAS; Mundt *et al.*, 2002). These sessional measures enabled a granular examination of treatment response (e.g. trajectories of change, and precise estimation of post-treatment outcomes at the end of each patient's therapy) to be taken in the current study. Only secondary outcomes of the clinical trial were analysed in the current study.

Inclusion and exclusion criteria

In order to be eligible for inclusion, participants needed to: (1) be referred by a general medical practitioner or to self-refer for psychological treatment; (2) 18+ years old; (3) meet criteria for an anxiety disorder based on the MINI neuropsychiatric interview (Sheehan *et al.*, 1997); (4) score >10 on the BAI; (5) report anxiety as their primary reason to seek treatment (e.g. in cases with co-morbid depression symptoms). Patients were ineligible when co-morbid depression symptoms were more severe than anxiety symptoms, when they had a diagnosis for which GSH is contraindicated (e.g. social anxiety disorder and post-traumatic stress disorder), when an interpreter was required, and when they were unable to read and/or write.

Current study sample

Data from a sub-sample of participants from the clinical trial were used. The sample was constrained by (a) random session selection and (b) some recording files were corrupt/inaudible. The number of recordings available for rating was $n = 94$, as 16 were excluded from the original 110 total due to being inaudible. During the trial, $n = 180$ chose to receive CAT-GSH compared with $n = 73$ who chose to receive CBT-GSH and so there were more CAT-GSH ($n = 71$) than CBT-GSH sessions ($n = 23$) to rate. Many patients chose CAT-GSH because they were returning to the service and did not want to access the same intervention they had before (i.e. typically a CBT-based intervention). In the CAT-GSH group, 49.3% had previously received a psychological intervention compared with 21.7% of the CBT-GSH group (see Table 1). Eligibility for the trial

Table 1. Summary of patient sample characteristics overall and per treatment condition

		Overall (<i>n</i> = 94)	CBT-GSH (<i>n</i> = 23)	CAT-GSH (<i>n</i> = 71)	Test of difference		
Previous psychological intervention (%)	Yes	40 (42.6)	5 (21.7)	35 (49.3)	$\chi^2 (1) = 5.4, p = 0.020^*$		
	No	54 (57.4)	18 (78.3)	36 (50.7)			
Treatment allocation method (%)	Patient choice	88 (93.6)	21 (91.3)	67 (94.4)	$\chi^2 (1) = 0.27, p = 0.602$		
	Randomised	6 (6.4)	2 (8.7)	4 (5.6)			
Mean age (<i>SD</i>)		36.61 (14.11)	37.04 (14.68)	36.46 (14.02)	$t_{92} = 0.17, p = 0.433$		
Gender (%)	Male	21 (22.3)	6 (26.1)	15 (21.1)	$\chi^2 (1) = 0.25, p = 0.620$		
	Female	73 (77.7)	17 (73.9)	56 (78.9)			
Psychotropic medication (%)	Yes	48 (51.1)	14 (60.9)	34 (47.9)	$\chi^2 (1) = 1.17, p = 0.279$		
	No	46 (48.9)	9 (39.1)	37 (52.1)			
Ethnicity (%)	White British	86 (91.5)	19 (82.6)	67 (94.4)	$\chi^2 (5) = 13.91, p = 0.016^*$		
	White Irish	1 (1.1)	1 (4.3)	0			
	Pakistani	2 (2.1)	2 (8.7)	0			
	Mixed: White and Black Caribbean	3 (3.2)	0	3 (4.2)			
	Caribbean	1 (1.1)	0	1 (1.4)			
	Not stated	1 (1.1)	1 (4.3)	0			
	Single	4 (4.3)	1 (4.3)	3 (4.2)			
Marital status (%)	Married/civil partnership	11 (11.7)	2 (8.7)	9 (12.7)	$\chi^2 (4) = 3.65, p = 0.456$		
	Separated	1 (1.1)	0	1 (1.4)			
	Widowed/surviving civil partner	1 (1.1)	1 (4.3)	0			
	Not stated	77 (81.9)	19 (82.6)	58 (81.7)			
	Full-time employment	50 (53.2)	13 (56.5)	37 (52.1)			
Employment status (%)	Part-time employment	10 (10.6)	2 (8.7)	8 (11.3)	$\chi^2 (7) = 9.62, p = 0.211$		
	Unemployed	16 (17.0)	2 (8.7)	14 (19.7)			
	Self-employed	2 (2.1)	0	2 (2.8)			
	Student	8 (8.5)	3 (13)	5 (7.0)			
	Retired	4 (4.3)	3 (13)	1 (1.4)			
	Volunteer	2 (2.1)	0	2 (2.8)			
	Not stated	2 (2.1)	0	2 (2.8)			
	Minimal (1–2 sessions)	3 (3.2)	0 (0)	3 (4.2)			
	Session attendance groupings (%)	Moderate (3–5 sessions)	18 (19.1)	8 (34.8)		10 (14.1)	$\chi^2 (2) = 5.47, p = 0.065$
		Full (6+ sessions)	73 (77.7)	15 (65.2)		58 (81.7)	
Stepped-up (%)	Yes	15 (16.0)	1 (4.3)	14 (19.7)	$\chi^2 (1) = 3.06, p = 0.080$		
	No	79 (84.0)	22 (95.7)	57 (80.3)			
Dropped out (%)	Yes	11 (11.7)	2 (8.7)	9 (12.7)	$\chi^2 (1) = 0.27, p = 0.606$		
	No	83 (88.3)	21 (91.3)	62 (87.3)			

*Significant at $p < .05$ threshold; significant at Bonferroni adjusted threshold for multiple testing in bold (though none is significant at the adjusted threshold).

required that participants should not access other concurrent psychological interventions while participating in the study.

Practitioners and raters

Interventions were delivered by $n=19$ psychological wellbeing practitioners (PWPs) with postgraduate qualifications. PWPs were mostly females ($n=18$), with a mean age of 31.31 ($SD=5.07$), and a mean of 4.75 ($SD=4.43$, range 1–15) years of post-qualification clinical experience. All PWPs participated in two clinical skills groups (one for each type of GSH) and accessed weekly case management supervision. Six expert independent raters were recruited from outside the service that sponsored the trial to provide independent ratings of a sub-sample of recordings to determine inter-rater reliability. All independent raters were senior PWPs in University teacher roles, who had between 2 and 6 years of experience rating trainee and qualified PWP audio recordings using a validated competence measure described below.

Measures

Practitioner competence

Competence was rated using the Low Intensity Treatment Competency Scale (LITC; Kellett *et al.*, 2021b). The six items cover the following domains: (1) focusing the treatment session; (2) maintaining engagement; (3) interpersonal skills; (4) information gathering specific to change; (5) delivery of a low intensity change method; and (6) homework planning. Given that the LITC was originally developed for rating CBT-GSH, the treatment-specific item (i.e. delivery of a low-intensity change method) was adapted for rating the CAT-GSH sessions to ensure treatment-specific adherence and skills were captured in competence rating. The LITC has good internal consistency ($\alpha=.84$), moderate-to-good inter-rater reliability (.74), adequate construct and predictive validity and a total score of 18 differentiates competent GSH.

Anxiety

Anxiety was measured with the GAD-7 before each GSH session. The GAD-7 is a 7-item measure based on the *Diagnostic and Statistical Manual (DSM-IV*; American Psychiatric Association, 2000). Scores range from 0 to 21 with higher scores indexing greater anxiety severity. The GAD-7 has good internal consistency ($\alpha=.92$; Spitzer *et al.*, 2006). A change score of ≥ 5 points is indicative of reliable change (Richards and Borglin, 2011) and significant anxiety disorder symptoms are identified with a score of ≥ 8 points (National Collaborating Centre for Mental Health, 2020).

Depression

Depression was measured with the PHQ-9 before each GSH session. The PHQ-9 is a 9-item measure of depression with items being derived from *DSM-IV* (American Psychiatric Association, 2000). Scores range from 0 to 27, with higher scores indicating greater depression severity. The scale has good internal reliability ($\alpha=.87$; Kroenke *et al.* 2001). A change score of ≥ 6 points is a reliable change (Richards and Borglin, 2011) and clinically significant symptoms of major depressive disorder are screened using a score of ≥ 10 points (Kroenke *et al.*, 2007).

Functional impairment

Functional impairment was measured using the WSAS before each GSH session. The WSAS is an 8-item measure of disability or functional impairment. Total scores range from 0 to 40, where higher scores indicate greater functional impairment across multiple life domains (e.g. work,

self-care, relationships, hobbies, etc.). The WSAS has good internal consistency ($\alpha = .87$) and test-retest reliability ($r = 0.73$; Mundt *et al.*, 2002).

Session attendance, drop-out, and need for further intervention rates

A treatment completer was defined as a participant that attended at least 6 sessions of GSH or who had recovered from anxiety symptoms before session 6. Session attendance was categorised as: minimal attendance (1–2 sessions), moderate attendance (3–5 sessions) or full attendance (6–8 sessions). Drop-out was defined as premature discontinuation of treatment initiated by the patient. The need for further intervention rate was the percentage of patients referred on for a traditional psychotherapy within the service (i.e. the ‘stepping-up rate’).

Definition of clinical outcomes

On the GAD-7, PHQ-9 and WSAS, a pre–post treatment change score was calculated by subtracting a patient’s baseline score from their last observed score. Measures of reliable change (RC) and reliable and clinically significant improvement (RCSI) were calculated for GAD-7 and PHQ-9. RC indicates whether a patient’s score has changed more than would be expected by measurement error, when comparing the initial score and the last observed (i.e. post-treatment) score (Jacobson and Truax, 1991). Patients that did not have RC were classed as ‘no change’. RCSI is a strict outcome definition that requires a patient to have (1) attained RC and (2) to have a post-treatment score in the sub-clinical range according to the diagnostic cut-off for the relevant questionnaire (Jacobson *et al.*, 1984). It was not possible to carry out WSAS RC/RCSI calculations due to the lack of established cut-off values.

Procedure

Treatment and competency rating

Both forms of GSH involved 6–8 weekly sessions and used highly structured workbooks with homework exercises. The PWP received two days of training on CAT-GSH. This was divided into four sections: (1) key theoretical concepts of reciprocal roles called relationship roles in the patient workbook and procedural sequences called patterns in the patient workbook, (2) the relationship focus of CAT-GSH, (3) modelling and practising sessions 1–3 and (4) modelling and practising sessions 4–6. The sessions contained key themes that PWP know how to deliver GSH and that the training was supporting them to deliver GSH with a different theoretical focus and that adherence to the patient manual was key. The first author (N.P.) and the six independent PWP raters attended separate 3-hour LITC training sessions led by two of the developers of the measure (S.K. and M.S.B.). The PWP educators had experience in using the LITC and the scoring was the same for both versions of GSH. The only difference was that in the change methods section (LITC Item 6) the CBT-GSH and CAT-GSH versions of the measure specified the use of model specific change methods. Therefore, the use of the competency measure did not differ during the assessment of CAT-GSH greatly from the assessment of CBT-GSH. This emphasises the psychoeducational component of each intervention being similar in style but differing in content. The PWP educator training focused on introducing the CAT-GSH protocol and calibrating ratings of two CAT-GSH session recordings that were excluded from the rest of the study.

N.P. rated all available sessions ($n = 94$). Raters were randomly allocated to two rating groups (Groups A and B) and each rating group rated the same six sessions (i.e. two rated as low competence, two rated medium, two rated as high by the primary rater). These sessions were randomly selected from the full sample until two sessions were in each competency category and each category had one from CAT-GSH and one from CBT-GSH. This created a ‘fully crossed’ design of 12 sessions enabling systematic bias between raters to be controlled for in estimating

intraclass correlations (ICC; Hallgren, 2012). The sequence in which raters rated sessions was counterbalanced by generating a random rating sequence for each rater. Inter-rater agreement was calculated by generating a two-way (target \times judges) random-effects ICC (Shrout and Fleiss, 1979) using 'absolute agreement'. Internal consistency of the LITC for the CAT-GSH sessions was determined through calculating Cronbach's alpha (scores above .70 are considered acceptable; Cronbach, 1951).

Data analysis

A series of chi-square, correlation and regression analyses were conducted. G-Power (version 3.1.9.6) indicated that minimal samples sizes of $n=67$ (for correlations) and $n=85$ (for regressions) were required. Statistical significance thresholds for chi-square and correlation analyses were adjusted for multiple tests. Pearson product-moment correlation coefficients tested the association between overall practitioner competence and end of treatment clinical outcomes (i.e. end of treatment score and pre-post change for GAD-7, PHQ-9 and WSAS) for each treatment condition and the overall sample (Hypothesis 1). Correlations for the overall and CAT-GSH samples were sufficiently powered, but the within-group CBT-GSH correlations were not. Cohen's definitions were used to interpret correlation coefficients: 'small' ($r=.10$), 'medium' ($r=.30$) and 'large' ($r=.50$) associations (Cohen, 1992). To establish whether different competence 'levels' were associated with clinical recovery, 'high', 'medium' and 'low' competence levels were created as the nominal variable of RCSI analyses. Competence levels were created using 'competence quartiles' (low = <25%, medium = 25–75%, high = >75%) of overall competence scores (Branson *et al.*, 2018). Chi-squared tests were then performed between competence level and RCSI status for anxiety and depression outcomes, for treatment conditions and the overall sample (Hypothesis 1). To further explore the association between PWP competence and anxiety outcome, a standard multiple regression was conducted. The model used the outcome variable 'end of treatment GAD-7 score', predictor variables 'overall competence' and 'treatment condition' and 'baseline GAD-7 score' and 'session number rated' as covariates (Hypothesis 1).

Chi-square tests were used to determine whether there was a significant association between competence level and session attendance (minimal, moderate or full attendance; Hypothesis 2). A logistic regression tested whether competence ratings were associated with patient drop-out, where the outcome variable was 'treatment drop-out' (yes/no), predictor variables were 'overall competence' and 'treatment condition' and covariates were 'baseline GAD-7 score' and 'session number rated' (Hypothesis 2). Sensitivity analyses were conducted to establish the effect of the timing of competence-rating on the overall competence-outcome and competence-drop-out associations. These analyses were carried out by re-running the multiple regression and logistic regression analyses including only cases that had a competence-rating based on recordings of an 'early' treatment session (session 1 or 2) and comparing this model with the original models.

Chi-square tests also investigated associations between 'competence level' and 'stepping up' (a nominal yes/no variable; Hypothesis 3). Exploratory longitudinal linear mixed models (LMM) were developed to explore the rate of sessional change in anxiety outcomes and the impact of practitioner competence on rate of change (Hypothesis 4). LMM were conducted with maximum likelihood (ML) estimation (a summary of the LMM model-building procedure is available in the Supplementary material online). It was expected that the outcome data contained missing data points due to, for instance, non-attendance and treatment drop-out. For all analyses except LMM, missing outcome data were handled by the last observation carried forward (LOCF) procedure. In terms of LMM models, these procedures are sufficiently flexible to allow for unbalanced data (e.g. unequal sample sizes between groups and missing data) through using maximum likelihood estimation (Rabe-Hesketh and Skrondal, 2008; Shek and Ma, 2011).

Results

Sample description

Participant characteristics are summarised in Table 1 showing that characteristics were comparable between each version of GSH, with no significant differences in terms of sessions attended, step-up rates, dropout rates, treatment allocation methods, nor previous psychological intervention.

Competence ratings

Inter-rater reliability was calculated within each rating group (groups A and B) and across the rating group and the primary rater. All raters in both rating groups (groups A and B) achieved an acceptable level of agreement in competency ratings (see Table S1 in Supplementary material), with all ICCs between 0.85 ('good') and 0.99 ('excellent'; Koo and Li, 2016). The independent ratings therefore indicated that the primary ratings were highly reliable. Competency-items and overall practitioner competence ratings for the total sample and within the two formats of GSH are displayed in Table 2; competence did not differ between CBT-GSH and CAT-GSH. LITC scores from the full sample of $n = 70$ CAT-GSH sessions had acceptable internal consistency ($\alpha = 0.78$).

Clinical outcomes

Table 3 summarises the clinical outcomes. All outcomes in both formats of GSH improved, with evidence of a decreasing trajectory of symptom severity over the course of treatment. Outcomes were not significantly different between the two formats of GSH after adjusting for multiple testing. Similarly, the RC and RCSI outcomes (i.e. shown in Table S2 of the Supplementary material) were not statistically significant between the two versions of GSH.

Competence and outcome (Hypothesis 1)

Table 4 reports the correlation coefficients between overall competence and clinical outcome variables. Neither the end of treatment nor the pre–post change anxiety, depression or functioning outcomes were significantly correlated with overall practitioner competence. There was a moderate correlation ($r = 0.34$) for CBT-GSH between competence and end of treatment anxiety, and a small correlation ($r = 0.16$) in the overall sample (including both types of GSH) between competence and end of treatment anxiety, although none were statistically significant. There were also small non-significant correlations observed for CBT-GSH between competence and pre–post change in anxiety ($r = 0.12$), and for CAT-GSH between competence and end of treatment anxiety ($r = 0.15$) and depression ($r = 0.14$). These correlation coefficients were in the opposite direction as predicted (i.e. they indicated that higher practitioner competence was associated with lower clinical improvement).

There was no association between practitioner competence level and patient recovery rates (see Table S3 in the Supplementary material). The overall multiple regression model explained 15% of the variance in post-treatment anxiety symptoms. After initially controlling for the covariates (which explained 13% of the variance), practitioner competence did not significantly improve model fit (i.e. only accounting for an additional 2% of the variance in anxiety outcomes). The final step assessing whether there was an interaction effect between competence and type of GSH was not statistically significant, indicating that the association between overall competence and end of treatment GAD-7 score did not differ across CBT-GSH and CAT-GSH ($B = -0.28$, $SE = 0.45$, $\beta = -0.49$, $p = 0.540$). See Table S4 in the Supplementary material for results of the regression model. A sensitivity analysis was conducted with only sessions that were rated early in treatment (session 1 or 2) included in the multiple regression model ($n = 39$). The main effect of overall competence on end of treatment GAD-7 score remained non-significant ($B = 0.16$ (95%

Table 2. Descriptive statistics of item and overall PWP competence ratings per treatment condition and for the overall sample

Competence rating item	Overall <i>M</i> (<i>SD</i>) (<i>n</i> = 94)	CBT-GSH <i>M</i> (<i>SD</i>) (<i>n</i> = 23)	CAT-GSH <i>M</i> (<i>SD</i>) (<i>n</i> = 71)	Test of difference
1: Focusing the session	2.47 (0.86)	2.7 (0.86)	2.41 (0.86)	$t_{92} = 1.39, p = 0.167$
2: Continued engagement	3.38 (0.57)	3.37 (0.5)	3.39 (0.59)	$t_{92} = -0.13, p = 0.898$
3: Interpersonal competencies	3.47 (0.68)	3.73 (0.68)	3.4 (0.67)	$t_{92} = 1.68, p = 0.097$
4: Information gathering	3.12 (0.78)	3.3 (0.72)	3.06 (0.8)	$t_{92} = 1.29, p = 0.201$
5: Within session self-help change method	3.61 (0.6)	3.7 (0.54)	3.58 (0.62)	$t_{92} = 0.77, p = 0.444$
6: Planning and shared decision making	3.27 (0.56)	3.46 (0.62)	3.21 (0.53)	$t_{92} = 1.86, p = 0.066$
Overall score	19.34 (2.85)	20.2 (2.69)	19.06 (2.87)	$t_{92} = 1.68, p = 0.096$

Significant at Bonferroni adjusted threshold for multiple testing in bold (though none is significant at the adjusted threshold).

Table 3. Descriptive statistics of outcome measure scores for baseline, end of treatment and pre-post change by treatment condition and overall sample

	Baseline	End of treatment	Pre-post change
Anxiety (GAD-7)			
Overall sample	15.85 (3.95)	8.38 (5.32)	-7.47 (5.84)
CBT-GSH	16.3 (3.35)	7.61 (4.01)	-8.7 (5.83)
CAT-GSH	15.7 (4.13)	8.63 (5.69)	-7.07 (5.83)
Test of difference	$t_{92} = 0.63, p = 0.529$	$t_{92} = -0.8, p = 0.43$	$t_{92} = -1.16, p = 0.25$
Depression (PHQ-9)			
Overall sample	14.98 (4.78)	9.63 (6.08)	-5.35 (5.98)
CBT-GSH	15.96 (3.56)	8.09 (4.81)	-7.87 (4.93)
CAT-GSH	14.66 (5.1)	10.13 (6.39)	-4.54 (6.1)
Test of difference	$t_{92} = 1.13, p = 0.262$	$t_{92} = -1.41, p = 0.163$	$t_{92} = -2.39, p = 0.019^*$
Functioning (WSAS)			
Overall sample	19.96 (7.41)	14.48 (9.08)	-5.48 (8.92)
CBT-GSH	23.30 (6.67)	12.1 (7.96)	-8.3 (7.69)
CAT-GSH	18.87 (7.35)	15.25 (9.34)	-4.56 (9.14)
Test of difference	$t_{92} = 2.57, p = 0.012^*$	$t_{92} = -1.47, p = 0.147$	$t_{92} = -1.77, p = 0.08$

*Significant at $p < .05$ threshold Significant at Bonferroni adjusted threshold for multiple testing in bold (though none is significant at the adjusted threshold).

CI = -0.63-0.95), $SE = 0.39, p = 0.68$) with treatment, baseline GAD-7 score and session number rated being controlled for (see Table S5 in Supplementary material).

Competence and attendance (Hypothesis 2)

There was no significant association between practitioner competence level and session attendance (full details of chi-square analyses are displayed in Table S6 of the Supplementary material). The overall logistic regression model explained 28% of the variation in drop-out. After initially controlling for covariates (i.e. which explained 23% of the variance), practitioner competence did not significantly improve model fit and only accounted for an additional 1% of the variance in drop-out. The final step assessing whether there was an interaction effect between competence and type of GSH was not statistically significant, indicating the association between overall competence and drop-out did not differ between CBT-GSH *versus* CAT-GSH ($B = -0.47, SE = 0.41, \text{Exp}(B) = 0.62, p = 0.246$). See Table S7 in the Supplementary material for results of the

Table 4. Pearson product-moment correlation coefficients between overall competence and clinical outcome variables (post-treatment score and pre–post change)

		Overall sample (<i>n</i> = 94)	CBT-GSH condition (<i>n</i> = 23)	CAT-GSH condition (<i>n</i> = 71)
		<i>r</i> (<i>p</i>)	<i>r</i> (<i>p</i>)	<i>r</i> (<i>p</i>)
Anxiety	Post-Tx GAD-7	0.16 (0.058)	0.34 (0.058)	0.15 (0.107)
	GAD-7 change	0.02 (0.441)	0.12 (0.287)	0.01 (0.461)
Depression	Post-Tx PHQ-9	0.09 (0.197)	0.00 (0.499)	0.14 (0.116)
	PHQ-9 change	-0.04 (0.344)	-0.1 (0.331)	0.02 (0.426)
Functioning	Post-Tx WSAS	0.03 (0.378)	0.08 (0.357)	0.06 (0.325)
	WSAS change	-0.03 (0.398)	0.14 (0.256)	-0.03 (0.402)

Correlations were interpreted using the following thresholds: ‘small’ ($r = .10$), ‘medium’ ($r = .30$) and ‘large’ ($r = .50$) correlations (Cohen, 1992). Tx, treatment. Significance level one-tailed. Significant at Bonferroni adjusted threshold for multiple testing in bold (though none is significant at the adjusted threshold).

Table 5. Fixed effects in the final LMM model

	Coefficient	95% CI	SE	<i>p</i>
Intercept	16.49	13.38, 19.60	1.57	<0.001
Time log	-3.63***	-4.22, -3.03	0.30	<0.001
Competence	-0.37*	-0.68, -0.05	0.16	0.022
Session number rated	0.10	-0.33, 0.54	0.22	0.637
Treatment	-0.27	-1.98, 1.44	0.86	0.756
Competence × Time log	0.07	-0.13, 0.27	0.10	0.509

CI, confidence intervals; SE, standard error. * $p < .05$, *** $p < .001$.

logistic regression model. A sensitivity analysis was conducted, where only the sessions rated early in treatment (session 1 or 2) were included in the model (see Supplementary material Table S8). Again, practitioner competence did not significantly improve model fit (1% added variance explained). The interaction effect between competence and type of GSH was also non-significant in the sensitivity analysis ($B = -0.52$, $SE = 0.41$, $\text{Exp}(B) = 0.6$, $p = 0.238$).

Competence, need for further intervention, and rate of sessional change (Hypotheses 3 and 4)

There was no significant association between practitioner competence level and whether patients required further intervention in the service (see Table S9 in Supplementary material). The LMM was developed to assess the impact of competence on the longitudinal rate of change in anxiety symptoms over the course of GSH sessions, controlling for treatment and session number rated. The best fitting model applied a log-linear growth trend with a AR1 heterogeneous covariance structure (see Supplemental material for results on model building and selection). The coefficients of the final model are displayed in Table 5. There was a significant main effect of competence, indicating that initial anxiety level and competence were associated (e.g. patients with higher baseline anxiety severity were treated by therapists with higher competence ratings). However, the main coefficient of interest was the competence-by-time interaction term, which was non-significant, indicating that practitioner competence was not associated with rate of change during treatment.

Discussion

This study aimed to investigate associations between practitioners' competence in the delivery of guided self-help interventions with patients' clinical outcomes (anxiety, depression and functioning), attendance, drop-out and the need for further intervention.

Contrary to our hypotheses, the results indicate that practitioner competence was not significantly associated with clinical outcomes, drop-out, session attendance or the need for further intervention. Furthermore, practitioner competence was not significantly associated with the rate of change in anxiety severity over the course of treatment. LMM analyses showed that baseline (pre-treatment) anxiety severity and practitioner competence were associated. However, practitioner competence was not associated with the rate of change over the course of treatment. The LITC competence measure was found to be internally consistent when measuring competence to deliver GSH and was reliable across different raters. Correlations suggest small to moderate effects in the opposite direction of pre-defined hypotheses (i.e. higher practitioner competence was associated with higher post-treatment symptoms). Whilst counter-intuitive, these correlations were not statistically significant, and the directionality of associations is plausibly explained by the observation that more competent therapists also treated more anxious patients (who would finish therapy with higher-than-average scores, as they were also more severely impaired at the start of treatment).

These findings contrast with the wider psychotherapy literature, which indicates that higher therapist competence is significantly associated with better treatment outcomes during cognitive therapy (e.g. Strunk *et al.*, 2010; Weck *et al.*, 2011; Weck *et al.*, 2015), cognitive behavioural therapy (e.g. Norrie *et al.*, 2013), cognitive processing therapy (e.g. Marques *et al.*, 2019) and dynamic interpersonal therapy (e.g. Wurman, 2019). It is plausible that there was a narrow range (i.e. variability) of competence ratings in the current study. All practitioners performed at least to a 'good enough' skill level, as they were all qualified, highly experienced in facilitating GSH and were closely supervised. It is also possible that due to the LITC consisting of five non-specific items and only one treatment-specific item, the limited range and comparable ratings across both versions of GSH may reflect comparable and sufficient skills in non-specific/common factors, rather than model distinctive/specific skills. A limited range of competence ratings is a common limitation for therapist competence-outcome studies (e.g. Bisseling *et al.*, 2019; Wurman, 2019) and lack of competence ranges is often observed during clinical trials where clinical performance is more highly scrutinised (Branson *et al.*, 2015).

The finding that the longitudinal rate of clinical change was not associated with practitioner competence was also in contrast with previous findings using traditional psychotherapies (e.g. Strunk *et al.*, 2010; Wurman, 2019). The brevity of GSH means that accelerated change is required to attain a positive outcome during a narrow intervention window. The LMM analyses lacked statistical power and therefore non-significant findings do not provide firm evidence for a lack of a significant effect. Practitioner competence may not influence patient outcome as strongly compared with traditional psychotherapies because the GSH practitioner is required to follow highly structured workbooks (Shafran *et al.*, 2021). Whilst the skills of highly competent GSH may well differ from highly competent traditional psychotherapy, they should be equally valued.

One interpretation of the significant main effect of competence in the longitudinal LMM is that more severely anxious patients were systematically assigned to therapists with higher competence. The reverse-causality interpretation seems less likely (e.g. initial patient anxiety influencing competence ratings), given the direction of this association. If higher initial symptom severity may lead practitioners to deviate from GSH protocols, we would expect to see an association in the opposite direction (e.g. a negative coefficient) to that observed in this sample. Furthermore, there may have been other confounding variables affecting both patient anxiety and practitioner competence but were not accounted for, such as practitioner years of experience, or patient clinical complexity (e.g. co-morbid axis I and/or axis II disorders).

Limitations

The study had several methodological limitations. LOCF procedures were employed to offer protection against attrition bias. The LOCF approach is known to have limitations and may result in biased clinical outcomes (Molnar *et al.*, 2008). Nevertheless, the LMM analysis is robust to missing data points and makes best-use of the available session-by-session measures.

The LITC tool used in the current study was initially designed to measure CBT-GSH practitioner competence (Kellett *et al.*, 2021a). Prior to the current study, this measure did not have psychometric properties in relation to measuring practitioner competence for delivering CAT-GSH and thus the lower rated competence observed in the CAT-GSH treatment arm may have been due to the tool not measuring CAT-GSH competence in a reliable and valid way. Specific amendments were required to the LITC during the current study including amending item 5 of the LITC to capture CAT-GSH specific change techniques. Good inter-rater reliability and internal consistency for CAT-GSH sessions was nevertheless found, with a very high level of agreement across both rating groups.

Previous studies of GSH competence-outcome associations (Branson *et al.*, 2018; Delgadillo *et al.*, 2020) were based on trainee role-plays using a measure that has not been psychometrically validated. The methodology used here is therefore a positive step for the field and further assessment of psychometric properties of the LITC during other versions of GSH will be an important focus for future studies. The study had unequal treatment group sizes and this was due to a large proportion of the patients preferring CAT-GSH. This was possibly due to prior engagement in the service (and so prior experience of CBT-GSH) and so may have introduced sampling bias.

The multi-level modelling analyses in the current study did not have sufficient statistical power. This means that the lack of significant association found between competence and rate of clinical change may have been due to a lack of statistical power to detect an association, as opposed to providing firm evidence for the lack of a meaningful association.

Future research/directions

Sampling bias would be remedied in future studies by using a more conventional random allocation RCT design. Developing and testing measures of GSH adherence (as opposed to competence or fidelity) may be better suited to the style of GSH, given that practitioner competence and patient outcome in GSH did not appear to be meaningfully associated in the current study. Finally, investigations into the mediators and moderators of patient outcome during GSH are needed. These may include practitioner characteristics (e.g. personality traits), patient characteristics (e.g. patient co-morbidity/complexity, duration of anxiety disorder, literacy skills, motivation to change), or completion (or non-completion) of between-session tasks. The LITC is a relatively recently developed measure of GSH practitioner competence. Future research should continue to assess the reliability and validity of the LITC in measuring practitioner competence when delivering GSH to patients with a range of clinical problems.

Clinical implications

CAT-GSH can be delivered by qualified PWP's with comparable levels of competence to CBT-GSH with only brief additional training and ongoing supervision support. This is presumably because CAT-GSH is a *bona fide* version of GSH, and practitioners are simply using differing workbooks and are not learning a completely new clinical method. CAT-GSH presents the opportunity to broaden the availability of different evidenced-based GSH interventions to patients with anxiety. The style and content of CAT-GSH differs from CBT-GSH and so offers patients an opportunity to receive a more analytically informed version of GSH containing an explicit

past-present focus (Wray *et al.*, 2022). The preference rates suggest that analytically informed GSH has some attraction to patients, particularly when CBT-GSH has been tried before.

Conclusions

The current study suggests that qualified GSH practitioners can quickly and successfully be trained to deliver a CAT-GSH treatment protocol with comparable competence to the delivery of CBT-GSH. Practitioner competence was not significantly associated with patients' clinical outcomes or treatment engagement, although a limited range of competence observed may have restricted the likelihood of significant associations being found. Finally, there was no evidence that practitioner competence significantly predicted rate of clinical change across GSH treatments, but these analyses were under-powered and thus do not provide firm evidence for a lack of association between competence and rate of GSH change. Further investigations of the competence-outcome relationship during GSH interventions are clearly indicated and this will match the scale at which these interventions are now being delivered in routine services. As the style, approach, philosophy and duration of GSH markedly differs from traditional psychotherapies, it would appear wise to develop a GSH specific evidence base.

Key practice points

- (1) GSH interventions are being increasingly used in Primary Care psychological services.
- (2) Due to their brevity and low cost relative to more intensive psychotherapies, GSH interventions have great potential to increase access to psychological support at a population level. Hence, understanding how to optimise the effectiveness of GSH is an important goal.
- (3) Research in the wider field of psychotherapy indicates that competent treatment delivery is associated with better treatment outcomes. However, clinical outcomes of GSH interventions do not appear to be associated with competence ratings established using a reliable and validated methodology.
- (4) The LITC competence measure can be used to reliably measure competence of qualified practitioners delivering GSH informed by CBT and CAT modalities.

Further reading

Kellett, S., Bee, C., Smithies, J., Aadahl, V., Simmonds-Buckley, M., Power, N., . . . & Delgadillo, J. (2023). Cognitive-behavioural versus cognitive-analytic guided self-help for mild-to-moderate anxiety: a pragmatic, randomised patient preference trial. *British Journal of Psychiatry*, 1–8. <https://doi.org/10.1192/bjp.2023.78>

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S1754470X24000369>

Data availability statement. Data are available upon request from the authors.

Acknowledgements. None.

Author Contributions. **Niall Power:** Conceptualization (lead), Data curation (lead), Formal analysis (lead), Project administration (lead), Writing - original draft (lead), Writing - review & editing (lead); **Mel Simmonds-Buckley:** Formal analysis (supporting), Methodology (supporting), Supervision (equal), Writing - review & editing (supporting); **Jaime Delgadillo:** Writing - review & editing (supporting); **Jess Smithies:** Data curation (supporting); **Charlotte Bee:** Data curation (supporting); **Stephen Kellett:** Conceptualization (supporting), Methodology (supporting), Supervision (lead), Writing - review & editing (supporting).

Financial support. No funding to report.

Competing interests. There are no competing interests relating to the publication of this manuscript.

Ethical standards. The primary trial from which data from the current study was obtained was ethically approved (IRAS ref: 240751). The consent form and ethical approval for the trial stated that one session would be audio recorded, analysed and

published. The authors confirm that they have abided by the Ethical Principles of Psychologists and Code of Conduct as set out by the BABCP and BPS.

References

- Ali, S., Rhodes, L., Moreea, O., McMillan, D., Gilbody, S., Leach, C., Lucock, M., Lutz, W., & Delgado J. (2017). How durable is the effect of low intensity CBT for depression and anxiety? Remission and relapse in a longitudinal cohort study. *Behaviour Research and Therapy*, 94, 1–8. doi: [10.1016/j.brat.2017.04.006](https://doi.org/10.1016/j.brat.2017.04.006)
- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th edn, text revision).
- Baguley, C., Farrand, P., Hope, R., Leibowitz, J., Lovell, K., Lucock, M., O'Neill, C., Paxton, R., Pilling, S., Richards, D., Turpin, G., White, J., & Williams, C. (2010). *Good Practice Guidance on the use of SelfHelp Materials within IAPT Services*. IAPT.
- Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of Consulting and Clinical Psychology*, 56, 893–897. <https://doi.org/10.1037/0022-006X.56.6.893>
- Bisseling, E. M., Schellekens, M. P., Spinhoven, P., Compen, F. R., Speckens, A. E., & van der Lee, M. L. (2019). Therapeutic alliance – not therapist competence or group cohesion – contributes to reduction of psychological distress in group-based mindfulness-based cognitive therapy for cancer patients. *Clinical Psychology & Psychotherapy*, 26, 309–318. <https://doi.org/10.1002/cpp.2352>
- Branson, A., Myles, P., Mahdi, M., & Shafran, R. (2018). The relationship between competence and patient outcome with low-intensity cognitive behavioural interventions. *Behavioural and Cognitive Psychotherapy*, 46, 101–114. <https://doi.org/10.1017/S1352465817000522>
- Branson, A., Shafran, R., & Myles, P. (2015). Investigating the relationship between competence and patient outcome with CBT. *Behaviour Research and Therapy*, 68, 19–26. <https://doi.org/10.1016/j.brat.2015.03.002>
- Chan, S. W., & Adams, M. (2014). Service use, drop-out rate and clinical outcomes: a comparison between high and low intensity treatments in an IAPT service. *Behavioural and Cognitive Psychotherapy*, 42, 747–759. doi: [10.1017/S1352465813000544](https://doi.org/10.1017/S1352465813000544)
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Cronbach L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334. <https://doi.org/10.1007/BF02310555>
- Cuijpers, P., Donker, T., van Straten, A., Li, J., & Andersson, G. (2010). Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. *Psychological Medicine*, 40, 1943–1957. <https://doi.org/10.1017/S0033291710000772>
- Delgado J., Branson, A., Kellett, S., Myles-Hooton, P., Hardy, G. E., & Shafran, R. (2020). Therapist personality traits as predictors of psychological treatment outcomes. *Psychotherapy Research*, 30, 857–870. <https://doi.org/10.1080/10503307.2020.1731927>
- Delgado J., Kellett, S., Ali, S., McMillan, D., Barkham, M., Saxon, D., ... & Lucock, M. (2016). A multi-service practice research network study of large group psychoeducational cognitive behavioural therapy. *Behaviour Research and Therapy*, 87, 155–161. <https://doi.org/10.1016/j.brat.2016.09.010>
- Firth, N., Barkham, M., Kellett, S., & Saxon, D. (2015). Therapist effects and moderators of effectiveness and efficiency in psychological wellbeing practitioners: a multilevel modelling analysis. *Behaviour Research and Therapy*, 69, 54–62. <https://doi.org/10.1016/j.brat.2015.04.001>
- Gellatly, J., Bower, P., Hennessy S., Richards, D., Gilbody, S., & Lovell, K. (2007). What makes self-help interventions effective in the management of depressive symptoms? Meta-analysis and meta-regression. *Psychological Medicine*, 37, 1217–1228. <https://doi.org/10.1017/S0033291707000062>
- Hallam, C., Simmonds-Buckley, M., Kellett, S., Greenhill, B., & Jones, A. (2021). The acceptability, effectiveness, and durability of cognitive analytic therapy: systematic review and meta-analysis. *Psychology and Psychotherapy: Theory, Research and Practice*, 94, 8–35. <https://doi.org/10.1111/papt.12286>
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8, 23–34. <https://doi.org/10.20982/tqmp.08.1.p023>
- Ho, F. Y. Y., Chan, C. S., Lo, W. Y., & Leung, J. C. Y. (2020). The effect of self-help cognitive behavioral therapy for insomnia on depressive symptoms: an updated meta-analysis of randomized controlled trials. *Journal of Affective Disorders*, 265, 287–304. <https://doi.org/10.1016/j.jad.2020.01.062>
- Ho, F. Y. Y., Yeung, W. F., Ng, T. H. Y., & Chan, C. S. (2016). The efficacy and cost-effectiveness of stepped care prevention and treatment for depressive and/or anxiety disorders: a systematic review and meta-analysis. *Scientific Reports*, 6, 29281. <https://doi.org/10.1038/srep29281>
- Jacobson, N. S., Follette, W. C., & Revenstorff, D. (1984). Psychotherapy outcome research: methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15, 336–352. [https://doi.org/10.1016/S0005-7894\(84\)80002-7](https://doi.org/10.1016/S0005-7894(84)80002-7)
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to denning meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19. <https://doi.org/10.1037//0022-006x.59.1.12>

- Kellett, S., Bee, C., Aadahl, V., Headley, E., & Delgadillo, J. (2021a). A pragmatic patient preference trial of cognitive behavioural versus cognitive analytic guided self-help for anxiety disorders. *Behavioural and Cognitive Psychotherapy*, 49, 104–111. <https://doi.org/10.1017/S1352465820000442>
- Kellett, S., Bee, C., Smithies, J., Aadahl, V., Simmonds-Buckley, M., Power, N., . . . & Delgadillo, J. (2023). Cognitive-behavioural versus cognitive-analytic guided self-help for mild-to-moderate anxiety: a pragmatic, randomised patient preference trial. *British Journal of Psychiatry*, 1–8. <https://doi.org/10.1192/bjp.2023.78>
- Kellett, S., Simmonds-Buckley, M., Limon, E., Hague, J., Hughes, L., Stride, C., & Millings, A. (2021b). Defining the assessment and treatment competencies to deliver low-intensity cognitive behavior therapy: a multi-center validation study. *Behavior Therapy*, 52, 15–27. <https://doi.org/10.1016/j.beth.2020.01.006>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15, 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16, 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine*, 146, 317–325. <https://doi.org/10.7326/0003-4819-146-5-200703060-00004>
- Marques, L., Valentine, S. E., Kaysen, D., Mackintosh, M. A., Dixon De Silva, L. E., Ahles, E. M., . . . & Wiltsey-Stirman, S. (2019). Provider fidelity and modifications to cognitive processing therapy in a diverse community health clinic: associations with clinical change. *Journal of Consulting and Clinical Psychology*, 87, 357–369. <https://doi.org/10.1037/ccp0000384>
- Meadows, J., & Kellett, S. (2017). Development and evaluation of cognitive analytic guided self-help (CAT-SH) for use in IAPT services. *Behavioural and Cognitive Psychotherapy*, 45, 266–284. <https://doi.org/10.1017/S1352465816000485>
- Molnar, F. J., Hutton, B., & Fergusson, D. (2008). Does analysis using “last observation carried forward” introduce bias in dementia research?. *CMAJ*, 179, 751–753. <https://doi.org/10.1503/cmaj.080820>
- Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. M. (2002). The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *British Journal of Psychiatry*, 180, 461–464. <https://doi.org/10.1192/bjp.180.5.461>
- National Collaborating Centre for Mental Health (2020). *The Improving Access to Psychological Therapies Manual* [version 4]. <https://www.england.nhs.uk/wp-content/uploads/2020/05/iapt-manual-v4.pdf>
- Norrie, J., Davidson, K., Tata, P., & Gumley, A. (2013). Influence of therapist competence and quantity of cognitive behavioural therapy on suicidal behaviour and inpatient hospitalisation in a randomised controlled trial in borderline personality disorder: Further analyses of treatment effects in the BOScot study. *Psychology and Psychotherapy: Theory, Research and Practice*, 86, 280–293. <https://doi.org/10.1111/papt.12004>
- Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology*, 75, 829–841. <https://doi.org/10.1037/0022-006X.75.6.829>
- Power, N., Noble, L. A., Simmonds-Buckley, M., Kellett, S., Stockton, C., Firth, N., & Delgadillo, J. (2022). Associations between treatment adherence-competence-integrity (ACI) and adult psychotherapy outcomes: a systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology*, 90, 427–445. <https://doi.org/10.1037/ccp0000736>
- Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and Longitudinal Modeling Using Stata*. STATA Press.
- Richards, D. A., & Borglin, G. (2011). Implementation of psychological therapies for anxiety and depression in routine practice: two-year prospective cohort study. *Journal of Affective Disorders*, 133, 51–60. <https://doi.org/10.1016/j.jad.2011.03.024>
- Shafraan, R., Myles-Hooton, P., Bennett, S., & Öst, L. G. (2021). The concept and definition of ‘low intensity’ cognitive behaviour therapy. *Behaviour Research and Therapy*, 103803. <https://doi.org/10.1016/j.brat.2021.103803>
- Sharpless, B. A., & Barber, J. P. (2009). A conceptual and empirical review of the meaning, measurement, development, and teaching of intervention competence in clinical psychology. *Clinical Psychology Review*, 29, 47–56. <https://doi.org/10.1016/j.cpr.2008.09.008>
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Janavs, J., Weiller, E., Keskiner, A., . . . & Dunbar, G. C. (1997). The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. *European Psychiatry*, 12, 232–241. [https://doi.org/10.1016/S0924-9338\(97\)83297-X](https://doi.org/10.1016/S0924-9338(97)83297-X)
- Shek, D. T., & Ma, C. (2011). Longitudinal data analyses using linear mixed models in SPSS: concepts, procedures and illustrations. *The Scientific World Journal*, 11, 42–76. <https://doi.org/10.1100/tsw.2011.2>
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428. <https://doi.org/10.1037/0033-2909.86.2.420>
- Simmonds-Buckley, M., Osivwemu, E. O., Kellett, S., & Taylor, C. (2022). The acceptability of cognitive analytic therapy (CAT): meta-analysis and benchmarking of treatment refusal and treatment dropout rates. *Clinical Psychology Review*, 102187. <https://doi.org/10.1016/j.cpr.2022.102187>
- Simmonds-Buckley, M., Osivwemu, E. O., Kellett, S., & Taylor, C. (2022). The acceptability of cognitive analytic therapy (CAT): meta-analysis and benchmarking of treatment refusal and treatment dropout rates. *Clinical Psychology Review*, 96, 102187. doi: 10.1016/j.cpr.2022.102187.

- Singla, D. R., Kohrt, B. A., Murray, L. K., Anand, A., Chorpita, B. F., & Patel, V. (2017). Psychological treatments for the world: lessons from low and middle-income countries. *Annual Review of Clinical Psychology*, 8, 149–181. doi: [10.1146/annurev-clinpsy-032816-045217](https://doi.org/10.1146/annurev-clinpsy-032816-045217)
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166, 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>
- Strunk, D. R., Brotman, M. A., DeRubeis, R. J., & Hollon, S. D. (2010). Therapist competence in cognitive therapy for depression: predicting subsequent symptom change. *Journal of Consulting and Clinical Psychology*, 78, 429–437. <https://doi.org/10.1037/a0019631>
- Wakefield, S., Kellett, S., Simmonds-Buckley, M., Stockton, D., Bradbury, A., & Delgadillo, J. (2021). Improving Access to Psychological Therapies (IAPT) in the United Kingdom: a systematic review and meta-analysis of 10-years of practice-based evidence. *British Journal of Clinical Psychology*, 60, 1–37. <https://doi.org/10.1111/bjc.12259>
- Weck, F., Bohn, C., Ginzburg, D. M., & Stangier, U. (2011). Assessment of adherence and competence in cognitive therapy: comparing session segments with entire sessions. *Psychotherapy Research*, 21, 658–669. <https://doi.org/10.1080/10503307.2011.602751>
- Weck, F., Richtberg, S., Jakob, M., Neng, J. M., & Höfling, V. (2015). Therapist competence and therapeutic alliance are important in the treatment of health anxiety (hypochondriasis). *Psychiatry Research*, 228, 53–58. <https://doi.org/10.1016/j.psychres.2015.03.042>
- Wray, A., Kellett, S., Bee, C., Smithies, J., Aadahl, V., Simmonds-Buckley, M., & McElhatton, C. (2022). The acceptability of cognitive analytic guided self-help in an Improving Access to Psychological Therapies service. *Behavioural and Cognitive Psychotherapy*, 50, 493–507. <https://doi.org/10.1017/S1352465822000194>
- Wurman, T.V. (2019). *Therapist competence in dynamic interpersonal therapy and its association with treatment outcome*. Unpublished doctoral dissertation, University College London.

Cite this article: Power N, Simmonds-Buckley M, Delgadillo J, Smithies J, Bee C, and Kellett S (2024). Practitioner competence and clinical outcomes during cognitive behavioural and cognitive-analytic guided self-help for anxiety. *The Cognitive Behaviour Therapist*. <https://doi.org/10.1017/S1754470X24000369>