Differential effect of early response on outcomes in person-centered experiential therapy and cognitive behavioral therapy for the treatment of adult moderate or severe depression

Kerry Ardern

Clinical and Applied Psychology Unit, School of Psychology, University of Sheffield, UK

Scott A Baldwin

Department of Psychology, Brigham Young University, Utah

David Saxon

Clinical and Applied Psychology Unit, School of Psychology, University of Sheffield, UK

Ben Lorimer

Clinical and Applied Psychology Unit, School of Psychology, University of Sheffield, UK

Gillian E Hardy

Clinical and Applied Psychology Unit, School of Psychology, University of Sheffield, UK

Michael Barkham

Clinical and Applied Psychology Unit, School of Psychology, University of Sheffield, UK

Author version

Objective To investigate if session 1–4 PHQ-9 scores are associated with treatment outcome and if there is a differential effect between person-centered experiential therapy (PCET) and cognitive behavioral therapy (CBT).

Methods A secondary data analysis of a prospectively registered and ethically approved pragmatic, non-inferiority randomized controlled trial comparing PCET and CBT for the treatment of moderate or severe depression. Latent Growth Curve Modelling was applied to data from 274 patients who received ≥ five sessions of therapy to investigate the association between change in session 1–4 Patient Health Questionnaire (PHQ-9) scores on a binary end-of-treatment outcome (i.e., reliable and clinically significant improvement; RCSI) and on final-session PHQ-9 scores. Estimated power was 80%.

Results Change in session 1–4 PHQ-9 scores were significantly associated with the probability of RCSI in the PCET condition (p = .002) but not the CBT condition (p = .156). Specifically, greater early-treatment improvement and higher PHQ-9 scores at session 1 were significantly associated with obtaining RCSI in PCET, but not in CBT; this relationship differed significantly between conditions (p = .007). Greater early-treatment improvement was also significantly associated with lower final-session PHQ-9 scores (p<.001), but this relationship did not significantly differ across conditions (p =.121).

Conclusions Early session scores are associated with final-session depression scores, though PCET and CBT manifest distinctively different trajectories for patients achieving RCSI.

Public health significance statement Routine outcome monitoring may be essential to detect early signs of patient-therapy misfit by session 4 in PCET as a process-marker for reviewing and possibly increasing process-guiding interventions that provides clearer structuring during therapy. In CBT specifically, reviewing at session 4 may help reduce the potential for patients dropping out, although early response may provide less of a cue to later outcomes for more severe patients.

Keywords: depression; psychological therapies; treatment outcomes; change patterns; early response.

In general, the earlier stages of psychological treatment are where most gains in symptomology tend to occur (Barkham et al., 2006; Rubel et al., 2015), with early response being significantly associated with better treatment outcomes (e.g., Duffy et al., 2022; Stulz et al., 2007). A recent meta-regression (Klein et al., 2024), comprising >7,000 patients from 72 randomized controlled trials (RCTs) of cognitive behavioral therapy (CBT) versus any comparator treatment, found depression symptom improvement to be greatest during the initial eight sessions of CBT therapy, with a similar result for non-CBT therapies. However, the potential confounds regarding between-study comparisons in meta-analyses have long been recognized (see Baldwin & Imel, 2020). In addition, the authors of this meta-regression note that (1) studies could have benefited from more frequent (i.e., session-by-session) symptom assessments and (2) treatments as conducted within RCTs may not mirror or be wholly representative of psychosocial treatments as delivered in real-world settings.

Related to such criticisms, Röhrig et al. (2024) reported on outcomes from CBT delivered in a natural setting and found that early response was associated with better therapeutic outcomes. Yet, again, this research remained limited by the absence of session-by-session symptom assessments, as well as the exclusion of patients who received less than 10 sessions; particularly, as meta-analytic findings show that patients with an early response at or after session 4 are at least four times more likely to have a positive treatment outcome (Beard & Delgadillo, 2019).

Treatment outcomes by session 4 are pertinent, not only due to the early response literature, but also because of systematic review findings whereby receiving four sessions of psychological therapy has been considered the minimally accepted dose of treatment (Robinson et al., 2020). Yet, change patterns may be nonlinear (Hayes et al., 2007) and they appear most pronounced and more diverse in the early stages of treatment (Rubel et al., 2015). What contributes to differential within-treatment responses remains unclear, and little

attention has been paid to the potential role of therapeutic modality. Overall, there is a need to better understand how differential responses may be related to treatment modality and patient outcomes.

In light of the methodological issues raised, a more systematic approach to the collection of routine data has been achieved by the English National Health Service (NHS) Talking Therapies (NHS-TT) program, formerly known as the Improving Access to Psychological Therapies (IAPT) program (Clark, 2018). A unique feature of the NHS-TT program is that symptom measures focusing on depression, anxiety, and social functioning are completed at every attended session, thereby yielding very large datasets monitoring the course of routine practice.

Within this national program, varying formats of CBT constitute over 80% of all psychological therapies provided (NHS Digital, 2024). Furthermore, since the national rollout of the NHS-TT program in 2007, additional therapy modalities have been included with person-centered experiential therapy (PCET), formerly termed Counselling for Depression, being the second most delivered evidence-based high-intensity treatment after CBT. PCET differs from generic counseling in that therapists work more actively with a client's emotions (Murphy, 2019) which is considered paramount for PCET effectiveness (Bohart & Watson, 2011). Two recent meta-analyses, one of non-directive support therapy (Cuijpers et al., 2024) and one of humanistic-experiential therapies (Duffy et al., 2023) have shown that when these classes of intervention are delivered as bona-fide therapies, there are only marginal differences in outcome between experiential therapy and CBT.

Research results derived from observational studies using data from the NHS-TT (and prior IAPT) program have linked sessional data to differential outcomes. For example, a secondary analysis of 33,243 patients across 103 IAPT (as it was then named) services found the overall effectiveness of CBT and counseling, a precursor to PCET, to be comparable,

though earlier versus later outcomes highlighted differential treatment responses (Pybis et al., 2017). Counseling had greater gains up to session seven, whereas CBT had greater gains after session seven, suggesting a crossover effect. More specifically for early response, counseling was significantly more effective than CBT for patients who received two sessions. Another study using the NHS-TT national dataset also reported differences between treatments as a function of treatment duration in which, for moderately-severe and severe patients, there was a crossover in effectiveness; PCET was more effective up to 6 sessions, and CBT more effective thereafter (Saxon et al., 2024).

However, these studies demonstrating a crossover effect between CBT and PCET have employed nonrandomized research designs which may have contributed to results being a function of confounding variables (e.g., selective treatment assignment). Applying trial methodology to the question of differential early response to treatments is therefore warranted. Moreover, applying such randomization procedures within the context of a trial *embedded* within the routine delivery of these therapy modalities, with session-by-session data collection, would appear to provide a unique lens on the question of early response and differential treatment effects.

Aims

The aim of the current paper was to investigate whether treatment modality moderates the relationship between early symptom change and post-treatment outcomes. We build upon previous research by investigating this relationship in the context of a large, pragmatic non-inferiority randomized trial comparing CBT and PCET embedded within the NHS-TT program and thereby utilizing session-by-session symptom assessments.

Methods

Design and Recruitment: Original Trial

The current study used existing data from a pragmatic non-inferiority randomized controlled trial comparing person-centered experiential therapy (PCET) and cognitive behavioral therapy (CBT) for the treatment of moderate or severe depression ('PRaCTICED'; see Barkham et al., 2021); the trial was prospectively registered and ethically approved (ISRCTN06461651).

The primary trial was embedded within a local NHS-TT primary care service situated in Sheffield, South Yorkshire, UK. Underpinned by the stepped-care approach, the clinical organization focused on treatments for common mental health difficulties comprising a range of low-intensity (e.g., guided self-help, educative) interventions at Step 2 delivered by Psychological Wellbeing Practitioners (PWPs), while Step 3 comprised high-intensity evidence-based psychological therapies (e.g., CBT and PCET) delivered by cognitive behavioral therapists and counselors trained to nationally approved standards. The clinical organization operated through multiple local general practitioner (GP) services grouped into four geographical sectors (i.e., areas) and complimented by an additional city center base.

In line with usual practice and consistent with a stepped-care model, all patients were initially assessed by a low-intensity Psychological Wellbeing Practitioner (PWP) to determine their suitability for receiving a high-intensity treatment. Patients over 18 were referred to the trial if: (1) their main self-determined presenting problem was depression, (2) they scored ≥ 12 on the Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001), and (3) they had no strong preference for either CBT or PCET such that they would decline treatment if offered the treatment not of their choice. At screening for the trial, patients completed the Clinical Interview Schedule-Revised (CIS-R; Lewis et al., 1992) and were accepted into the trial if the CIS-R determined a primary diagnosis of moderate or severe depression. Patients were excluded if they had a previous diagnoses of personality disorder,

schizophrenia, or bipolar, had long-term health conditions, elevated risk, or dependence on alcohol or drugs.

Eligibility and Randomization: Current Study

For the original trial, a total of 761 participants were assessed between November 2014 and August 2018, of which 510 were randomly assigned to either PCET (n = 254) or CBT (n = 256). To investigate how early session scores were associated with treatment outcomes, participants for the current study were required to have received a sufficient dose of treatment judged to be necessary for the proposed analysis. With (up to) four sessions as the independent variable/s, eligible participants were required to have attended at least five sessions. Of the 510 participants randomized into the original trial, 24 participants changed treatment modality for clinical reasons during the course of the trial and were therefore excluded. Of the remaining 486 participants who received randomized treatment, 356 had at least two sessions and were classified as receiving a course of treatment, as defined by the NHS-TT criteria (National Collaborating Centre for Mental Health, 2021). Of those who received such a course of treatment, 68 (19.1%) participants (55.9% CBT; 44.1% PCET) received ≤ four sessions and were therefore excluded. Prior to session 4, there was no significant difference in dropout rates between therapies ($\chi^2(1) = 0.001$, $p = .974$).
The PHQ-9 scores of a further 14 patients (12 CBT, 2 PCET) fell to below clinical caseness (i.e., ≤ 10) before their first session and were also excluded. Thus, for the current study, 274 participants had ≥ five sessions and this therefore comprises the study sample.

Participants

Patients

Patient demographics for the current study are presented in Table 1. Differences between treatment groups for all demographic variables were non-significant (left portion of

Table 1). Participants were mostly White-British, middle-aged and employed, and there were more females than males.

Counselors and Therapists

The original trial employed therapists and counselors who met nationally-approved standards regarding their qualification to deliver either PCET or CBT. Experienced, national trainers assessed treatments as delivered at sessions two, six, and 12, if available, using adherence scales designed specifically for each modality; for PCET, the 10-item Person-Centred Experiential Psychotherapy Scale (PCEPS; Freire et al., 2014) was used and, for CBT, the 12-item Cognitive Therapy Scale-Revised (CTS-R; Blackburn et al., 2001) was used. A score of 40 of 60 was considered a pass mark on the PCEPS and a score of 36 of 72 was considered a pass mark on the CTS-R. The mean total scores for sampled adherence ratings were 39.3 (*SD* 10.0; *n*=60) for PCET and 40.8 (*SD* 9.7; *n*=72) for CBT.

The present study employed data from 42 of these therapists and counselors (28 CBT, 14 PCET) and analyses showed small differences between these two groups with counselors being significantly older, having a greater number of years professional experience, and more working part-time. There were no significant differences between therapist groups in the (distribution of) number of patients seen. (see Table A in Supplemental Materials).

Person-centered experiential therapy (PCET)

In the study-specific sample, PCET was delivered by 14 PCET-trained counselors: 12 females and two males. All counselors were accredited by a recognized counseling body. On average, they had been working in their current role for 9.89 (*SD* = 5.33) years and 17.43 (*SD* = 6.02) years of experience as a practicing counselor. In their PCET training, they had completed recordings of 80 sessions, of which four were randomly assessed for adherence by expert trainers using the PCEPS (Freire et al., 2014). Only counselors who completed training and whose sessions passed the required adherence checks were included in the trial.

Cognitive behavioral therapy (CBT)

Beckian CBT was delivered by 28 CBT therapists: 22 females and six males. All CBT therapists were accredited by the British Association for Behavioral and Cognitive Psychotherapy. On average, they had been working in their current role for 7.89 (*SD* = 2.59) years and had 8.46 (*SD* = 3.20) years professional experience. All CBT therapists received Beckian CBT refresher training before participating in the trial.

Supervision

Therapists received individual and group supervision, equating to 4.5 hours of supervision per month. For the counselors, supervisors were PCET-qualified, and the CBT therapists received supervision from trained CBT supervisors.

Measures

As the primary trial was embedded in a routine service, data collected by the service was recorded, including sessional measures but only at the level of total scores, and were subsequently downloaded for trial patients to the research team. The measures comprised the following:

Demographic Characteristics

Demographics of participating patients included gender, age, ethnicity, employment status, medication status, and the Index of Multiple Deprivation (IMD; Ministry of Housing, Communities and Local Government, 2019) – a UK measure of deprivation based on small geographical areas. IMD scores range from 1 (most deprived) to 10 (least deprived).

Depression Symptoms

The Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001), a 9-item self-report psychological screening instrument, measured symptoms of depression at each session. Scores range from 0-27, with higher scores indicating higher severity. In line with the Talking Therapies program, national data, and guidance, RCSI was operationalized as a

PHQ-9 of ≤ 9 at the end of treatment with a change score ≥ 6, compared to the first session PHQ-9 score. Scores above ≥ 10 (i.e., NHS-TT clinical cut off) have a sensitivity and specificity of 88% for major depressive disorder, with high internal and test-retest reliability (Kroenke et al., 2001). The PHQ-9 at the initial screening assessment for the trial was used to calculate the internal consistency of the measure as session-specific item-level data was not available. Cronbach's alpha was .72 for the 510 original trial participants and .71 for the 274 participants included in the current analysis.

Anxiety Symptoms

The Generalized Anxiety Disorder-7 (GAD-7; Spitzer et al., 2006), a 7-item self-report questionnaire, measured symptoms of generalized anxiety at each session. Scores range from 0-21, with higher scores indicating higher severity. A threshold score of 10 has a specificity of 82% and a sensitivity of 89% (Spitzer et al., 2006). The clinical cut off within the Talking Therapies program is scores ≥ 8.

Impairment of Functioning

The Work and Social Adjustment Scale (WSAS; Mundt et al., 2002), a 5-item self-report Likert scale, measured impairment of functioning at each session. Each item is scored between 0 (not at all) and 8 (very severely), with higher scores indicating a higher impact on functioning. The WSAS is a reliable and valid tool with good internal consistency and test-retest reliability (Mundt et al., 2002).

Data Analysis

Latent growth curve modeling (Bollen & Curran, 2006) examined how early symptom change was associated with (1) end-of-treatment reliable and clinically significant improvement (RCSI) and (2) final session PHQ-9 score. Figure 1 depicts the growth model predicting RCSI. An identical model using final session PHQ-9 was also estimated. A

multiple group approach was used to test whether parameters differed across the CBT and PCET conditions.

First, ignoring treatment condition, a linear growth curve was fit to describe change in the PHQ-9 over sessions 1–4. The results suggested that session specific residuals fit significantly better than a single residual value across time ($\chi^2(3) = 13.03, p = 0.005$). Thus, all primary analyses used session-specific residuals. Then, a model was fit to determine whether the growth parameters for sessions 1–4 differed across treatment conditions. Time was scaled so that the mean and variance of the latent intercept (CBT: μ_1, σ_1^2; PCET: μ_3, σ_3^2) represent the average session 1 and variability of session 1 scores across all participants. Likewise, the mean and variance of the latent slope (CBT: μ_2, σ_2^2; PCET: μ_4, σ_4^2) represent the average rate of change for every additional session from sessions 1–4, and the variability of that change across participants. As change in the PHQ-9 was modeled, higher values for the latent intercept indicate more severe depressive symptoms than lower values. Negative values for the latent slope indicate a reduction in symptoms on a session-to-session basis; thus, the more negative the latent slope, the greater the reduction in symptoms from one session to the next.

For the sessions 1–4, steps described in Bollen and Curran (2006) were followed to determine which parameters could be constrained across groups. First, a fully unconstrained model was fit; second, the latent means across groups were constrained ($\mu_1 = \mu_3$ and $\mu_2 = \mu_4$); third, the latent variances across groups were constrained ($\sigma_1^2 = \sigma_3^2$ and $\sigma_2^2 = \sigma_4^2$); and finally the residual variances across the groups were constrained ($\sigma_{e1}^2 = \sigma_{e5}^2$, $\sigma_{e2}^2 = \sigma_{e6}^2$, $\sigma_{e3}^2 = \sigma_{e7}^2$, and $\sigma_{e4}^2 = \sigma_{e8}^2$). Fit was compared using a likelihood ratio test.

Models were then extended to include treatment outcome. The latent intercept and slope for each treatment were used to predict (1) end-of-treatment RCSI (using a logistic model within the structural equation model) and (2) final session PHQ-9 score (using a

Gaussian model; cf. Figure 1). Again, a multiple group model examined whether the regression coefficients differed across treatment groups. First, the intercept for the outcome was constrained (RCSI or final session PHQ-9 score, hereafter referred to as the constant) across groups; then, the slope for the association between change over sessions 1–4 and outcome were constrained ($\beta_3 = \beta_4$), and then the slope for the association between session 1 PHQ-9 score and outcome were constrained ($\beta_1 = \beta_2$). Fit was compared using a likelihood ratio test. All models were estimated in sem or gsem in Stata version 18 (StataCorp, 2023).

With the primary analysis, logistic regression functions in G*Power were used to perform a sensitivity analysis (Faul et al., 2007) for the logistic regression predicting RCSI. The sensitivity analysis (i.e., detectable difference analysis) examined the odds-ratio that describes the interaction between early change and treatment condition. In the multiple group analysis, that interaction is represented by the difference between the early change-RCSI relationship in the two conditions. In this analysis, an odds-ratio between 0 and 1 indicates that an early improvement (i.e., a reduction in session 1-4 PHQ-9 scores) was more strongly related to RCSI in the PCET condition than the CBT condition; an odds-ratio above 1 would indicate the oppositive. The results of the sensitivity analysis suggested that, with 274 patients, we had 80% power to detect population odds-ratios ranging from 0.1 to 0.6. See supplementary Table B and Figure A for more information.

Results

There were no significant differences in patient demographics between treatment groups. When comparing patient demographics based on end-of-therapy outcomes, however, there were significant differences in employment status ($p = .014$) and gender ($p = .004$). Those participants who obtained RCSI were significantly more likely to be female and less likely to be employed or sick/disabled. All other demographic variables were non-significant (see right-hand portion of Table 1).

Table 2 outlines the clinical characteristics of participants. The mean number of therapy sessions was significantly lower in the PCET group (M = 11.36) than the CBT group (M = 12.60; p = .049). There were no significant differences in session 1 PHQ-9, GAD-7, or WSAS scores between treatment groups, or in the rate of PHQ-9 change between session 1 to 4. Fewer participants dropped out or declined treatment in the PCET group (14.9%) than the CBT group (24.2%) but the difference was not significant (p = 0.51).

When comparing outcome groups (i.e., participants meeting RCSI or not), there was no significant difference in RCSI between therapy groups (PCET/CBT), session 1 PHQ-9 or GAD-7 scores, or number of sessions attended. However, there were significant differences between session 1 WSAS scores and session 1–4 PHQ-9 change scores. Those participants who recovered had significantly lower WSAS scores at session 1 (p = .011) and significantly more change in PHQ-9 scores for sessions 1–4 (p <.001). Additionally, participants who recovered were significantly less likely to have dropped out or declined treatment (p <.001).

Change over Sessions 1–4

Table 3 presents the fit statistics and model comparison tests used to determine whether growth model parameters—latent means, latent variances, latent covariance, and residual errors—differed across treatment conditions in sessions 1–4. Higher values of the comparative fit index (CFI) and Tucker-Lewis index (TLI) indicate better fit than low values. Lower values of the root mean square error of approximation (RMSEA), Akaike information criteria (AIC), and Bayesian information criteria (BIC) suggest preferred fit (comparative to higher values). The null hypothesis of the model comparison tests is that introducing constraints does not degrade fit as compared to a less constrained model. The fit indices all suggest that a model where all growth model parameters are constrained across groups fits the data the best. Likewise, the model comparison test suggests that a fully constrained model across groups fits as well as any of the other models that have at least one of the parameter

types free to vary across conditions. Thus, all models for the last session analysis constrained the growth model parameters across condition.

The average starting point for patients at session 1 was 18.0, and there was variability in this starting point $\sigma_1^2 = \sigma_3^2 = 14.5$. The average rate of change (i.e., change expected for one additional session) was -1.03 ($p < .001$), indicating that on average after the first four sessions, patients in both conditions would have obtained approximately a 3-point reduction in PHQ-9 scores. There was also variability in the rate of change across patients of 1.4 PHQ-9 points. Thus, there was a significant reduction in PHQ-9 scores during sessions 1–4 but while the amount of change varied across patients, the amount did not vary across treatment condition.

Predicting Reliable and Clinically Significant Improvement (RCSI)

Table 4 presents the fit statistics and model comparison tests used to determine whether the relationship between the latent intercepts and slopes and last session RCSI differed across treatment groups. The results suggest that the best fitting model was one where the constant for RCSI was constrained to be equal across groups, as any differences among the fit indices for models 1-3 (see Table 4) were small. The AIC, BIC, and model comparison tests showed that constraining slope-RCSI and intercept-RCSI significantly degraded fit.

The right-hand side of Table 5 provides the model coefficients for the best fitting RCSI model. The odds-ratio for the latent slope-RCSI relationship in the CBT condition was not significant ($e^{\beta_3} = 0.734$, $p = 0.156$), indicating no relationship between session 1–4 change and odds of RCSI. The odds-ratio for the latent slope-RCSI relationship in the PCET condition was significant ($e^{\beta_4} = 0.258$, $p = 0.002$), indicating less change between session 1–4 was associated with reduced odds of the scores obtaining RCSI. As previously stated, model comparison test indicated that the difference between the treatment conditions (β_3 and

β_4) was statistically significant (see Table 4; cf. Gelman & Stern, 2006; Nieuwenhuis et al., 2011).

Finally, the odds-ratio for the latent intercept-RCSI relationship in the CBT condition was not significant ($e^{\beta_3} = 0.963$, $p = 0.387$), indicating no relationship between session 1 PHQ-9 score and odds of RCSI. The odds-ratio for the latent intercept-RCSI relationship in the PCET condition was significant ($e^{\beta_4} = 0.008$, $p = 0.002$), indicating that patients with a higher session 1 PHQ-9 score had lower odds of RCSI. As noted previously (see Table 4), this difference between CBT and PCET was statistically significant.

Figure 2 shows the predicted probability of RCSI in each treatment condition (y-axis) as a function of improvement during sessions 1–4 (i.e., latent-slope; x-axis). That is, each data point summarizes the change for each individual participant comprising (an average of) the four PHQ-9 scores across the initial four sessions. The plot shows that CBT was fairly robust to change that was less than the mean change (about -1 PHQ-9 unit per session) when assessing the probability of RCSI. In contrast, PCET patients whose change on the PHQ-9 was *less* than the mean change across the initial four sessions had a lower probability of RCSI. However, the probability increases notably in the PCET condition as a function of the amount of change *above* the mean, though the differences in probability of RCSI between CBT and PCET conditions are less pronounced when there is more change (i.e., to the right of the plot) compared with where there is less change (i.e., to the left of the plot). As a sensitivity analysis to rule out therapist effects, we reran the structural equation model analyses with robust standard errors and Wald tests to evaluate the constraints on prediction models; findings did not meaningfully change.

Predicting Last Session Scores

Table 6 presents the fit statistics and model comparison tests used to determine whether the association between the latent intercepts and slopes and last session PHQ-9

differed across groups. The results suggest that the best fitting model was one where the constant and slope of last PHQ-9 coefficients (β_1 and β_2) were constrained to be equal across treatment groups, as differences among the fit indices for models 1-3 (see Table 6) were small. Adding a constraint to the latent intercept-last PHQ-9 coefficients (β_1 and β_2) to be equal across groups consistently degraded fit according to the fit indices, though the differences were not large. Additionally, constraining β_2 and β_4 resulted in a significant likelihood-ratio test ($\chi^2(1) = 6.03$, $p = 0.014$).

The left-hand side of Table 5 provides the model coefficients for the best fitting last PHQ-9 model. The bottom portion of Table 5 provides the outcome portion of the analysis. The coefficient for the latent intercept for the CBT condition was $\beta_1 = 0.531$, $p < 0.001$, indicating that higher session 1 PHQ-9 scores were associated with higher last session PHQ-9 scores. The coefficient for the latent intercept for the PCET condition was $\beta_2 = 0.623$, $p < 0.001$, indicating that higher values on session 1 PHQ-9 were associated with higher values on final-session PHQ-9; this difference between CBT and PCET treatment conditions was statistically significant (see Table 6). The coefficient for the latent slope ($\beta_3 = \beta_4$) = 1.993, $p < 0.001$, shows that that the larger the rate of change value (larger values indicate less change) the larger the last session PHQ-9 score was. This coefficient was the same (i.e., did not significantly differ) across treatment groups ($p = .121$); see supplementary Figure B which visually depicts the findings from this secondary analysis.

Discussion

In the context of secondary data analyses of a randomized controlled trial, this study investigated the prognostic value of early session scores, comparing person-centered experiential therapy (PCET) and cognitive behavioral therapy (CBT) for the treatment of moderate and severe depression and its relationship to a potential crossover effect. Results indicated that an improvement in session 1–4 depression scores significantly increased the

odds of improvement on the final-session PHQ-9 score. Further, in both PCET and CBT, higher PHQ-9 scores at session 1 were significantly associated with higher final-session PHQ-9 scores.

However, although there was no significant difference in (overall) rates of RCSI between PCET and CBT, treatment mode significantly moderated the association with early-session depression scores and end-of-treatment RCSI. In the PCET group, patients with higher (i.e., worse) scores on a depression measure at session 1 were less likely to obtain reliable and clinically significant improvement. However, in the CBT group, PHQ-9 scores at session 1 were not significantly associated with end-of-treatment RCSI. These effects yield a differential finding with implications for these two treatment modalities early on in the course of treatment.

When comparing the effect of session 1–4 change scores on end-of treatment RCSI, the difference took the form of a crossover effect. Minimal gains or deterioration by participants receiving PCET, as compared with CBT, was predictive of a low probability of subsequently obtaining RCSI. By contrast, greater gains achieved in PCET in the initial four sessions had a higher probability of achieving RCSI than CBT (recall Figure 2). Conceptually, findings suggest that participants were more likely to achieve RCSI in PCET if they reported approximately ≥ 2-point improvement in their depression scores, per session, over sessions 1–4. Conversely, those who made improvements below the average rate of change, or deteriorated further, were more likely to achieve RCSI in CBT.

While current differential findings, between treatments, were dependent of whether the PHQ-9 outcome was binary (i.e., RCSI) or continuous, any benefits to statistical gains with a continuous outcome measure are negated by the aggregation of mean symptom severity ratings at the group-level, as this obscures within-individual change. Pertinently, this within-individual change matters when it comes to clinical decision making and the

identification of differential responses to treatment. In psychological therapies research, the RCSI criterion is commonly applied and cited (Jacobson & Truax, 1991) and given that partial symptom remission is a well-documented risk factor for short-term depression relapse (e.g., Wojnarowski et al., 2019), the clinically stringent and statistically conservative RCSI outcome, that prioritises full remission of symptoms, is necessary. Further, clinical interpretations are limited when research fails to operationalize therapeutic outcomes as they are understood and reported clinically. The absence of clinically graspable findings remains a shortfall in research, particularly when using practice-based data with complex statistical analyses.

The current findings have implications not only for initial patient-treatment matching (i.e., assignment), but also for early treatment monitoring and therapist responsiveness regarding, in the current study, person-centered experiential therapy, that may generalize to other non-CBT therapies. The psychological therapies literature provides examples of stage-wise models reflecting the differing phases of therapy, for example the phase model (e.g., Howard et al., 1993). It may be that the initial stage of the phase model (remoralization) has a more immediate impact on some patients receiving PCET, sufficient to yield greater or lesser change over the initial sessions.

In terms of clinical strategies for responding to such situations, the model of context-responsive psychotherapy integration (e.g., Constantino et al., 2021) postulates an *if-then* clinical contingency, which in the current context for PCET would approximate to *if* ROM data in sessions 1–4 does not show improvement, *then* review with the patient a range of options that could comprise the following: (a) floundering and needing greater process guiding and model explanation within the PCET model (i.e., enhance key clinical action); (b) check with patient possible mismatch between depression score and their experience (i.e., PHQ-9 items may not be tapping their core concerns); or (c) review the possibility of a

therapy-patient mis-match (i.e., would they rather receive CBT). These clinical options may equally be options for the person-centered experiential therapist empathically tracking the patient and being attuned to their discomfort with the therapy (Watson, 2021).

Interestingly, by contrast, the current study yields very little evidence that would guide flexibility of therapeutic actions during CBT treatment as it appears robust to the influence of initial scores and early response. This is in contrast with many previous studies where early symptom improvement (i.e., up to four sessions) was significantly associated with RCSI in CBT (Beard & Delgadillo, 2019; Duffy et al., 2022). However, it may be that the sensitivity of the early response was dampened in CBT due to the patient sample as a whole being more severe and CBT tending to focus on the later remediation phase (Howard et al., 1993); this is consistent with previously reported crossover effects whereby favourable outcomes in CBT, for moderately severe or severely depressed patients, happen in the later stages of therapy (Saxon et al., 2024). Possibly, when CBT therapists become aware that progress is slow, they change their approach and utilize different CBT strategies.

Alternatively, it may be that rather than therapist-initiated actions, patients act themselves by dropping out of therapy, as indicated by the 9.3% higher dropout during treatment in CBT than PCET. If this is the case, reviewing progress at session 4 is important for CBT therapists as well in order to identify and discuss options with patients who are considering dropping out of therapy.

To the authors' knowledge, this study is the first to specifically investigate, in a controlled setting, if the effect of early session depression scores on outcome significantly differed between PCET and CBT during treatment for moderate-severe depression. While there was no main effect of therapy type on treatment outcomes, the association between early session depression scores and RCSI depended on the therapy received. Understanding how treatment mode affects the relationship between early session scores and treatment

outcome is important for future investigations. These findings begin to unpack the importance of reviewing progress in the context of therapy type, as well as patient and process factors already suggested in the routine outcome monitoring and feedback literature (Barkham et al., 2023)

Lastly, it is important to consider the measurement used, as PCET places more value on the client's narration of their own experiences (Angus, 2012), whereas CBT focuses on symptom reduction and behavioral change (Tolin, 2010). Thus, the PHQ-9 may complement CBT treatments, but they do not align with the ethos of PCET. Within-session progress feedback has been found to improve outcomes and reduce dropouts, though pertinently, these effects were moderated by the feedback instrument used (Barkham et al., 2023; de Jong et al., 2021).

Limitations and Strengths

Randomized data reduces the bias that may present when clients are selectively assigned to treatments; however, these data often lack 'real world' applicability. Therefore, using routine but randomized practice data was a strength of the present study. While this study is the first, to the authors' knowledge, to compare differential session-specific outcomes in PCET and CBT within a trial design, excluding patients with ≤ four sessions was a limitation of the study design – especially considering the finding that early session effects on RCSI were moderated by treatment modality. However, our decision was to test our research questions utilizing trial data that was more stringent as a test of treatment received given it matched our definition of the per protocol condition in the original trial (defined as in receipt of 4 or more therapy sessions).

Further, The LGCMs included session-by-session scores and treatment mode; while parsimonious models are encouraged in research so that variance can be explained with as few covariates as possible (Field, 2018), controlling for possible confounds would have

reduced the likelihood of biased estimations (Kahan et al., 2014; McNamee, 2005). That said, reporting on models for both a binary and continuous outcome was a strength of the study design, as doing so highlighted that differential findings were also dependent on how treatment outcome was operationalised.

Conclusion

In contrast to the body of literature attesting either to no or small differences between psychological therapies based on group means (e.g., Wampold et al., 1997), pursuing a strategy of determining predictive models appears to capture nuances that are, effectively, washed out in more traditional research approaches. It also appears more consistent with an agenda focused on precision (i.e., where are the differences?) rather than either suppressing them or arguing them away as being a function of traditionally under-valued small effects. In addition, the study shows the merits of cycling around findings from practice-based data and then extending them within the stringent setting of a trial.

Acknowledgements

We thank members of the Psychotherapy Evaluation and Research Lab at Sheffield (PEARLS) for helpful discussions. Please refer to the publication of the primary trial for full acknowledgments (see Barkham et al., 2021).

References

Angus, L. (2012). Toward an integrative understanding of narrative and emotion processes in Emotion-focused therapy of depression: Implications for theory, research and practice. *Psychotherapy Research*, *22*(4), 367–380. https://doi.org/10.1080/10503307.2012.683988

Baldwin, S. A., & Imel, Z. E. (2020). Studying specificity in psychotherapy with meta-analysis is hard. *Psychotherapy Research*, *30*(3), 294-296. https://doi.org/10.1080/10503307.2019.1679403

Barkham, M., Connell, J., Stiles, W. B., Miles, J. N. V., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose-effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology*, *74*(1), 160–167. https://doi.org/10.1037/0022-006X.74.1.160

Barkham, M., De Jong, K., Delgadillo, J., & Lutz, W. (2023). Routine outcome monitoring (ROM) and feedback: Research review and recommendations. *Psychotherapy Research*, *33*(7), 841–855. https://doi.org/10.1080/10503307.2023.2181114

Barkham, M., Saxon, D., Hardy, G. E., Bradburn, M., Galloway, D., Wickramasekera, N., Keetharuth, A., Bower, P., King, M., Elliot, R., Gabriel, L., Kellett, S., Shaw, S., Wilkinson, T., Connell, J., Harrison, P., Ardern, K., Bishop-Edwards, L., Ashley, K.,…& Brazier, J. E. (2021). Clinical and cost-effectiveness of person-centred experiential therapy vs. cognitive behavioural therapy for moderate and severe depression delivered in the English Improving Access to Psychological Therapies national programme: A pragmatic randomised non-inferiority trial [PRaCTICED]. *The Lancet Psychiatry*, *8*(6), 487-499. https://doi.org/10.1016/S2215-0366(21)00083-3

Beard, J. I. L., & Delgadillo, J. (2019). Early response to psychological therapy as a predictor of depression and anxiety treatment outcomes: A systematic review and meta-analysis. *Depression and Anxiety*, *36*(9), 866–878. https://doi.org/10.1002/da.22931

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(4), 431-446.

Bohart, A. C., & Watson, J. C. (2011). Person-centered psychotherapy and related experiential approaches. In S. B. Messer & A. S. Gurman (Eds.), *Essential psychotherapies: Theory and practice* (pp. 1–38). Guilford Press.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective* (Vol. 467). John Wiley & Sons.

Clark, D. M. (2018). Realizing the mass public benefit of evidence-based psychological therapies: the IAPT program. *Annual Review of Clinical Psychology, 14*(1), 159-183. https://doi.org/10.1146/annurev-clinpsy-050817-084833

Constantino, M. J., Goodwin, B. J., Muir, H. J., Coyne, A. E., & Boswell, J. F. (2021). Context-responsive psychotherapy integration applied to cognitive behavioral therapy. In J. C. Watson & H. Wiseman (Eds.), T*he responsive psychotherapist: Attuning to clients in the moment*. APA. https://doi.org/10.1037/0000240-008

Cuijpers, P., Miguel, C., Ciharova, M., Harrer, M., & Karyotaki, E. (2024). Non-directive supportive therapy for depression: A meta-analytic review. *Journal of Affective Disorders, 349*, 452–461. https://doi.org/10.1016/j.jad.2024.01.073

de Jong, K., Conijn, J. M., Gallagher, R. A., Reshetnikova, A. S., Heij, M., & Lutz, M. C. (2021). Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis. *Clinical Psychology Review*, *85*, 2-19. https://doi.org/10.1016/j.cpr.2021.102002

Duffy, K. E. M., Simmonds-Buckley, M., Haake, R., Delgadillo, J., & Barkham, M. (2023). The efficacy of individual humanistic-experiential therapies for the treatment of depression: A systematic review and meta-analysis of randomized controlled trials. *Psychotherapy Research*, 1–16. Advance online publication. https://doi.org/10.1080/10503307.2023.2227757

Duffy, K. E. M., Simmonds-Buckley, M., Saxon, D., Delgadillo, J., & Barkham, M. (2022). Early response as a prognostic indicator in person-centered experiential therapy for depression. *Journal of Counseling Psychology*, *69*(6), 803–811. https://doi.org/10.1037/cou0000633

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3 : A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. https://doi.org/10.3758/bf03193146

Field, A. (2018). *Discovering statistic using IBM SPSS statistics*. Sage Publications.

Freire, E., Elliott, R., & Westwell, G. (2014). Person-Centred and Experiential Psychotherapy Scale: Development and reliability of an adherence/competence measure for person-centred and experiential psychotherapies. *Counselling and Psychotherapy Research*, *14*(3), 220–226. https://doi.org/10.1080/14733145.2013.808682

Gelman, A., & Stern, H. (2006). The difference between “significant” and “not significant” is not itself statistically significant. *The American Statistician, 60*(4), 328-331. https://doi.org/10.1198/000313006X152649

Hayes, A. M., Laurenceau, J.-P., Feldman, G., Strauss, J. L., & Cardaciotto, L. (2007). Change is not always linear: The study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review*, *27*(6), 715–723. https://doi.org/10.1016/j.cpr.2007.01.008

Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy outcome: causal mediation of change. *Journal of consulting and clinical psychology*, *61*(4), 678.

Jacobson, N., and Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.

Kahan, B. C., Jairath, V., Doré, C. J., & Morris, T. P. (2014). The risks and rewards of covariate adjustment in randomized trials: An assessment of 12 outcomes from 8 studies. *Trials*, *15*(1), 139. https://doi.org/10.1186/1745-6215-15-139

Klein, T., Breilmann, J., Schneider, C., Girlanda, F., Fiedler, I., Dawson, S., Crippa, A., Priebe, S., Barbui, C., Becker, T., & Kösters, M. (2024). Dose–response relationship in cognitive behavioral therapy for depression: A nonlinear metaregression analysis. *Journal of Consulting and Clinical Psychology, 92*(5), 296–309. https://doi.org/10.1037/ccp0000879

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lewis, G., Pelosi, A. J., Araya, R., & Dunn, G. (1992). Measuring psychiatric disorder in the community: A standardized assessment for use by lay interviewers. *Psychological Medicine*, *22*(2), 465–486. https://doi.org/10.1017/S0033291700030415

McNamee, R. (2005). Regression modelling and other methods to control confounding. *Occupational and Environmental Medicine*, *62*(7), 500–506. https://doi.org/10.1136/oem.2002.001115

Ministry of Housing, Communities & Local Government. (2019). *English indices of deprivation 2019*. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019

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(5), 461–464. https://doi.org/10.1192/bjp.180.5.461

Murphy, D. (2019). *Person-centred experiential counselling for depression* (2nd ed.). Sage Publishing.

National Collaborating Centre for Mental Health. (2021). The Improving Access to *Psychological Therapies Manual.* NHS England. https://www.england.nhs.uk/wp-content/uploads/2018/06/the-nhs-talking-therapies-manual-v6.pdf

National Health Service Digital. (2024). *NHS Talking Therapies: Supplementary analysis of therapy-based ouctomes, England 2022-2023.* NHS England. https://digital.nhs.uk/data-and-information/publications/statistical/nhs-talking-therapies-for-anxiety-and-depression-annual-reports/2022-23

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E. J. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nature Neuroscience*, *14*(9), 1105-1107. https://doi.org/10.1038/nn.2886

Pybis, J., Saxon, D., Hill, A., & Barkham, M. (2017). The comparative effectiveness and efficiency of cognitive behaviour therapy and generic counselling in the treatment of depression: Evidence from the 2nd UK National Audit of psychological therapies. *BMC Psychiatry*, *17*(1), 215. https://doi.org/10.1186/s12888-017-1370-7

Robinson, L., Delgadillo, J., & Kellett, S. (2020). The dose-response effect in routinely delivered psychological therapies: A systematic review. *Psychotherapy Research*, *30*(1), 79–96. https://doi.org/10.1080/10503307.2019.1566676

Röhrig, B., Witthöft, M., Herpertz, S., & Petrak, F. (2024). “The early bird catches the worm”—Prediction of early response and its importance for treatment outcome in a naturalistic setting of cognitive behaviour therapy. *Clinical Psychology & Psychotherapy, 31*(1) https://doi.org/10.1002/cpp.2956

Rubel, J., Lutz, W., & Schulte, D. (2015). Patterns of change in different phases of outpatient psychotherapy: A stage-sequential pattern analysis of change in session reports. *Clinical Psychology & Psychotherapy*, *22*(1), 1–14. https://doi.org/10.1002/cpp.1868

Saxon, D., Broglia, E., Duncan, C., Barkham, M. (2024). Variability in treatment effects in an English national dataset of psychological therapies: The relationships between severity, treatment duration, and therapy type. *Journal of Affective Disorders, 362*, 244-255. https://doi.org/10.1016/j.jad.2024.06.115

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder. *Archives of Internal Medicine*, *166*(10), 1092 - 1097. https://doi.org/10.1001/archinte.166.10.1092

StataCorp. (2023). *Stata 18 Structural Equation Modeling Reference Manual*. Stata Press.

Stulz, N., Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of early change in psychotherapy under routine outpatient conditions. *Journal of Consulting and Clinical Psychology*, *75*(6), 864–874. https://doi.org/10.1037/0022-006X.75.6.864

Tolin, D. F. (2010). Is cognitive–behavioral therapy more effective than other therapies?: A meta-analytic review. *Clinical Psychology Review*, *30*(6), 710–720. https://doi.org/10.1016/j.cpr.2010.05.003

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H.-n. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes." *Psychological Bulletin, 122*(3), 203–215. https://doi.org/10.1037/0033-2909.122.3.203

Watson, J. C. (2021). Responsiveness in emotion-focused therapy. In J. C. Watson & H. Wiseman (Eds.), *The responsive psychotherapist: Attuning to clients in the moment*. APA. https://doi.org/10.1037/0000240-009

Wojnarowski, C., Firth, N., Finegan, M., & Delgadillo, J. (2019). Predictors of depression relapse and recurrence after cognitive behavioural therapy: a systematic review and meta-analysis. *Behavioural and Cognitive Psychotherapy*, *47*(5), 514-529. https://doi.org/10.1017/S1352465819000080

Table 1

Patient Demographics

Demographics	**Full Sample (*n* = 274)**	**PCET (*n*= 149)**	**CBT (*n* = 125)**	**Test Statistic**	***p***	**RCSI (*n* = 155)**	**NonRCSI (*n* = 119)**	**Test Statistic**	***p***
Age (M, SD)	39.22 (12.81)	40.01 (13.27)	38.26 (12.21)	$t(272) = -1.13$	.261	39.68 (12.71)	38.61 (12.96)	$t(272) = -0.69$	.490
Gender (%)									
Female	56.9	54.4	60.0	$\chi^2(1) = 0.88$	.348	64.5	47.1	$\chi^2(1) = 8.37$	**.004**
Male	43.1	45.6	40.0			35.5	52.9		
Ethnicity (%)									
White British	86.9	83.9	90.4	$\chi^2(5) = 8.11$	.150	83.9	90.8	$\chi^2(5) = 5.30$	.380
Asian/ Asian British	1.1	1.3	0.8			1.3	0.8		
Black/ Black British	1.5	2.7	-			2.6	-		
Mixed Ethnicity	2.9	2.0	4.0			3.9	1.7		
White Other	2.6	2.7	2.4			3.2	1.7		
Not asked	5.1	7.4	2.4			5.2	5.0		
IMD (M, SD)	5.48 (3.27)	5.39 (3.29)	5.60 (3.27)	$t(271) = 0.54$	.590	5.72 (3.33)	5.17 (3.18)	$t(271) = -1.39$	.167
Employment Status (%)									
Employed	56.9	57.0	56.8	$\chi^2(7) = 1.83$	.969	59.4	53.8	$\chi^2(7) = 17.63$	**.014**
Unemployed Seeking Work	8.4	8.1	8.8			5.8	11.8		
Homemaker	2.2	2.7	1.6			1.9	2.5		
Sick or Disabled	9.9	10.1	9.6			5.8	15.1		
Student	5.8	5.4	6.4			7.7	3.4		
Retired	2.2	1.3	3.2			3.9	-		
Not seeking work	0.7	0.7	0.8			1.3	-		
Missing	13.9	14.8	12.8			14.2	13.4		
Psychotropic Status (%)									
Prescribed and taking	56.2	52.3	60.8	$\chi^2(3) = 4.88$	.181	53.5	59.7	$\chi^2(3) = 7.37$	.061
Not prescribed	34.3	34.9	33.6			38.7	28.6		
Prescribed, not taking	3.6	5.4	1.6			4.5	2.5		
Patient Unsure	5.8	7.4	4.0			3.2	9.2		

Abbreviations. CBT = Cognitive Behavioral Therapy; IMD = Index of Multiple Deprivation; PCET = Person Centered Experiential Therapy

Table 2

Patient Clinical Characteristics

		Treatment Group				Outcome Group			
Clinical Characteristics	Full Sample (n = 274)	PCET (n= 149)	CBT (n = 125)	Test Statistic	p	RCSI (n = 155)	NonRCSI (n = 119)	Test Statistic	p
Session 1, PHQ-9 M (SD)	18.03 (4.11)	18.06 (4.31)	17.99 (3.86)	$t(257) = -0.13$	.900	17.88 (3.80)	18.21 (4.49)	$t(216)^{\dagger} = 0.63$	.532
Session 1, GAD-7 M (SD)	13.56 (4.50)	13.54 (4.86)	13.58 (4.07)	$t(251) = 0.07$	.948	13.41 (4.31)	13.76 (4.75)	$t(251) = 0.61$	.543
Session 1, WSAS M (SD)	24.48 (6.93)	24.12 (7.04)	24.89 (6.82)	$t(227) = 0.84$	.401	23.47 (6.92)	25.81 (6.76)	$t(227) = 2.56$	**.011**
Session 1-4 PHQ-9 change M (SD)	3.08 (4.90)	3.26 (4.93)	2.85 (4.90)	$t(248) = -0.66$	.513	4.20 (5.16)	1.61 (4.13)	$t(247)^{\dagger} = -4.28$	**<.001**
Number of sessions M (SD) Range	11.93 (5.20) 5 – 28	11.36 (5.23) 5 – 28	12.60 (5.10) 5 – 24	$t(272) = 1.97$	**.049**	11.88 (4.90) 5 - 21	11.99 (5.59) 5 – 24	$t(272) = 0.18$	.857
Attrition (%)									
Did not drop out / decline	80.9	85.1	75.8	$\chi2(1) = 3.78$	.051	92.9	65.3	$\chi^2(1) = 32.92$	**<.001**
Dropped out/ declined	19.1	14.9	24.2			7.1	34.7		
Therapy Group		-	-	-	-				
PCET	149					77	72	$\chi^2(1) = 3.18$	.075
CBT	125					78	47		

Abbreviations. CBT = Cognitive Behavioural Therapy; PCET = Person Centred Experiential Therapy; GAD-7 = Generalised Anxiety Disorder Measure; PHQ-9 = Patient Health Questionnaire; RCSI = Reliable and Clinically Significant Improvement; WSAS = Work and Social Adjustment Scale. † Equal variance not assumed

Table 3

Model Fit and Model Comparison Analyses for Multiple Group Growth Curve Model across Treatment Conditions

			Fit				
Constraints	$\chi^2(df)$	p	**CFI**	**TLI**	**RMSEA**	**AIC**	**BIC**
1. Unconstrained	30.35(10)	.001	0.966	0.959	0.122	5776.124	5841.16
2. Means	32.01(12)	.001	0.966	0.966	0.111	5773.779	5831.589
3. Means, Variances	32.88(15)	.005	0.970	0.976	0.093	5768.651	5815.622
4. Means, Variances, Errors	34.43(19)	.016	0.974	0.984	0.077	5762.204	5794.722

Model Comparison		
	$\chi^2(df)$	p
1 vs 2	1.655(2)	.437
2 vs 3	0.872(2)	.832
3 vs 4	1.552(4)	.817

Note. CFI = Comparative Fit Index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square Error of Approximation, AIC = Akaike Information Criterion, BIC = Bayesian Information Criterion

Table 4

Model fit and Model Comparison Analyses for Multiple Group Growth Curve Model predicting last session RCSI category across Treatment Conditions

		Fit	
Constraints	LL(df)	AIC	BIC
1. Unconstrained	-3039.239(15)	6108.478	6162.675
2. Constant Constrained	-3039.417(14)	6106.834	6157.418
3. Slope & Constant	-3043.097(13)	6112.194	6159.165
4. Intercept, Slope, & Constant	-3045.825(12)	6155.65	6159.008
		Model Comparison	
	$\boldsymbol{\chi^2}$**(df)**	***p***	
1 vs 2	0.36(1)	.5509	
2 vs 3	7.36(1)	**.0067**	
3 vs 4	5.46(1)	**.0195**	

Note. LL = log-likelihood; AIC = Akaike Information Criterion, BIC = Bayesian Information Criterion; RCSI = Reliable and Clinically Significant Improvement for the Patient Health Questionnaire-9

Table 5

Coefficients for best fitting model for last session PHQ-9 and RCSI analysis

		Last PHQ-9			RCSI		
	Coefficient	**Est.**	**Upper Limit**	**Lower Limit**	**Est.**	**Upper Limit**	**Lower Limit**
Growth Model	Intercept Mean (μ_1, μ_3)	17.998**	17.5	18.5	18.003**	17.502	18.505
	Slope Mean (μ_2, μ_4)	1.030**	-1.224	-0.837	-1.033**	-1.228	-0.839
	Intercept Variance (σ_1^2, σ_3^2)	14.467	11.717	17.863	14.609	11.862	17.992
	Slope Variance (σ_2^2, σ_4^2)	1.474	0.999	2.173	1.508	1.031	2.208
	Covariance (σ_{13}, σ_{24})	-0.544	-1.575	0.487	-0.633	-1.656	0.390
Outcome Model	**CBT**						
	Constant	0.978	-2.703	4.66	2.458$_a$	0.493	12.248
	Random Intercept (β_1)	0.531**	0.335	0.727	0.963$_a$	0.884	1.049
	Random Slope (β_3)	1.993**	1.225	2.760	0.734$_a$	0.479	1.125
	PCET						
	Constant	0.978	-2.703	4.66	2.458$_a$	0.493	12.248
	Random Intercept (β_2)	0.623**	0.429	0.816	0.882*$_a$	0.804	0.967
	Random Slope (β_4)	1.993**	1.225	2.760	0.258*$_a$	0.111	0.601

Note. ** p < 0.01, ***p <0.001, a = exponentiated coefficient (β_1, β_2, β_3, β_4 are odds-ratios in the RCSI model); CBT = Cognitive Behavioral Therapy; PCET = Person-centered Experiential Therapy; last PHQ = Last Session PHQ-9 Score; RCSI = Reliable and Clinically Significant Improvement on the Patient Health Questionnaire-9.

Table 6

Model Fit and Model Comparison Analyses for Multiple Group Growth Curve Model predicting last session PHQ-9 across conditions.

			Fit				
Constraints	$\boldsymbol{\chi^2(df)}$	**p**	**CFI**	**TLI**	**RMSEA**	**AIC**	**BIC**
1. Unconstrained	35.953(23)	.042	0.980	0.983	0.064	7407.528	7468.951
2. Constant Constrained	36.741(24)	.046	0.981	0.984	0.062	7406.316	7464.126
3. Slope & Constant	39.134(25)	.036	0.978	0.983	0.064	7406.709	7460.906
4.Intercept, Slope, & Constant	45.168(26)	.011	0.971	0.977	0.073	7410.743	7461.327
	Model Comparison						
	$\boldsymbol{\chi^2(df)}$	$\boldsymbol{p}$					
1 vs 2	0.79(1)	.3748					
2 vs 3	2.39(1)	.1219					
3 vs 4	6.03(1)	**.0140**					

Note. CFI = Comparative Fit Index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square Error of Approximation, AIC = Akaike Information Criterion, BIC = Bayesian Information Criterion; PHQ-9 = Patient Health Questionnaire-

Figure 1

Multiple-Group Latent Growth Curve Model predicting Reliable and Clinically Significant Improvement (RCSI).

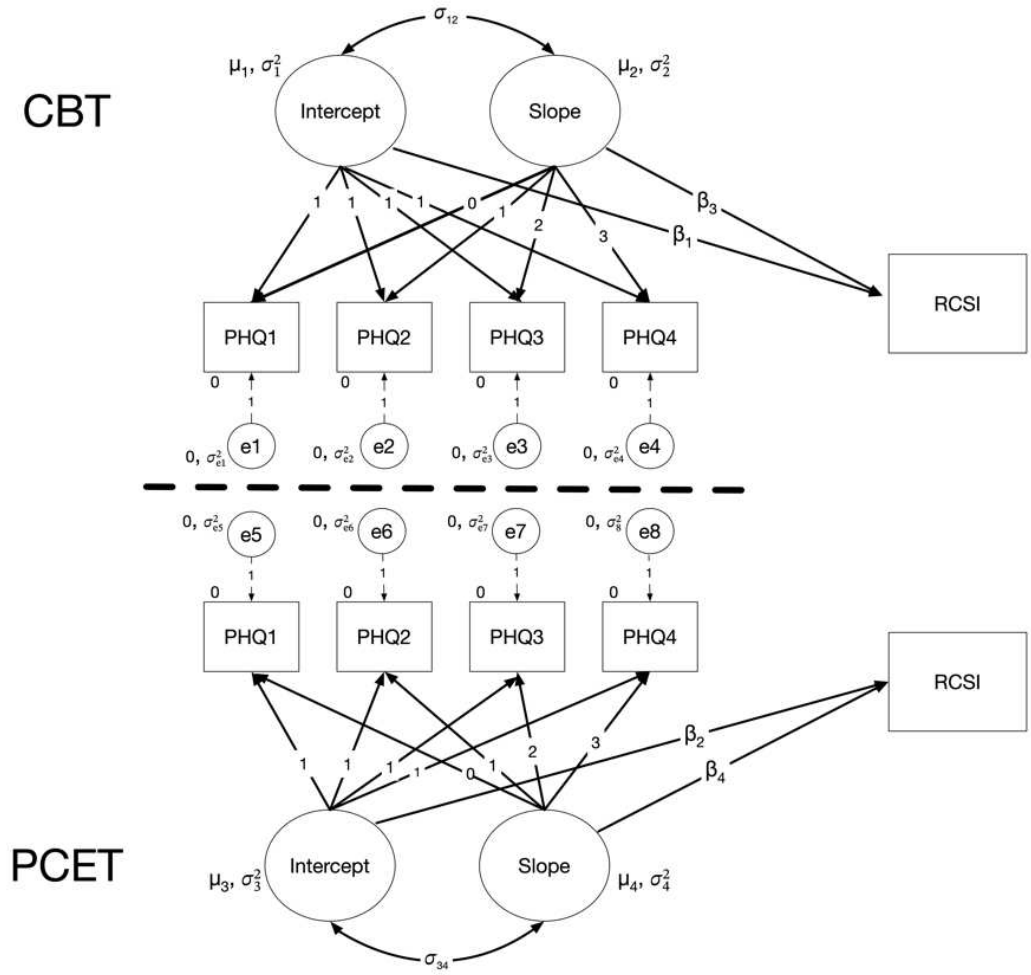

Figure 2

Predicted Probability of Reliable and Clinically Significant Improvement (RCSI) to PCET and CBT as a function of improvement (in PHQ-9 score) during sessions 1–4

Figure 2 Color (for the print)

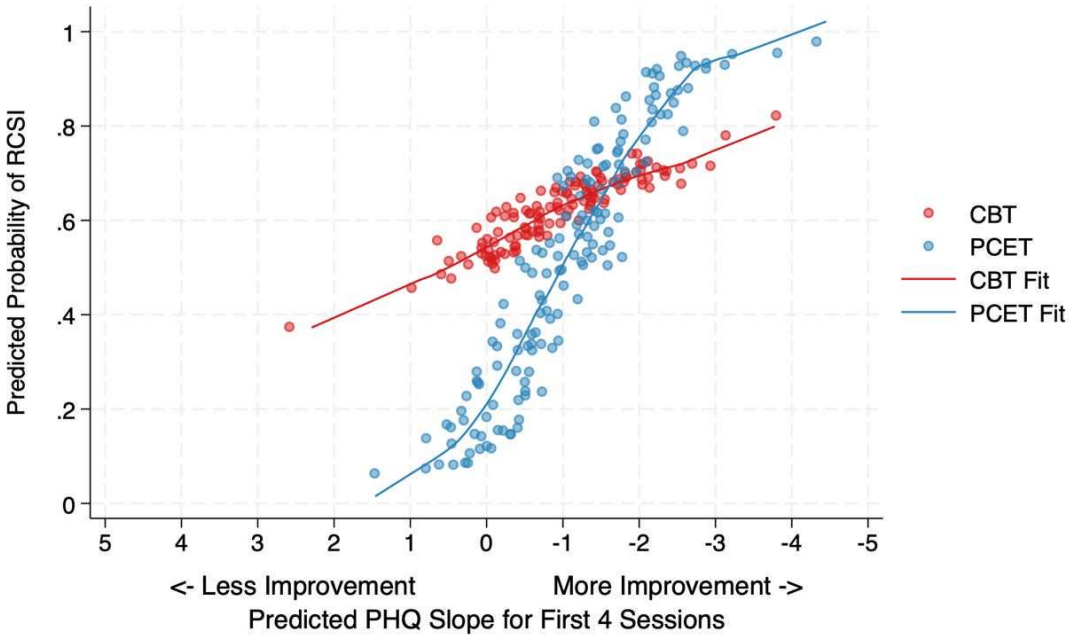

Figure 2 Greyscale (for online).