Early response as a prognostic indicator in person-centered experiential therapy for depression

Kate E M Duffy[1]

Melanie Simmonds-Buckley[2]

David Saxon[2]

Jaime Delgadillo[2]

Michael Barkham[2]

[1] Department of Psychology, University of Sheffield, UK

[2] Clinical and Applied Psychology Unit, Department of Psychology, University of Sheffield

Correspondence concerning this article should be addressed to Melanie Simmonds-Buckley, Clinical and Applied Psychology Unit, Department of Psychology, University of Sheffield, Cathedral Court, 1 Vicar Lane, Sheffield, S1 2LT, United Kingdom. Email: m.simmonds-buckley@sheffield.ac.uk

Abstract

Objectives: Currently, no reports exist on the phenomenon of early response in humanistic-experiential therapies. This study investigated the prognostic value of early response on post-treatment outcomes in person-centered experiential therapy (PCET) for depression within the English Improving Access to Psychological Therapies (IAPT) program.

Design: Retrospective observational cohort study.

Method: Routine clinical data were drawn from N=3321 patients with depression symptoms. The primary outcome was reliable and clinically significant improvement (RCSI) on the Patient Health Questionnaire-9 (PHQ-9) self-report depression measure at the end of treatment. Early response was operationalized as reliable improvement, defined as a PHQ-9 change score ≥6 from baseline to session four. Early response was examined as a predictor of RCSI using logistic regression controlling for baseline depression severity. In sensitivity analyses, therapist effects were controlled using multi-level modeling.

Results: A total of 38.7% of patients met the criterion for early response. Patients who experienced an early response to treatment were six times more likely to recover at the end of treatment compared to patients who did not have an early response. The early response effect was still evident after accounting for individual variability between therapists. However, a quarter of patients displayed a pattern of eventual response, reaching recovery at end of treatment despite not experiencing an initial improvement early in therapy.

Conclusion: Early response to PCET is a reliable predictor of treatment outcome. Different response patterns evidenced in this study indicate that identifying subgroups of patients associated with early and eventual response could support clinical decision making.

Keywords: early response, depression, person-centered experiential therapy, IAPT, therapist effects.

Public significance statement:

This study extends the investigation of patterns of early response into person-centered experiential therapy. Early response was a strong predictor of better post-treatment outcomes in depression. These findings are important, especially for practitioners, in understanding patterns of patient improvement and deterioration and also to provide further clinical support where necessary.

The cumulative body of research on the psychological therapies attests to the view that client outcomes are broadly similar or, when not, slightly favor cognitive behavioral therapies (Barkham & Lambert, 2021). While outcome differences between psychological therapies may be small, there is acknowledgement that technical diversity exists between differing theoretical models in terms of their content and processes (e.g., Stiles et al., 1986). For example, in addition to processes and content that are common across different models of therapy, modality-specific competencies and techniques have been identified regarding a range of diverse therapies including cognitive behavioral therapy (CBT; Blagys & Hilsenroth, 2002; Roth & Pilling, 2007; Taylor et al., 2020), as well as other leading therapeutic approaches such as psychodynamic/psychoanalytic therapies (Lemma et al., 2008) and humanistic-experiential therapies (Roth et al., 2009).

One explanation for the broad equivalence of outcomes is that change is evaluated at the level of group means, thereby masking differential effects over the course of therapy. Using session-by-session data over the course of therapy, a recent report comprising a large sample of clients receiving either CBT or person-centered experiential therapy – a form of humanistic-experiential therapy – found outcomes for moderate or severe depression in earlier sessions to favor person-centered experiential therapy while later sessions favored CBT (Pybis et al., 2017). Such a finding may be indicative of a particular propensity of person-centered experiential therapy to yield a robust early response. Some researchers have argued that early response is a result of the successful sequencing of specific theoretical techniques, particularly in CBT (e.g., Fennell & Teasdale, 1987), although such accounts within CBT have been challenged (see Ilardi & Craighead, 1994).

In contrast to CBT, in humanistic-experiential therapies the therapeutic relationship is viewed as both central and potentially curative with the direction of therapy being client-led and focusing on emotional processing of the client's experiencing (Elliott et al., 2021;

Murphy, 2019; Watson, 2018). The process is premised on validating the client's experiences, emotions, and them as a person. Accordingly, early sessions of humanistic-experiential therapies differ technically from those in CBT. Although some researchers have argued that early response is a function of common factors (e.g., Lambert, 2005), there is sufficient distinction in the content and processes of CBT and humanistic-experiential therapies to support the notion of technical diversity in early sessions of therapy.

A recent systematic review and meta-analysis of early response to psychological therapy as a predictor of depression and anxiety outcomes showed a positive relationship between early response and final treatment outcomes yielding an odds ratio (OR) of 4.84, indicating that early responders were close to five times as likely to have a favorable post-treatment outcome contrasted with non-early responders (Beard & Delgadillo, 2019). However, of the 25 studies included in the review, the majority of studies with depression as the primary disorder utilized CBT-based interventions. Only two of the 25 studies in their meta-analysis focused on or included a non-CBT intervention (Crits-Christoph et al., 2001; Van et al., 2008). In addition, only nine studies were observational (i.e., not RCTs), with three being university-based clinics, five studies comprising CBT-based interventions, and the remaining study denoted as 'psychotherapy'.

Hence, while the phenomenon of early response has been shown to present in CBT and also in a small number of studies of psychodynamic therapies, no study has examined the phenomenon of early response in humanistic-experiential therapies. The one exception is a study of adolescents receiving non-directive supportive therapy and showing rapid response (i.e., by session two) who were found to yield superior outcomes at one-year (Renaud et al., 1998). There is, therefore, an evidence gap concerning whether the phenomenon of early response occurs in humanistic-experiential therapies in adult clinical samples. The aim of the present study was to investigate the prognostic value of early response on post-treatment

outcomes in patients receiving person-centered experiential therapy (PCET), which has been found to be broadly equivalent to CBT at the close of therapy (Barkham et al., 2021). Importantly, sudden gains are a related but distinct phenomenon as they denote statistically significant and rapid improvement between two successive therapy sessions (Tang & DeRubeis, 1999) as opposed to early response which occurs over the course of a number of sessions in the initial phase of therapy.

In terms of methodology, the disparity between the definitions of early response raises the issue of differential criteria used in research studies relating to (a) the time frame of early response (i.e., the number of sessions), and (b) the method used to determine the extent of change. In terms of the former, researchers have used various numbers of sessions: two-session has been used by Renaud et al. (1998), three-sessions by Haas et al. (2002) because the authors "felt that an early response would be identifiable within this time frame" (p. 1161), and also by Rubel et al. (2015) because three sessions was the minimum data requirement for growth mixture modeling to yield a log-linear trend. However, a majority of studies (60%) in a systematic review of early response used four-sessions (Beard & Delgadillo, 2019). In addition, a treatment duration of four sessions has also been identified as the minimally acceptable dose of therapy in a systematic review of dose-effect relations (Robinson et al., 2020). Given its majority use and association with a minimum therapy dose, combined with our focus on time-limited therapies, we adopted four sessions as the time frame for early response.

In terms of the extent of change required to denote early change, three candidate criteria include: (1) a specified degree of response, for example, 25% or 50% reduction in baseline score; (2) meeting the threshold for remission, which is usually based on the cut-off score indicating clinical significance; or (3) the criterion for reliable change, in which the extent of improvement since baseline exceeds the measurement error of the instrument being

used, thereby providing confidence that the change is not due to chance or natural fluctuation; that is, it is a reliable index of change (Jacobson & Truax, 1991). Considering each criteria in turn: (1) a specified percentage of improvement in relation to baseline places the emphasis on the intake (baseline) score with greater change required for patients presenting with a higher baseline score, with no clear rationale existing for the selection of the percentage reduction; (2) a criterion of meeting remission is equivalent to the final aim of therapy and, therefore, would appear to be equating early response with the end-point score; and (3) a criterion of reliable change is premised on the psychometric properties of the outcome measure of interest, with the extent of change excluding the possibility of measurement error accounting for the change in score from baseline. Considering each of these criteria, reliable change is the only one that takes account of the psychometric properties of a measure whilst not confounding it with the final outcome. It therefore has specific properties that make it a credible criterion for determining early response.

In determining the clinical context for the present study, we heeded the call that the phenomenon of early response needs to be investigated in routine practice (Schlagert & Hiller, 2017). Hence, the study comprised a large routine psychological service setting within the English Improving Access to Psychological Therapies (IAPT; Clark, 2018) program, similar to the clinical setting used in the Pybis et al. (2017) study. However, in such settings, there is evidence of a significant therapist effect (i.e., variability between the outcomes of therapists) accounting for approximately 5-7% of patient outcomes (Baldwin & Imel., 2013; Johns et al., 2019). As yet, no studies have investigated the influence of therapist effects on early response. Given the naturally clustered structure of routine practice data sets whereby patients are nested within therapists, failure to consider clustered data could potentially impact on the interpretation of results. Hence, the current study controlled for therapist effects.

The current study addresses two research questions: (1) Does early response in routine practice predict post-treatment outcome in PCET whilst controlling for baseline scores, on the criterion measure?, and (2) does the predictive value of early response in routine practice remain the same in magnitude when controlling for therapist effects? It was hypothesised that early response would predict post-treatment depression outcomes and that the predictive value of early response would remain similar after controlling for therapist effects.

Method

Design

This was a retrospective observational cohort study. Both the primary predictor and outcome variables were binary; the predictor variable specified as reliable improvement at session four, and the outcome variable specified as both reliable and clinically significant improvement as measured by a standard patient self-report outcome measure of depression. Both these variables are detailed later in this section.

Service setting

The study focused on an analysis of secondary routinely collected data from patients accessing a single service located in the North West of England within the Improving Access to Psychological Therapies (IAPT) program (Clark, 2018). The IAPT program is a stepped care model in which patients are either first offered the least intensive treatment (low intensity at step 2) that will likely yield a positive outcome and stepped up (to a high-intensity treatment at step 3) if improvement does not occur relatively quickly, or referred directly to a high-intensity treatment on the basis of severity or specific presenting condition (e.g., PTSD).

Most patients initially receive a low-intensity intervention in the form of a range of psycho-educational interventions from Psychological Wellbeing Practitioners and are only

stepped up to a high-intensity treatment if they do not show improvement. High intensity interventions comprise a number of IAPT-approved psychological therapies of which the two main modalities are cognitive-behavioral therapy (CBT) delivered by approved CBT therapists and PCET delivered by approved counselors. Major hallmarks of the IAPT program are its weekly supervision routine for CBT therapists and PCET counselors and the completion of routinely administered outcome measures by patients at each attended therapy session used to determine their progress and to aid decision making in either stepping up or discharge.

Data set and sample selection

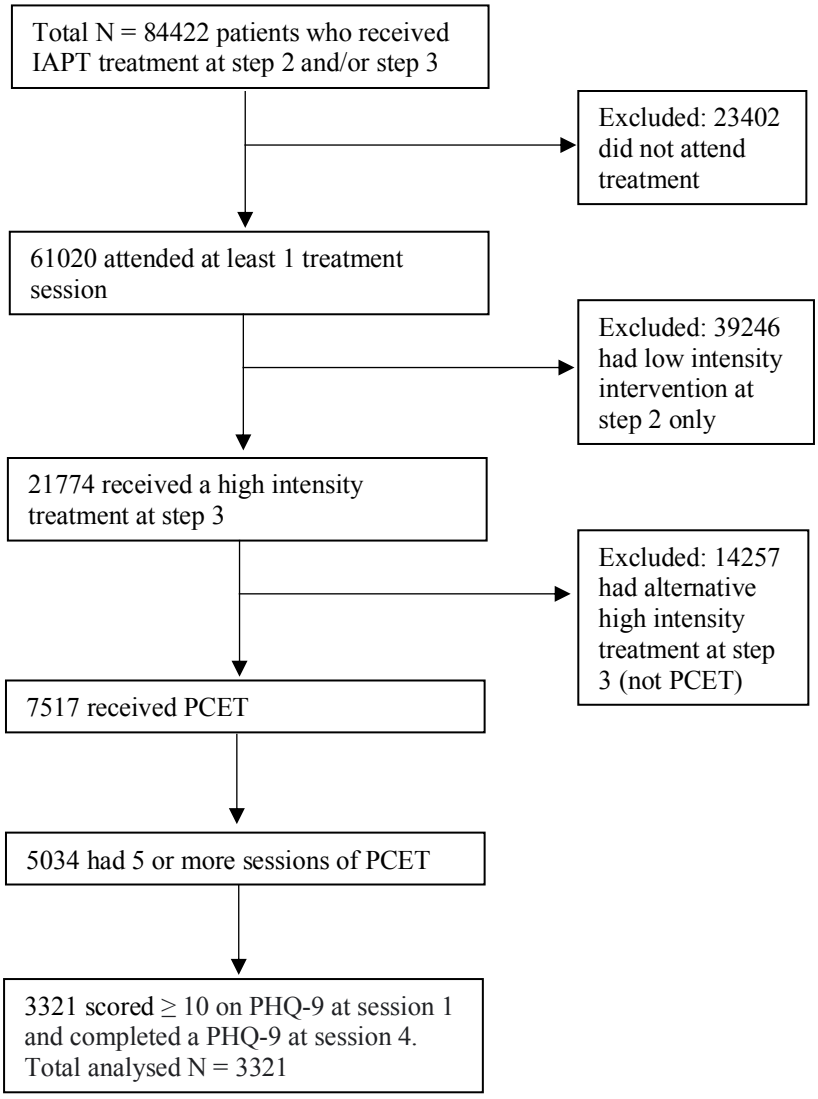
Data were drawn from an archival and pseudo anonymized routine practice data set collected under the auspices of a larger program of research funded by the UK National Institute for Health Research that collected data from nine IAPT services. Ethics approval was given by the Department of Psychology Research Ethics Committee at the University of Sheffield (13th May 2021, Ref: 040038).

The current project utilized data from a single service in the dataset, that was selected by the data custodian [initials masked], on the basis of including a sizable proportion of patients who accessed PCET. The data set had not previously been analysed and hence no other information was used to determine the selection of this data set. The single service data were filtered from a total N = 84,422 patients inclusive of IAPT low- and high-intensity interventions. The data were processed and cases that met any one of the following criteria were excluded: did not attend; only received step 2 low-intensity interventions (guided self-help), cases who had alternative step 3 high-intensity treatments to PCET (such as CBT); cases that had less than five sessions of PCET; those who scored < 10 on the PHQ-9 at session one; and those who did not complete a PHQ-9 measure at session four. Figure 1 displays a data flow diagram summarizing the filtering process. On the basis that early

response is defined as occurring during the initial four therapy sessions, the analysed data set comprised only cases that received a minimum of five therapy sessions.

Figure 1

Flow diagram illustrating data filtering process



Measures

As mandated in the IAPT program, patients routinely completed a number of standard outcome measures at each attended therapy. As the focus for the present study was on response in depression, only data from the depression measure were utilized.

The Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001).

The PHQ-9 is a self-administered measure of depression comprising nine items mapping onto the nine symptoms identified in the DSM-5 diagnosis of depression (e.g., poor appetite, poor concentration, trouble sleeping). The questionnaire measures the frequency of symptoms during the past two weeks using anchor points of 0 (*not at all*) to 3 (*nearly every day*) yielding a score ranging from 0 to 27. Reports of the PHQ-9 have identified severity bands as follows: mild (5-9); moderate (10-14), moderately-severe (15-19), and severe (20-27). The literature reports internal consistency for the PHQ-9 to be $a = .89$ (PHQ-9; Kroenke et al., 2001).

Study sample

The study sample included patients receiving high-intensity PCET at step 3 across a period of 42 months from June 2014 to November 2017 and comprised $N = 3321$ patients with a mean age at referral of 43.72 ($SD = 14.03$, range = 18 to 88 years). The study sample only included cases with a baseline PHQ-9 score within the clinical range (≥10) and those with a minimum of five sessions. Table 1 presents comparisons of descriptive characteristics of early responders, non-early responders, and the total sample. Non-early responders were significantly more likely to be unemployed, taking medication, have lower baseline depression severity but higher impaired functioning at baseline.

Intervention

The format of humanistic-experiential therapy in the planned study was PCET, one of the IAPT-approved high intensity interventions that was developed specifically for the treatment of depression at step 3 (Sanders & Hill, 2014). It was originally termed *Counselling for Depression* but has subsequently been referred to as person-centred experiential therapy in order to better reflect the theoretical origins of the intervention (Barkham et al., 2021). It originated in response to the 2009 UK National Institute for Health and Care Excellence (NICE) guideline for depression that identified a paucity of research evidence for humanistic

therapies, leading to the development of a research-informed model of traditional person-centred therapy enhanced with components drawn from emotion focused therapy (Elliott et al., 2004; see also Elliott et al., 2021). As a result, its distinctive features include the therapist being more active than in traditional person-centred therapy, particularly in being more process-guiding but not directive. Hence, in contrast to traditional person-centred therapy, it is best described as process-guiding rather than non-directive. NICE guidelines state that the model as delivered within the English IAPT program can have a maximum of 20 individual sessions although reports indicate that the average number of sessions administered is approximately seven (e.g., Barkham et al., 2021). The data set comprised a total of 307 therapists, but no information was available on the therapists.

Statistical analysis

Descriptive statistics comprised means (*Ms*) and standard deviations (*SDs*) of PHQ-9 scores presented for baseline, session four, and final session. Pre-post treatment (uncontrolled) effect sizes (ESs) were calculated to assess the general effectiveness of PCET in the current sample. There were no missing data in the primary analysis because the data were initially filtered (as seen in Figure 1) on the basis that every participant had three scores: a PHQ-9 score at session one, session four, and final session. However, some of the demographic variables contained missing data as indicated by the differing denominators in Table 1. Demographics were calculated based on available data with the proviso that percentages excluded cases with missing data.

The main predictor variable, early response, was computed using the reliable improvement (RI) criterion (Jacobson & Truax, 1991) at session four, defined as a PHQ-9 change score ≥ 6, based on the reliable change index recommended by the IAPT program (National Collaborating Centre for Mental Health, 2018). A categorical variable was

computed, where participants who met this early response criterion were coded 1 and those who did not show early response were coded 0.

Table 1

Sample characteristics of early responders, non-early responders, and all cases at entry to treatment

	Early responder cases (n = 1285)	**Non-early responder cases (n = 2036)**	**All cases (n = 3321)**	**Difference between groups *p*[a]**
Demographics				
Gender				.254
Male	364/1281 (28.4%)	614/2031 (30.2%)	978/3312 (29.5%)	
Female	915/1281 (71.4%)	1411/2031 (69.5%)	2326/3312 (70.2%)	
Age (SD)	43.81 (14.61)	43.67 (13.66)	43.72 (14.03)	.786
Ethnicity				.507
White British	1073/1285 (83.5%)	1682/2036 (82.6%)	2755/3321 (83.0%)	
Other	212/1285 (16.5%)	354/2036 (17.4%)	566/3321 (17.0%)	
Employment status				**.011**
Employed or other	879//1122 (78.3%)	1300/1760 (73.9%)	2179/2882 (75.6%)	
Unemployed or long-term sick/disabled	243/1122 (21.7%)	460/1760 (26.1%)	703/2882 (24.4%)	
Psychotropic Medication				**.002**
Taking	694/1110 (62.5%)	1222/1773 (68.9%)	1916/2883 (66.5%)	
Not taking	416/1110 (37.5%)	551/1773 (31.1%)	967/2883 (33.5%)	
Treatment format				.788
Face to face	1120/1285 (87.2%)	1768/2036 (86.8%)	2888/3321 (87.0%)	
Telephone	165/1285(12.8%)	268/2036 (13.2%)	433/3321 (13.0%)	
Treatment modality				**.036**
One-to-one	1119/1120 (99.9%)	1757/1768 (99.4%)	2876/2888 (99.6%)	
Couples	1/1120 (0.1%)	11/1768 (0.6%)	12/2888 (0.4%)	
Clinical characteristics				
Baseline PHQ-9 score (SD)	18.33 (4.35)	17.38 (4.71)	17.75 (4.60)	**<.001**
Baseline GAD-7 score (SD)	15.11 (4.30)	15.0 (4.40)	15.04 (4.36)	.469
Baseline WSAS score (SD)	20.64 (9.63)	22.19 (9.34)	21.58 (9.48)	**<.001**
Mean number of sessions (SD)	6.94 (2.01)	7.52 (2.34)	7.30 (2.23)	**<.001**

Note: Percentages exclude cases with missing data; Abbreviations: GAD-7, General Anxiety Disorder-7 measure; PHQ-9, Patient Health Questionnaire-9 measure of depression; WSAS, Work and Social Adjustment Scale; SD, standard deviation. [a] Significance test based on Chi-squared (or Fishers exact test if cell counts less than 5) for categorical variables and independent samples T-test for continuous variables.

The primary outcome variable (i.e., reliable and clinically significant improvement; RCSI) was computed for each participant following the method proposed by Jacobson and Truax (1991). Participants were defined as having met RCSI if the following rules were met: (1) baseline PHQ-9 scores were within the clinical range ≥ 10; (2) the pre-post treatment PHQ-9 change score ≥ 6; and (3) the final PHQ-9 score < 10 (clinical cut-off). Participants were coded as 1 if they met RCSI criteria and 0 if not. We examined the cross-tabulations between the early response (signal) and RCSI (outcome) by partitioning the sample across four categories: (1) *true positives*, (2) *false positives*, (3) *false negatives*, and (4) *true negatives*. Mean scores in the PHQ-9 at three time-points (baseline, session 4, final session) were visualised to examine differences between these groups.

Data analysis comprised two stages: stage one applied a single-level logistic regression model and stage two used a multi-level model to control for therapist effects (variability attributable to therapists).

Stage 1: A logistic regression model was developed using SPSS (version 25) to investigate whether patients with an early response at session four (independent variable) were more likely to attain post-treatment RCSI (dependent variable), controlling for baseline severity of depression (PHQ-9). Odds ratios (OR) and 95% confidence intervals (CI) were computed to determine the effect size of the predictors in the model.

Stage 2: A multilevel logistic regression model was developed using MLwiN software to examine whether controlling for therapist effects would change the predictive value of early response on post-treatment RCSI, using the same independent and dependent variables described above. A random intercept model was developed using second-order penalized quasi-likelihood (PQL) estimation method as opposed to marginal quasi-likelihood (MQL) because PQL is a less biased method of estimation for binary outcomes (Rodriguez & Goldman, 2001). The intercept was set to vary randomly by therapists (level 2) in order to

control for therapist effects. An additional model allowing random slopes for therapists (level 2) was also tested.

A post hoc sensitivity analyses was conducted on the primary logistic regression analysis in order to compare the odds ratio for shorter versus longer duration PCET interventions. Using the mean number of sessions (7.3), the sample was split into two subgroups; shorter treatments (<8 sessions) and longer treatments (>=8 sessions). This enabled assessment of whether the effect of early response is merely a product of the classification rules. Finally, post hoc tests were carried out to determine the level of sensitivity, specificity, positive and negative predictive values for the prediction model.

Transparency and Openness

The study design, hypotheses and analysis plan were pre-registered on aspredicted.org on 30th July 2021 (See https://aspredicted.org/blind.php?x=DSC_Z21). Reporting standards are followed throughout, including clear reports of the dataset, sample selection, any exclusions and measures used. Data were analysed using SPSS version 25 (IBM Corp, 2017) and MLwiN (Charlton et al., 2020) software. The data and analysis code for this study are not available.

Results

In terms of general effectiveness of PCET, for the total sample of patients, the pre-post effect size, (ES) was $d = 1.75$. When calculated separately for early and non-early responders, the ESs were $d = 2.66$ and $d = 1.24$, respectively. Overall, 53.5% (n = 1776) of the whole sample achieved RCSI by the end of treatment, and 1.5% (n = 49) met threshold for reliable deterioration. Of the whole sample, 38.7% (n = 1285) met the criterion for early response at session four and 24.4% (n = 745) made both RI and RCSI by session four. Table 2 displays means (*Ms*) and standard deviations (*SDs*) for the PHQ-9 at baseline, session four, and final scores for early and non-early responders and for the whole sample. At baseline, the

numbers of participants in each of PHQ-9 severity band were as follows: 10–14, n = 942 (28.4%); 15–19, n = 1143 (34.4%); and 20–27, n = 1236 (37.2%).

Table 2

Means and standard deviations for PHQ-9 baseline and final scores across study sample

	Early responders (n = 1285)		Non-early responders (n = 2036)		Total sample (N = 3321)	
PHQ-9 score	*M*	*SD*	*M*	*SD*	*M*	*SD*
Baseline	18.33	4.35	17.38	4.71	17.75	4.60
Session 4	8.79	4.44	16.43	5.25	13.48	6.20
Final score	6.78	5.23	11.52	6.86	9.69	6.69

Note. PHQ-9 = Patient Health Questionnsire-9.

Table 3 shows the number of patients who met or did not meet the criteria for early response and RCSI. Table 3 shows four groups of patients: (1) early responders who attained RCSI [*true positive signal*], (2) early responders who did not attain RCSI [*false positive signal*], (3) those with no early response but who eventually attained RCSI [*false negative signal*], and (4) those neither met early response nor RCSI criteria [*true negative signal*].

Table 3

Cross-tabulation detailing the number of participants who met early response and RCSI criteria

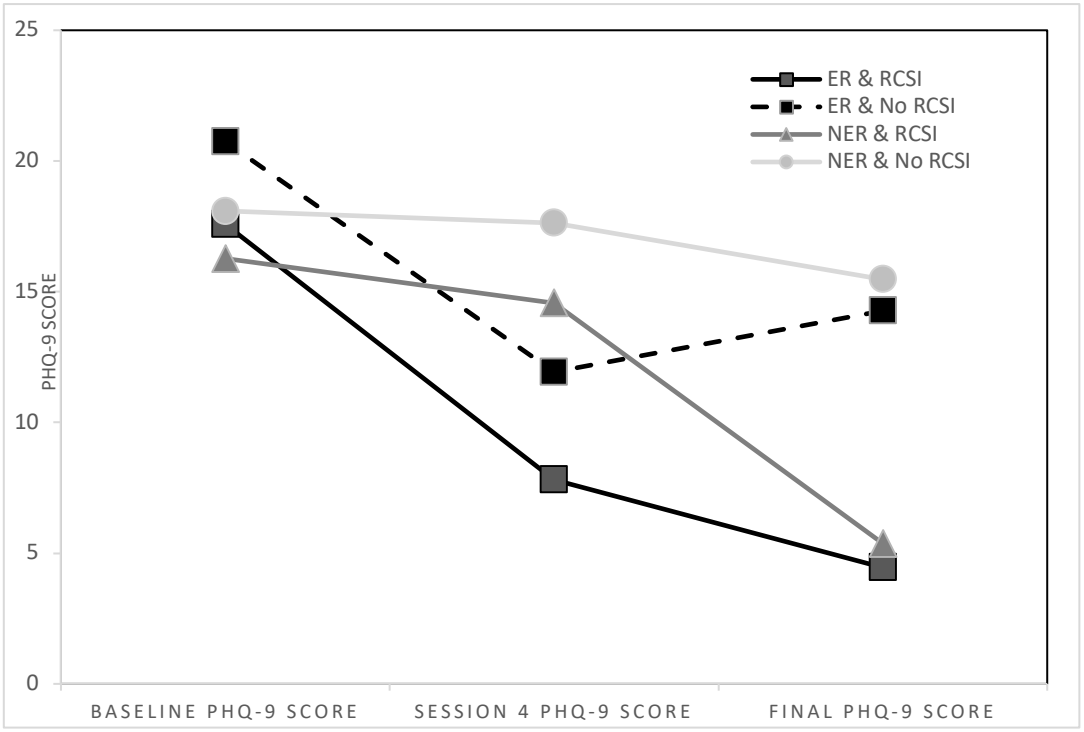
	RCSI	No RCSI	Total
	n (%)	n (%)	n
Early response	980 (29.5)	305 (9.2)	1285
No early response	796 (24.0)	1240 (37.3)	2036
Total	1776	1545	3321

Note. RCSI = Reliable and clinically significant improvement.

Figure 2 presents the PHQ-9 scores from baseline to session four and to final session across these four discrete groups. Figure 2 shows that two classes of patients had the best overall treatment outcomes: *early responders* who continued to improve over time (29.5%) and *eventual responders* who did not show early response but who eventually improved (24.0%). The graph also reveals that cases where early response was a *false positive signal* experienced (on average) a reversal of early gains.

Figure 2

Baseline, session 4, and final session PHQ-9 scores across four discrete patient groups



Note: ER & RCSI = early response and RCSI; ER & No RCSI = early response but no RCSI; NER & RCSI = no early response but later RCSI; NER & No RCSI = no early response nor RCSI.

The early responders who attained RCSI at the end of treatment made 74.4% of their overall gains by the first four sessions. By contrast, eventual responders made 15.6% of their overall gains by the first four sessions and 84.4% of their overall gains between session four

and the final session. Analysis of demographic and clinical characteristics associated with these four response patterns are reported in the supplementary materials. Relative to early responders, logistic regression analysis indicated eventual responders were more likely to be taking medication, have lower baseline depression severity and higher impaired functioning (see supplementary materials for full results).

Logistic regression analysis

The single-level logistic regression model explained 23% of the variance in RCSI (see Table 4), and both early response and baseline severity were significant predictors ($p < .001$). After controlling for baseline depression severity, participants with early response were six times (OR = 6.12) more likely to attain RCSI at the end of treatment compared to those without early response.

Table 4

Logistic regression model

Variable	Pseudo R^2 = .23					
	B	SE B	Wald X^2	*p*	OR	95% CI OR
Constant	1.486	0.15	92.60	<.001	4.42	
Baseline PHQ-9	-0.113	0.009	166.19	<.001	0.89	0.88 to 0.91
RI at session 4	1.812	0.085	456.43	<.001	6.12	5.18 to 7.23

Note. SEB = standard error of B; RI = reliable improvement; PHQ-9 = Patient Health Questionnaire; OR = odds ratio; 95% CI OR = 95% confidence interval for odds ratio; Pseudo R^2 statistic used was Nagelkerke.

Evaluation of model predictive performance demonstrated the summary Area Under the Curve (AUC) was .677. The model's sensitivity was 55% (95% CI [0.53, 0.58]) and its specificity was 80% (95% CI [0.78, 0.82]). For any given positive signal (early response), the positive predictive value was 76% (95% CI [0.74, 0.79]). For any given negative signal (no early response), the negative predictive value was 60% (95% CI [0.59, 0.63]). (See supplementary material for full model performance output).

A post hoc sensitivity analysis of the logistic regression was conducted to compare the odds ratio for the less-than and more-than average duration PCET interventions (see supplemental Table S1 and S2 for the full model results). In less-than average duration interventions, the logistic regression resulted in an OR for early response at session four of 6.92 (95% [CI 5.61, 8.53], $p < .001$). For more-than average duration interventions, the OR for early response at session four was 4.99, (95% [CI 3.77, 6.61], $p < .001$).

Multi-level logistic regression analysis

Results of the multi-level logistic regression model (Table 5) confirmed that early response was associated with a higher probability of RCSI, after controlling for baseline depression severity and therapist effects, with similarly large effect size (OR = 6.18) as in the single-level regression model. There was a significant therapist effect within the model accounting for 10.96% of the variance. A further model was created to test a random slope in reliable improvement at session 4 but it was not significant (see supplemental Table S3 for full model results).

Table 5

Multi-level logistic regression model based on full sample

Fixed	*B*	SE B	*p*	OR	95% CI OR
Constant	-0.797	0.088	<.001	0.45	0.38 to 0.54
Baseline PHQ-9	-0.108	0.009	<.001	0.90	0.88 to 0.91
RI at session 4	1.822	0.089	<.001	6.18	5.19 to 7.36
Random					
Constant	0.405	0.096			

Note. SEB = standard error of B; RI = reliable improvement; PHQ-9 = Patient health Questionnaire-9; OR = odds ratio; 95% CI OR = 95% confidence interval for odds ratio.

A post hoc sensitivity analysis was conducted to determine the effect after removing any cases with a PHQ-9 score less than 10 at session four (Table 6). The OR for early response at session four was reduced to 3.21 in this sub-sample.

Table 6

Multi-level logistic regression model with sample of removed cases with PHQ-9 score less than 10 at session 4

Fixed	*B*	SE B	*p*	OR	95% CI OR
Constant	-0.886	0.094	<.001	0.41	0.34 to 0.50
Baseline PHQ-9	-0.082	0.011	<.001	0.92	0.90 to 0.94
RI at session 4	1.167	0.115	<.001	3.21	2.56 to 4.02
Random					
Constant	0.445	0.111			

Discussion

This study investigated the prognostic value of early response during person-centered experiential therapy (PCET). Results indicated that early response was a strong predictor of better post-treatment outcomes, whereby early responders were six times more likely to achieve reliable and clinically significant improvement (RCSI) at the end of treatment compared to those without an early response. Furthermore, the magnitude and statistical significance of the early response effect was robust after controlling for therapist effects, a well-documented predictor of treatment response in psychotherapy studies (Baldwin & Imel, 2013). There was no evidence of an effect of therapist on the relationship between early response and patient outcome, indicated by a non-significant random slope for reliable improvement at session 4. The sensitivity analyses examining the prognostic value of reliable improvement in PCET interventions of less-than and more-than average treatment duration revealed that although the OR in the latter was reduced, the prognostic value was still significant and large, thereby suggesting the findings are robust across a range of treatment durations.

Given the current study is, to the best of our knowledge, the first to investigate the phenomenon of early response in a form of humanistic-experiential therapy, there is no direct comparison to be made with effect sizes of other studies in the same class of therapy. However, when contrasted against studies of mainly CBT, the current finding exceeded the OR of 4.84 reported in Beard and Delgadillo's (2019) systematic review but less than the OR of 8.75 found by Schlagert and Hiller (2017). However, the latter study used noticeably different criteria for defining early response including, for example, defining early response as a $\geq 50\%$ reduction in the BDI by session 10 in a therapy comprising an average of 40 sessions within a German University outpatient clinic. Overall, the present results indicate that early response is a reliable prognostic indicator for PCET, and the effect of early response may be larger in this form of therapy compared to other interventions such as CBT.

The gains made in the initial four sessions by early responders who subsequently met RCSI at the end of therapy (29.5% of the sample) accounted for 74% of their overall gains. However, almost a quarter of patients (24.0%) that we refer to as *eventual responders* showed that 84% of their overall gains occurred after session four. To our knowledge, this *eventual responder* phenotype is less well documented and understood in the psychological therapy literature. Taking this specific finding into consideration, it would be hasty to conclude that patients who do not show early response by session four are non-responders, since some may in fact be eventual responders. This prognostic error is reflected in the negative predictive value, which indicates that judging a case without early response to be a non-responder would only be correct in around 60% of cases. In order to avoid such a prognostic error, future studies could investigate if eventual responders could be reliably differentiated from non-responders by comparing their baseline features (i.e., symptomatic/ demographic profiles) or process measures (i.e., alliance trajectories). Initial exploratory analyses reported in the supplementary materials provided preliminary evidence that eventual

responders are more likely to be less depressed at intake but taking medication and experiencing worse functioning, thereby suggesting a different clinical presentation and, accordingly, a slower response to therapy.

Although early response is a ubiquitous prognostic indicator across various forms of psychological therapy, relatively little is known about what accounts for such an effect. There is a debate in the literature, with some arguing for the role of techniques (e.g., Fennell & Teasdale, 1986; Tang & DeRubeis, 1999) and others for the role of common factors (e.g., Lambert, 2005). Although it is not possible to disentangle the separate contributions made by technical or common factors to the phenomenon of early response, one feature of PCET in the early sessions may be the validating environment that provides psychological space for patients to narrate their personal experiences – that is, to tell their story – and it may be that it is this process for patients who are not more functionally impaired that yields early benefits from depression. The importance of narrative discourse and expression in humanistic-experiential therapies has been noted (Angus, 2012) and the recognition that personal story disclosures are central to the development of meaning and understanding between patients and therapists (Angus et al., 2004). Nevertheless, intensive process-outcome research (i.e., analysis of session recordings) is necessary to gain greater insight into the determinants of early response.

In terms of the clinical implications of the current findings, within a stepped care model, one suggested approach is to take the absence of early response as a signal or flag and 'step-up' the patient to a more intense intervention (Singla et al., 2019). However, such a strategy ignores the finding from the current study that a considerable portion of patients not showing early response, did ultimately meet the criteria for recovery. Hence, the clinical strategy at session four needs to be able to better distinguish these two groups. One approach would be to adapt strategies used in the feedback component of routine outcome monitoring

such that the absence of early response at session four is taken as a signal or alert for administering what have been termed clinical support tools (Lambert et al., 2015) that would provide additional information about key areas that may be stalling patient progress. Examples of such assessments include the Assessment for Signal Clients (Lambert et al., 2015), which taps four domains: (1) therapeutic alliance, (2) motivation, (3) social support, and (4) coping with problematic life events. Other feedback systems have assessed similar domains such as (1) risk/suicidality, (2) motivation/therapy goals, (3) therapeutic alliance, (4) social support/critical life events, and (5) emotion regulation/self-regulation (Lutz et al., 2019). In addition, the finding that a subgroup of patients who show early response but have elevated baseline scores subsequently fail to maintain those gains can also act as an alert for therapists to view such gains with caution and similarly implement an assessment of these domains.

A significant strength of the current study is the sample size of patients drawn from a routinely collect data set within a single IAPT service in England. Comparing the sample size against those included in Beard and Delgadillo's (2019) review, it would be considered one of the largest samples. Furthermore, the data spans a time frame of 42 months, thereby acknowledging temporal effects and signals the robustness of the findings. The application of multi-level modeling further strengthens the results because therapist effects in routine data sets have consistently explained between 5% and 10% of variability in treatment outcomes (Saxon et al., 2017) and few studies investigating the relationship between early response and final depression outcomes have accounted for this source of variability.

Notwithstanding the strengths of the study, there are a number of caveats. The index of change was a standardized measure of depression symptoms and it might be argued that, as such, this was not the most sensitive outcome to capture patients' gains in a form of humanistic-experiential therapy. However, the measures were implemented by a central

national body focusing on symptom gains in depression and anxiety. Administering measures focusing on interpersonal relations or quality of life would be informative. No longer-term follow-up data were available after the end of treatment limiting assessment of the durability of the early response effect or additional analysis of early withdrawers of treatment, again a function of the nationally implemented program as was the absence of information regarding therapists (i.e., their demographic or professional characteristics). Given that therapists effects accounted for 11% of the variance, future research would benefit from understanding how therapist characteristics contribute to the outcomes and improved early response effects in PCET.

In conclusion, the present study indicates that early response occurs in a sizeable proportion of patients receiving person-centered experiential therapy and is a reliable prognostic indicator, although at least a quarter of patients may display a more eventual pattern of treatment response.

Author Note

None of the authors have any conflicts of interest to declare.

The study design, hypotheses and analysis plan were pre-registered (See https://aspredicted.org/blind.php?x=DSC_Z21). The data and analysis code for this study are not available.

The present research is derived from the first authors dissertation submitted for the completion of an MSc qualification. There has been no prior public dissemination of the data or narrative interpretations of the manuscript.

Acknowledgements

We would like to thank PCMIS for facilitating access to the data set and specifically to Pennine Care NHS Foundation Trust, UK, for permission to use their pseudo-anonymized service data for the purposes of the current research.

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

Angus, L., Lewin, J., Bouffard, B., & Rotondi-Trevisan, D. (2004). What's the story?'': Working with narrative in experiential psychotherapy. In L. Angus & J. McLeod (Eds.), *Handbook of narrative and psychotherapy: Practice, theory and research* (pp. 87–101). Thousand Oaks, CA: Sage.

Baldwin, S., & Imel, Z. E. (2013). Therapist effects: findings and methods. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change, 6th Edition.* (pp. 258–297). Wiley.

Barkham, M., & Lambert, M. J. (2021). The efficacy and effectiveness of psychological therapies. In M. Barkham, Lutz, W., & Castonguay, L. G. (Eds.). *Bergin and Garfield's handbook of psychotherapy and behavior change*. 7th Edition. Wiley.

Barkham, M., Saxon, D., Hardy, G. E., Bradburn, M., Galloway, D., Wickramasekera, N., Keetharuth, A. D., Bower, P., King, M., Elliott, 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 as delivered in the English Improving Access to Psychological Therapies (IAPT) national programme: A pragmatic randomised non-inferiority trial [PRaCTICED]. *The Lancet Psychiatry, 8*, 487–499. https://doi.org/10.1016/S2215-0366(21)00083-3

Beard, J., & 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

Blagys, M. D., & Hilsenroth, M. J. (2002). Distinctive activities of cognitive-behavioral therapy. A review of the comparative psychotherapy process literature. *Clinical Psychology Review, 22*(5), 671–706. https://doi.org/10.1016/s0272-7358(01)00117-9

Charlton, C., Rasbash, J., Browne, W.J., Healy, M. and Cameron, B. (2020) *MLwiN Version 3.05*. Centre for Multilevel Modelling, University of Bristol.

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

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037//0033-2909.112.1.155

Crits-Christoph, P., Connolly, M. B., Gallop, R., Barber, J. P., Tu, X., Gladis, M., & Siqueland, L. (2001). Early improvement during manual-guided cognitive and dynamic psychotherapies predicts 16-week remission status. *The Journal of Psychotherapy Practice and Research*, *10*(3), 145–154.

Delgadillo, J., McMillan, D., Lucock, M., Leach, C., Ali, S., & Gilbody, S. (2014). Early changes, attrition, and dose-response in low intensity psychological interventions. *The British Journal of Clinical Psychology*, *53*(1), 114–130. https://doi.org/10.1111/bjc.12031

Elliott, R., Watson, J. C., Goldman, R. N., & Greenberg, L. S. (2004). *Learning emotion-focused therapy: the process-experiential approach to change*. Washington, DC: American Psychological Association.

Elliott, R., Watson, J., Timulak, L., & Sharbanee, J. (2021). Research on humanistic-experiential psychotherapies: Updated review. In. M. Barkham, W. Lutz., & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change*. 7th ed., pp. 421–467. Wiley.

Fennel, M. J. V., & Teasdale, J. D. (1987). Cognitive therapy for depression: Individual differences and the process of change. *Cognitive Therapy and Research, 11*, 253–271.

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist*, *41*(2), 159–164.

IBM Corp. (2017). *IBM SPSS Statistics for Windows*. Armonk, NY: IBM Corp. Retrieved from https://hadoop.apache.org

Ilardi, S. S., & Craighead, W. E. (1994). The role of nonspecific factors in cognitive-behavior therapy for depression. *Clinical Psychology: Science and Practice, 1*(2), 138–156. https://doi.org/10.1111/J.1468-2850.1994.TB00016.X

Ilardi, S. S., & Craighead, W. E. (1999). Rapid early response, cognitive modification, and nonspecific factors in cognitive behaviour therapy for depression: a reply to Tang and DeRubeis. *Clinical Psychology: Science and Practice, 6*(3), 295–299. https://doi.org/10.1093/clipsy.6.3.295

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037//0022-006x.59.1.12

Johns, R. G., Barkham, M., Kellett, S., & Saxon, D. (2019). A systematic review of therapist effects: A critical narrative update and refinement to review. *Clinical Psychology Review*, *67*, 78–93. https://doi.org/10.1016/j.cpr.2018.08.004

Kadera, S. W., Lambert, M. J., & Andrews, A. A. (1996). How much therapy is really enough? A session-by-session analysis of the psychotherapy dose-effect relationship. *Journal of Psychotherapy Practice & Research, 5*(2), 132–151

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*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lambert M. J. (2005). Early response in psychotherapy: further evidence for the importance of common factors rather than "placebo effects". *Journal of Clinical Psychology, 61*(7), 855–869. https://doi.org/10.1002/jclp.20130

Lambert, M. J., Bailey, R. J., White, M., Tingey, K. M., & Stevens, E. (2015). *Clinical support tool manual (Brief Version-40*). Salt Lake City, UT: OQMeasures.

Lemma, A., Roth, A., & Pilling, S. (2008). *The competences required to deliver effective psychoanalytic/psychodynamic therapy*. London: Department of Health. https://www.ucl.ac.uk/pals/research/clinical-educational-and-health-psychology/research-groups/core/competence-frameworks-6

Lutz, W., Rubel, J. A., Schwartz, B., Schilling, V., & Deisenhofer, A. -K. (2019). Towards integrating personalized feedback research into clinical practice: Development of the Trier Treatment Navigator (TTN). *Behaviour Research and Therapy, 120*, 103438. https://doi.org/10.1016/j.brat.2019.103438

Murphy, D. (2019). *Person-centred experiential counselling for depression: a manual for training and practice*. London, UK: SAGE Publishing.

National Collaborating Centre for Mental Health. (2018). *The Improving Access to Psychological Therapies Manual*. https://www.england.nhs.uk/publication/the-improving-access-to-psychological-therapies-manual/

National Health Service (NHS) Digital (2019). *Psychological therapies, annual report on the use of IAPT services 2018–19*. https://digital.nhs.uk/data-and-information/publications/ statistical/psychological-therapies-annual-reports-on-the-use-of-iapt- services/annual-report-2018-19

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

Renaud, J., Brent, D. A., Baugher, M., Birmaher, B., Kolko, D. J., & Bridge, J. (1998). Rapid response to psychosocial treatment for adolescent depression: a two-year follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*(11), 1184–1190. https://doi.org/10.1097/00004583-199811000-00019

Rodríguez, G., & Goldman, N. (2001). Improved estimation procedures for multilevel models with binary response: a case study. *Journal of the Royal Statistical Society: Series A, 164*, Part 2, 339–355. https://doi.org/10.1111/1467-985X.00206

Roth, A., Hill, A., & Pilling, S. (2009). *The competences required to deliver effective humanistic psychological therapies.* London: Department of Health. https://www.ucl.ac.uk/pals/research/clinical-educational-and-health-psychology/research-groups/core/competence-frameworks-4

Roth, A., & Pilling, S. (2007). *The competences required to deliver effective cognitive and behavioural therapy for people with depression and with anxiety disorders*. London: Department of Health. https://www.ucl.ac.uk/pals/research/cehp/research-groups/core/competence-frameworks/cognitive-and-behavioural-therapy

Rubel, J., Lutz, W., Kopta, S. M., Köck, K., Minami, T., Zimmermann, D., & Saunders, S. M. (2015). Defining early positive response to psychotherapy: An empirical

comparison between clinically significant change criteria and growth mixture modeling. *Psychological Assessment*, *27*(2), 478–488. https://doi.org/10.1037/pas0000060

Sanders, P., & Hill, A. (2014). *Counselling for depression: a person-centred and experiential approach to practice.* London, UK: SAGE Publishing.

Schlagert, H. S., & Hiller, W. (2017). The predictive value of early response in patients with depressive disorders. *Psychotherapy Research, 27*(4), 488–500. https://doi.org/10.1080/10503307.2015.1119329

Singla, D. R., Hollon, S. D., Fairburn, C. G., Dimidjian, S., & Patel, V. (2019). The roles of early response and sudden gains on depression outcomes: Findings from a randomized controlled trial of behavioral activation in Goa, India. *Clinical Psychological Science, 7*(4), 768–777. https://doi.org/10.1177/2167702619825860

Stiles, W. B., Shapiro, D. A., & Elliott, R. (1986). "Are all psychotherapies equivalent?". *American Psychologist, 41*(2), 165–180. https://doi.org/10.1037//0003-066x.41.2.165

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

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, *67*(6), 894–904. https://doi.org/10.1037//0022-006x.67.6.894

Taylor, A., Tallon, D., Kessler, D., Peters, T. J., Shafran, R., Williams, C., & Wiles, N. (2020). An expert consensus on the most effective components of cognitive behavioural therapy for adults with depression: a modified Delphi study. *Cognitive Behaviour Therapy, 49*(3), 242–255. https://doi.org/10.1080/16506073.2019.1641146

Van, H. L., Schoevers, R. A., Kool, S., Hendriksen, M., Peen, J., & Dekker, J. (2008). Does early response predict outcome in psychotherapy and combined therapy for major depression? *Journal of Affective Disorders*, *105*(1-3), 261–265. https://doi.org/10.1016/j.jad.2007.04.016