The University Of Sheffield.

The Role of Emotional Processing Across Psychological Therapies and Its Influence on Client Outcomes

Jack Isgar

A thesis submitted in partial fulfilment of the requirements for the award of Doctorate in Clinical Psychology (DClinPsy)

Clinical and Applied Psychology Unit

Department of Psychology

The University of Sheffield

May 2024

Declaration

I declare that this thesis has been submitted in partial completion of the award of Doctorate in Clinical Psychology at the University of Sheffield. This thesis has not been submitted to any other institution, or for the purpose of obtaining any other qualifications.

Word Counts

Chapter One: Literature Review

Excluding abstract, figures, tables, references and appendices: 7,950 words

Including abstract, figures, tables, references and appendices: 16,217 words

Chapter Two: Research Report

Excluding figures, tables, references and appendices: 7,759 words

Including figures, tables, references and appendices: 12,054 words

Total Word Count

Excluding figures, tables, references and appendices: 15,709 words

Including figures, tables, references and appendices: 28,271 words

Lay Summary

Several studies have found associations between patient depth of experiencing and psychological therapy treatment outcomes. However, the overall number of studies is sparse and have often been small in sample size. While a previous meta-analysis found a small but significant association between depth of experiencing and outcomes, there were a number of limitations with its methodology.

Research has generally shown that individuals accessing experiential therapies explore their feelings in greater depth compared to those accessing cognitive-behavioural therapy. Additionally, depth of experiencing generally appears to increase from early to mid-therapy sessions. However, the research base for these conclusions is limited and there are psychological approaches in which this association has not been explored. One such example is person-centred experiential therapy, which is the focus of the current research.

This thesis had two broad aims: 1) to conduct an updated and refined systematic review and meta-analysis by investigating the relationship between depth of patient experiencing and treatment outcomes, and 2) an empirical study exploring the depth of experiencing in person-centred experiential therapy compared with cognitive-behavioural therapy in early therapy sessions versus mid-therapy sessions. The plan for the research was pre-registered so as to protect the integrity of the research.

The first section of this thesis reports on a systematic review of 30 studies, all of which explored the relationship between depth of patient experiencing (using the Experiencing Scale; Klein et al., 1969) and therapeutic outcomes. Two meta-analyses were conducted, using 13 papers comprising of 15 datasets that met inclusion criteria, with the aim of determining the strength of association between depth of experiencing and treatment outcomes. Results showed a significant association between depth of experiencing and therapy outcomes. The findings corroborate results from the previous meta-analysis. Further analyses suggested the extent of variability between

studies was small, with the exception of one outlier, and the included studies were representative of studies within the research field.

The empirical section reports on a secondary data analysis of the PRaCTICED dataset (Barkham et al., 2021). The study aimed to partially replicate an earlier study, conducted by Watson and Bedard (2006), by investigating the relationship between depth of experiencing in person-centred experiential therapy and cognitive-behavioural therapy. In addition, it aimed to establish whether depth of experiencing would increase in both therapies from early therapy to working (middle) therapy sessions. Results showed that depth of experiencing was significantly higher for clients who accessed person-centred experiential therapy in comparison to cognitive-behavioural therapy. This result supports findings from previous studies in which clients accessing experiential therapies explored their feelings in greater depth compared to those accessing cognitive-behavioural therapy. However, there were no significant changes in depth of experiencing from early to mid-therapy sessions. This result differs from the majority of findings in the literature. Limitations, clinical implications, and suggestions for future research are described.

Acknowledgements

Thank you to Jeanne Watson for your support and guidance during early discussions of my research. Thank you also to The International Focusing Institute, especially Elizabeth Cantor, for facilitating access to the training materials.

Thank you to my supervisors; Professor Gillian Hardy, Dr David Saxon and Professor Michael Barkham for your support and guidance. Thank you, also, to my collaborating researchers.

My greatest thanks go to my family. Words cannot describe the appreciation I have for the unwavering support you have offered across the years. Without your support, this wouldn't have been possible.

Lastly, to Laura, my greatest achievement of this course. I will forever be indebted to you for standing alongside me throughout this process, for your unrelenting love and for the encouragement you have provided me with across these three years. I cannot wait to start the next chapter together.

Table of Contents

Title Page…….....i

Declaration……. ii

Word Counts…….....iii

Lay Summary……. iv

Acknowledgements…….....vi

Section One: Literature Review

The Client "Experiencing" Scale (EXP) as a Predictor of Treatment Outcomes: An Updated and Extended Systematic Review and Meta-Analysis on Psychotherapy Process

Abstract.....1

Introduction.....3

Emotional Processing.....3

The Experiencing Scale.....3

Past Reviews on Emotional Processing Using the EXP and its Relationship to Treatment Outcome.....5

Aims of the Current Review.....7

Hypotheses.....7

Methods.....7

Search Strategy.....7

Inclusion and Exclusion Criteria.....9

Screening.....10

Data Extraction.....11

Quality Assessment.....11

Meta-analytic Strategy.....12

Heterogeneity.....14

Moderator Analysis.....15

Sensitivity Analysis.....16

Publication Bias.....16

Results.....17

Study Selection.....17

General Study Characteristics.....19

Intervention Characteristics.....19

Study Characteristics – Overlapping Datasets.....20

Methodological Quality.....26

Meta-Analysis Results.....27

Moderator Analysis.....33

Sensitivity Analysis.....33

Publication Bias.....34

Moderator Analysis.....36

Sensitivity Analysis.....36

Publication Bias.....36

Discussion.....35

Meta-analyses.....38

Sensitivity Analyses.....38

Strengths and Limitations.....39

Research and Clinical Implications.....41

Conclusions.....42

References.....42

Appendices 54

Section Two: Empirical Study

Clients' Emotional Processing in Psychotherapy: A Comparison Between Cognitive-Behavioural and Person-Centred Experiential Therapies Utilising the Archived PRaCTICED Trial Data Set

Abstract....73

Introduction....75

The Experiencing Scale....75

Emotional Processing Research....76

Emotional Processing and its Relation to Outcomes....77

Current Study....77

Clinical Rationale and Implications....78

Research Aims....79

Hypotheses....79

Methods....79

Design....79

Ethical Considerations and Open Science....80

Participants....80

Therapists and Counsellors....81

Procedure....81

Primary Measures in the Current Study....82

Data Security....83

Analysis....84

Quality Control....87

Service User Involvement....88

Results....88

Demographic Information....88

Descriptive Statistics....89

Reliabilities....89

Mean Modal Ratings....90

Post-Hoc Power Analyses....91

Post-hoc Analyses for Modal Emotional Processing Scores.....92

Mean Peak Ratings.....92

Post-Hoc Power Analyses.....94

Post-hoc Analyses for Peak Emotional Processing Scores.....94

Discussion.....95

Comparison to Current Literature.....96

Theory Development.....97

Limitations and Future Directions.....98

Clinical Implications.....100

Conclusions.....101

References.....101

Appendices.....113

Section One: Literature Review

The Client "Experiencing" Scale (EXP) as a Predictor of Treatment Outcomes: An Updated and Extended Systematic Review and Meta-Analysis on Psychotherapy Process

Abstract

Objectives: This review had three aims. First, to synthesise the literature on EXP and treatment outcomes. Second, to explore the strength of association between mid-therapy EXP and treatment outcomes. Third, to examine variables which influence this relationship.

Method: A systematic search of six databases using specific search terms was initially conducted. Papers were systematically reviewed using pre-determined inclusion and exclusion criteria. Data extraction, quality appraisal, and a narrative synthesis were completed on the papers remaining after screening. Two random-effects meta-analyses were completed, one using modal EXP scores and the other using peak EXP scores. Sensitivity analyses were conducted**,** with papers removed which used different methodologies such as therapist rated outcomes.

Results: The searches identified 30 studies meeting criteria for the systematic review, the majority of which were rated of moderate quality. Of the studies meeting criteria for the meta-analyses, a moderate relationship between EXP and treatment outcomes was found for modal EXP ($r = 0.33$, 95% CI [0.10, 0.52], $p = 0.001$) and peak EXP ($r = 0.34$, 95% CI [0.24, 0.42] $p < 0.0001$). Moderator analyses were not possible to address the third review objective due to the limited number of papers. Effect sizes remained significant after sensitivity analyses removed papers which used therapist rated outcomes or papers which did not explore mid-therapy sessions.

Conclusion: The results show a significant relationship between depth of experiencing and treatment outcomes, which corroborates with the results from the previous meta-analysis.

Key Words: Experiencing, Emotional Processing, EXP, Psychotherapy.

Practitioner Points

- Training of therapists should contain some focus on developing client experiencing.

- Supervision of therapists should include discussion on client experiencing and how to maximise this in therapy.

Introduction

Psychological therapies research focuses on events between therapists and clients during or between therapy sessions (Howard et al., 1986). Its purpose is to provide an understanding of and explain variations in patient outcomes through the exploration of factors that lead to, or prevent, individual change (Falkenstron et al., 2017). The hope is that with greater understanding of change mechanisms and change processes, researchers can aim to improve the quality of therapy for both clients and therapists, which should lead to improved therapeutic outcomes.

Emotional Processing

Emotional processing is the term used to describe a person's involvement in therapy, which includes their awareness of self throughout a session and the way their feelings are experienced (Klein et al., 1969). It has been recognised as a variable which influences outcomes across therapies (Town et al., 2017) and has therefore been widely studied, with Watson (2023) stating that understanding the role of emotional processing across therapies has been a key objective of their work in process research.

A range of measures have been created to measure emotional processing including, for example, the Classification of Affective Meaning States (CAMS; Pascual-Leone & Greenberg, 2005) and the Narrative Emotion Process Coding System 2.0 (NEPCS; Angus et al., 2017). However, The Experiencing Scale (EXP; Klein et al., 1969) is the most used measure and has been used to explore emotional processing in a number of therapies (Hendricks et al., 2009)

The Experiencing Scale

The EXP scale (Appendix A) measures emotional processing along a seven-point continuum (see Table 1), with each turn-take of speech scored on this scale. Thus, a passage of speech yields any given number of scores dependent on the number of turn-takes between patient and therapist. The most common score is known as the *modal* score, whilst the highest score (most in-depth level of experiencing) is known as the *peak* score. Anchor points for the scale are as follows: Stage 1 is characterised by speech which is impersonal (e.g., the speaker may focus on external events or other

people); Stage 7 is characterised by the speaker sharing and exploring an internal struggle, sharing different formulations about themselves, and reaching a conclusion that could incorporate a number of different solutions.

Published results suggest intraclass correlation reliabilities of the EXP scale are good, with scores ranging from .76 (Schoeninger et al., 1967) to .92 (Kiesler et al., 1964). Pascual-Leone et al., (2016) stated that psychotherapy process researchers viewed the EXP scale as the "gold standard" of good psychotherapy process.

Table 1

The Stages of The EXP Scale and a Description of Each Level.

Stage 1	The speaker talks in a detached way about external events, that are not about the speaker. There are no personal details in the content.
Stage 2	Events that the speaker discusses are still external, however the association between the content and speaker is clear. Their participation in the event is made clear, however their involvement does not go beyond a description of events.
Stage 3	The speaker continues to describe external events. Their participation in the events is clear. The speaker provides a limited self-description of their reaction to the event describe their reactions to the events, such as attaching a behavioural description of the feeling.
Stage 4	The speaker will describe the content completely from their point of view. The speaker will build on what was described at Stage 3 by speaking more deeply about their personal perspective, including attaching several feelings to the event, and sharing what they are like more generally.
Stage 5	The speaker begins to explore their feelings. The speaker must define a problem about themselves regarding their feelings and they must work with the problem in a personal way.

Stage 6	The speaker will vividly share their feelings, synthesising their increased understanding of their feelings and experiencing, to resolve an issue.
Stage 7	The content is expansive in nature, integrating elements of the above. The awareness of the client has increased, meaning they can shift from one inner reference to another. Expanding self-awareness provides a platform for further exploration.

Past Reviews on Emotional Processing Using the EXP Scale and its Relationship to Treatment Outcome

There have been a limited number of studies which have summarised the EXP literature. Earlier reviews were predominantly narrative in nature. Luborsky et al. (1988) provided the first review which disseminated results from 11 studies. Due to the qualitative nature of most of these studies (8 of 11), tentative conclusions were drawn regarding emotional processing as a predictor of outcomes.

Given the rise in the number of publications using the EXP Scale to over 100 papers by 2009 (Hendricks, 2009) most studies have not explored EXP as a predictor of treatment outcomes. Despite this situation, researchers routinely came to the same conclusion, namely that there was a positive association between EXP and treatment outcome (Elliott et al., 2013; Hendricks, 2002, 2009).

The most comprehensive review of the EXP Scale and its predictive value in relation to therapy outcomes was conducted by Yeryomenko (2012). The paper was later published (Pascual-Leone & Yeryomenko, 2016), with no further papers included. The study reviewed 10 papers and found that emotional processing (as measured using the EXP) was significantly related ($r = 0.25$, 95% CI [0.16, 0.33], $p < 0.001$) to improvement. The authors explored several moderating variables. They found that type of outcome measure moderated the effects, with higher effects obtained for observational measures compared with self-report measures. Treatment phase and

model of therapy did not moderate results. The authors raised a number of limitations with their study which included the small number of studies meaning that conclusions were "tenuous at best" and the lack of diversity in which samples were drawn from. Of note, the full-text paper review process as part of their search strategy resulted in eight papers being removed as they had used overlapping datasets. Although the paper stated which studies overlapped, it did not provide any further information on how they overlapped or the original datasets from which these papers data were drawn from.

However, the Pascual-Leone and Yeryomenko (2016) study contains at least six methodological limitations: (1) an unclear search strategy that only yielded 190 papers; (2) no quality appraisal of included studies; (3) no summary of overlapping datasets; (4) pooling correlations across a range of variables which measured very different ideas (e.g., early and late therapy, as well as modal and peak EXP); (5) no inclusion of a forest plot; and (6) only including one test of publication bias. Pascual-Leone and Yeryomenko (2016) recommended that there was a need for more papers which study several therapies, outcome measures, and participants before broader generalisations can be made. Since their publication, there have been a number of studies conducted that add to those included in their review.

More recently, Sonderland et al. (2023) conducted a systematic review and meta-analysis that aimed to summarise current knowledge on emotional change processes and mechanisms, and their relationship with outcomes in psychotherapy. In terms of psychotherapy for depression, they found experiencing was the emotional change process most strongly linked to outcome ($r = 0.44$, 95% CI [0.31, 0.55], $p < 0.0001$). It should be noted that the reviewers explored the main change processes and mechanisms in the literature and did not focus specifically on experiencing. Within their analyses on experiencing, they did not solely focus on the EXP Scale, focussing instead on a range of tools which measured experiencing. Treatment outcomes in relation to experiencing were only reported for depression. There were at least four methodological problems associated with the

Sonderland study, these included; (1) no second reviewer of papers, (2) exclusion of grey literature, (3) exclusion of case studies, and (4) no systematic forward and backward referencing.

Aims of The Current Review

Despite client experiencing being a therapeutic process associated with therapeutic outcome, there has been limited research aimed at increasing the understanding of its predictive power. Studies which have explored it had a number of methodological limitations which the current paper aimed to remedy.

The current study had three aims. The primary aim of the current review was to synthesise the literature on EXP and its relation to treatment outcomes. Given Pascual-Leone and Yeryomenko (2016) only discovered a significant portion of literature in the field was being published using the same dataset whilst conducting their analysis, the current study aimed to explore and understand which original datasets are being repeatedly used and the exact proportion of papers that have relied on the same datasets. The second aim was to examine the predictive qualities of the EXP Scale on treatment outcomes, running separate analyses for modal EXP and peak EXP. The third aim was to explore variables that may moderate the relationship between EXP and outcome, if there were enough papers to run moderator analyses. Subgroup analyses previously unexplored include presenting problem, depression versus other.

Hypotheses

It was hypothesised that treatment outcomes would be significantly better for those who explored their feelings in more depth (higher modal and peak EXP scores).

No specific hypotheses were made regarding the potential impact of variables on outcomes as it was unclear if there would be sufficient papers to run the required analyses.

Method

Search Strategy

The systematic review and meta-analysis were conducted adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, Appendix B) guidelines

(Liberati et al., 2009), as advised by the Centre for Reviews and Dissemination (CRD; 2009). The protocol for the review was pre-registered with PROSPERO (CRD42023467090).

The following bibliographic electronic databases were searched; PsycINFO via OVID, ERIC, Medline via OVID, CINAHL, Scopus and ProQuest: Dissertations and Theses. Databases chosen replicated the databases used in the previous review where possible (PsycINFO, ERIC and Medline) and incorporated new databases to ensure breadth and depth of searches. Publication dates were limited from 1969 (the year that the EXP Scale was published) to the year of the searches (2023). A manual review of references and citations within studies included for full review was conducted (i.e., forward and backward referencing). The Lead Reviewer also contacted authors who appeared prominent (contributed several studies) in the topic of research on experiencing.

The search terms and Boolean logic used are displayed in Table 2. Search terms were drawn from terms used in previous reviews and key words in in the emotional processing literature. Search terms were refined following a consultation with the university librarian. Searches were conducted between 9th - 10th October 2023. Searches were repeated on 6th May 2024, covering the timeframe of October 2023 to May 2024. No new papers were discovered.

Table 2

Search Syntax Used in Database Searching.

Construct	Search Terms
Emotional Processing	"experiencing scale" OR "emotional experiencing" OR "client experiencing" OR "patient experiencing" OR "depth of experiencing" OR "emotion* process*" OR "experien* process*"
Therapy	psychotherap* OR "psychotherap* process*" OR "psycho* therap*" OR "psych* intervention"

Note. Terms were searched as keywords. Constructs were combined using AND.

Inclusion and Exclusion Criteria

The Population, Intervention, Comparison, Outcomes and Study (PICOS) framework (Amir-Behghadami & Janati, 2020) was used to refine the inclusion and exclusion criteria which were subsequently used to identify papers that met the criteria to be included in the review (Table 3). Most of the inclusion criteria from the Pascual-Leone and Yeryomenko study were retained. Two of the inclusion criteria from the previous meta-analysis were removed. The first specified that studies must have non-overlapping dataset. This criterion was removed as one of the aims of the current study was to determine the proportion of literature using the same datasets as this was not explored in the last review. The second was that at least two independent raters need to code emotional processing using the EXP Scale. This criterion was removed to increase the likelihood of including all relevant papers. Additional inclusion criteria added were; (a) studies must have been published after 1969, and (b) studies must have been written in the English language.

Table 3

Inclusion and Exclusion Criteria for Study Selection Using the PICOS Framework.

Construct	Inclusion	Exclusion
Population	Individuals of any age accessing psychotherapy.	Couples accessing therapy.
Intervention	Any psychological therapy.	Non-psychological interventions.
Comparison	Due to the study exploring strength of association between EXP and clinical outcomes, this is not applicable.	Due to the study exploring strength of association between EXP and clinical outcomes, this is not applicable.
Outcome	Data from which effect sizes of the relationship between	

	EXP and treatment outcome can be calculated. Measures of in-session emotional processing using the EXP Scale and, pre and post therapy clinical treatment outcomes data.	
Study Type	Quantitative studies, published in the English language after 1969.	Qualitative studies.

Screening

Following the completion of the searches, the lead author independently exported all identified papers to the desktop reference manager software Rayyan (Ouzzani et al., 2016). The duplicator detector was used to highlight potential duplicates. The remaining titles and abstracts were reviewed by the Lead Researcher to determine if they met the inclusion criteria. A sub-sample (25%) of these papers were screened by an independent reviewer (LW) to ensure a consistent approach to inclusion and exclusion of papers. The level of agreement was 99.75%. The one paper that the lead reviewer and collaborator disagreed on was discussed until a resolution was reached.

All articles that met inclusion criteria underwent a full-text screening by the lead author. A sub-sample (25%) of these papers were reviewed by an independent reviewer (LW) to ensure a consistent approach. The level of agreement was 95%. If full text for papers could not be accessed, the author of that paper was contacted. Authors were given three weeks to respond. Papers were removed if there was no response.

Although the Burgess-Moser (2012) and Makinen and Johnson (2006) papers were included in the previous review, they were removed from the current review at the screening stage because they used participants accessing couples therapy.

Data Extraction

For those articles remaining after full-text screening, extraction was completed by the primary author using a data extraction tool created before the review was conducted. As recommended by Boland et al. (2017), the tool was piloted on several studies. The aim of this process was to understand how easy it was to extract the data required and to assess whether the desired data was included within the chosen studies. Data extracted partially replicated the previous meta-analysis's process, which included extrinsic characteristics (i.e., authors, date, publication status, country, presenting problem), treatment characteristics (modality, frequency, duration), methodological characteristics (method of session sampling, EXP sampling method), sample characteristics (sample size, age, gender), measurement characteristics (modal or peak scores; measurement during early, middle, or late treatment; reliability of data; percent of data reviewed), outcome measures and effect sizes. Additional data extracted comprised source of data (novel data or repeated data), study design, and clinic setting.

Data from 25% of the studies were additionally extracted by a collaborating researcher (LW). Data extracted by both researchers were compared to assess for reliability. Interrater reliability was 100%.

Quality Assessment

Boland et al. (2017) describe quality appraisal as the extent to which a study utilises different measures to reduce bias and error in its processes, such as design, implementation, and analysis. According to Dreier (2013), quality appraisal is a crucial process to complete when conducting a meta-analysis.

The Effective Public Health Practice Project Quality Assessment Tool (EPHPP; Thomas et al., 2004; Appendix C) was used to assess the quality of each paper. The EPHPP was designed to

assess quality of studies in “a wide range of health-related topics” (Thomas et al., 2004), thus making it suitable for the current research. The tool assesses six areas within a study; selection bias, study design, confounders, blinding, data collection methods, and withdrawals/drop-outs. Each area is coded on a three-point scale; strong, moderate, and weak. The authors created a supplementary guide to support raters in coding each of the six factors. Once the factors have been scored, a global rating is calculated and the paper is either rated strong, moderate, or weak. Strong papers do not contain any weak factors, moderate papers contain one, whilst weak papers contain two or more. The EPHPP has excellent inter-rater agreement (Armijo-Olivo et al., 2012) and good construct validity (Thomas et al., 2004).

Harrison et al. (2017) recommend that quality appraisal is completed by two individuals for two reasons. First, because of the subjectivity in quality appraisal tools, and second to improve the overall quality of the review. Therefore, initially the Lead Researcher completed the quality scoring of all included studies. To ensure reliability of the quality assessment procedures, a second reviewer (LW) assessed a random sub-sample (25%) of the articles. The level of agreement was 100%. Studies were not excluded based on their quality score.

Meta-analytic Strategy

During data extraction it became clear that some datasets had been used by multiple papers. On occasion, this included papers analysing the same participants and sessions, and using the same EXP codings. Therefore, when datasets had been used by more than one study, a systematic process was taken with regards to choosing which study to include in the meta-analyses. The primary paper, with the greatest number of participants was chosen for inclusion. This process is the same as that used by Pascual-Leone and Yeryomenko (2016) and has been recommended in the literature (Wood, 2008).

If studies reported Pearson product–moment correlation coefficients for the relationship between EXP score and therapeutic outcomes, the effect sizes were extracted. If studies reported alternative statistics, such as T-Tests, then a Pearson product–moment correlation coefficient was

calculated using The Campbell Collection Effect Size Calculator (Wilson, 2023). Similarly, standardised regression coefficients were converted into correlation coefficients. Although, regression coefficients were historically inputted as r-type effect sizes, more recent literature has suggested that this methodology inflates the effect sizes (Peterson & Brown, 2005). If authors reported partial correlations, the alternative partial correlation Meta-Essentials workbook version 1.5 (Suurmond et al., 2017) was used.

If there were not enough details to calculate a correlation coefficient, the author was contacted to obtain the required information. Authors were given three weeks to respond to such requests. If an author did not respond, then their work was still included within the systematic review but excluded from the meta-analysis.

Depending on the outcome measures used, some results suggested a positive effect size indicating an improvement in outcomes, whilst for other measures, a negative effect size indicated an improvement. Therefore, where appropriate, the effect sizes were switched so that a positive effect size indicated a positive relationship between higher level of EXP and improved therapy outcome, whilst a negative effect size indicated a negative relationship between higher EXP and worse therapy outcomes.

Some studies reported more than one effect size as they utilised more than one outcome measure. In studies that reported multiple effect sizes for different outcome measures, a systematic approach was used in terms of selecting the most relevant outcome measure to the presenting problem of the client in the first instance. For example, if the client presented with depression, then a measure of depression was prioritised.

A systematic approach was taken regarding data used from each dataset to ensure that, as far as possible, the same type of data was being analysed, in relation to timing of therapy. The decision was taken to use working therapy data where possible, this was because 1) EXP coding's at different stages of therapy were explored in the previous meta-analysis so it was deemed unnecessary to re-analyse this, 2) working therapy was previously found to have the strongest

association with outcome, and 3) the majority of papers used working therapy phases. If a paper explicitly reported on working therapy phase data, this was used in the first instance. For papers that had averaged EXP data from different stages of therapy, then this was automatically used also. If papers had only collected data at early or late therapy, they were included in the initial meta-analyses but removed as part of the follow-up sensitivity analysis.

If a paper reported effect sizes for both modal and peak EXP, then both effect sizes were extracted. Dependent on the number of studies with modal and peak EXP scores, it was planned to run two meta-analyses (one using modal results and the other using peak results).

Once all effect sizes were converted into the same metric, a random effects meta-analysis was conducted. This was deemed more appropriate than a fixed-effects meta-analysis, because heterogeneity across studies would invalidate the assumptions of a fixed-effects meta-analysis (Boland et al., 2017). The method of weighting used was the inverse-variance method, studies with larger samples have smaller standard errors and thus have a larger weighting. This appears to be the most optimal approach according to the literature (Marin-Martinez et al., 2010).

Interpretation was based on Cohen's (1988) recommendations of $r = 0.1$ signifying a small association, $r = 0.3$ a medium association, and $r = 0.5$ a large association.

Heterogeneity

A meta-analysis quantitatively integrates the effect sizes from different studies to understand the pattern of effects (Marin-Martinez et al., 2010). Heterogeneity explains the variation in outcomes between studies (Boland et al., 2017), and the extent of heterogeneity was used to inform the interpretation of the summary effect. Four methods were used to explore this; visual inspection of the forest plot, Q-statistic (Cochran, 1954), I^2 (Higgins & Thompson, 2002), and Tau.

Initially the forest plot was examined to assess for heterogeneity between studies. A limitation of this approach alone is that it is subjective and thus more robust methods are required to determine heterogeneity (Boland et al., 2017). Therefore, the Q-statistic test was used which assesses whether there is significant heterogeneity. However, a limitation of this approach is that it

does not provide any further information regarding the amount of heterogeneity, just whether it is present or not (Huedo-Medina et al., 2006).

The I^2 analysis measures the proportion of total variability due to between-study heterogeneity. The interpretation thresholds are 25% (low), 50% (moderate), and 75% (substantial) heterogeneity (Higgins et al., 2003). The limitation with this method is that it is not influenced by the number of the studies but by the precision of the studies (Thorlund et al., 2012). So, studies that have larger samples will have smaller sampling error.

Tau estimates and reports the between-study variance and standard deviation of effect sizes across studies. Tau is not influenced by number of studies or their precision.

Moderator Analysis

With heterogeneity analyses informing whether there are differences between studies, moderator analyses aim to explain any such differences (Borenstein et al., 2011). Potential moderator variables were listed within the original protocol registration, a priori. Due to the uncertainty regarding the number of studies that would be suitable for the meta-analysis, conducting moderator analyses were not finalised until the final number of studies was known and if significant heterogeneity was confirmed. The potential variables chosen to be moderators were based on the variables studied (and not studied) in the previous analyses and existing literature. All potential variables were categorical in nature, meaning subgroup analyses would be performed rather than meta-regressions.

Caution was taken to plan only a limited number of moderator analyses. The primary reason is that the risk of obtaining a false positive (Type 1 error) increases as a function of subgroup analyses being conducted (Wang & Ware, 2014). It was also known that moderator analysis numbers were going be restricted due to limited numbers of final papers. Card (2012) recommended that subgroup analyses were only performed when there was a minimum of 10 studies that could be included and where each subgroup had a minimum of three studies. This rule was applied once all studies had been reviewed.

Sensitivity Analysis

Sensitivity analyses aim to examine the robustness of results to the methodological decisions of the study (Deeks et al., 2019) and requires the re-running of analyses after altering a systematic decision, to examine whether this influences the results. Dependent on the ability to conduct moderator analyses and dependent on the number and variety of studies included, then sensitivity analyses would be conducted. These decisions were made a-posteriori, once it was known what the final sample of studies would be.

Publication Bias

Publication bias is the process whereby the likelihood of a study being published is increased when a significant finding is found, a positive result is discovered and there is a larger effect (Nair, 2019). The result of this is an upwards bias in the summary effects. As a result of publication bias, studies included in the analysis may not be representative of all studies within that research field. Given the current meta-analysis focussed on correlations, there was a possibility of bias towards the publication of papers that reported significant results. Despite the inclusion of grey literature, which is likely to reduce publication bias, it is still recommended to assess for publication bias.

There are several methods to explore, assess, and remove publication bias. Initially, visual analyses of funnel plots were conducted. However, this method alone can be subjective, so Egger's Regression Test (Egger et al., 1997) was also conducted to examine whether the association between estimated effects and study size was greater than expected to occur by chance. The trim-and-fill method was also used in combination with the fail-safe N (Rosenthal, 1979).

Results

Study Selection

Figure 1 provides an overview of the process of identifying the papers to be included in the current review. Following the completion of the searches, the Lead Reviewer independently

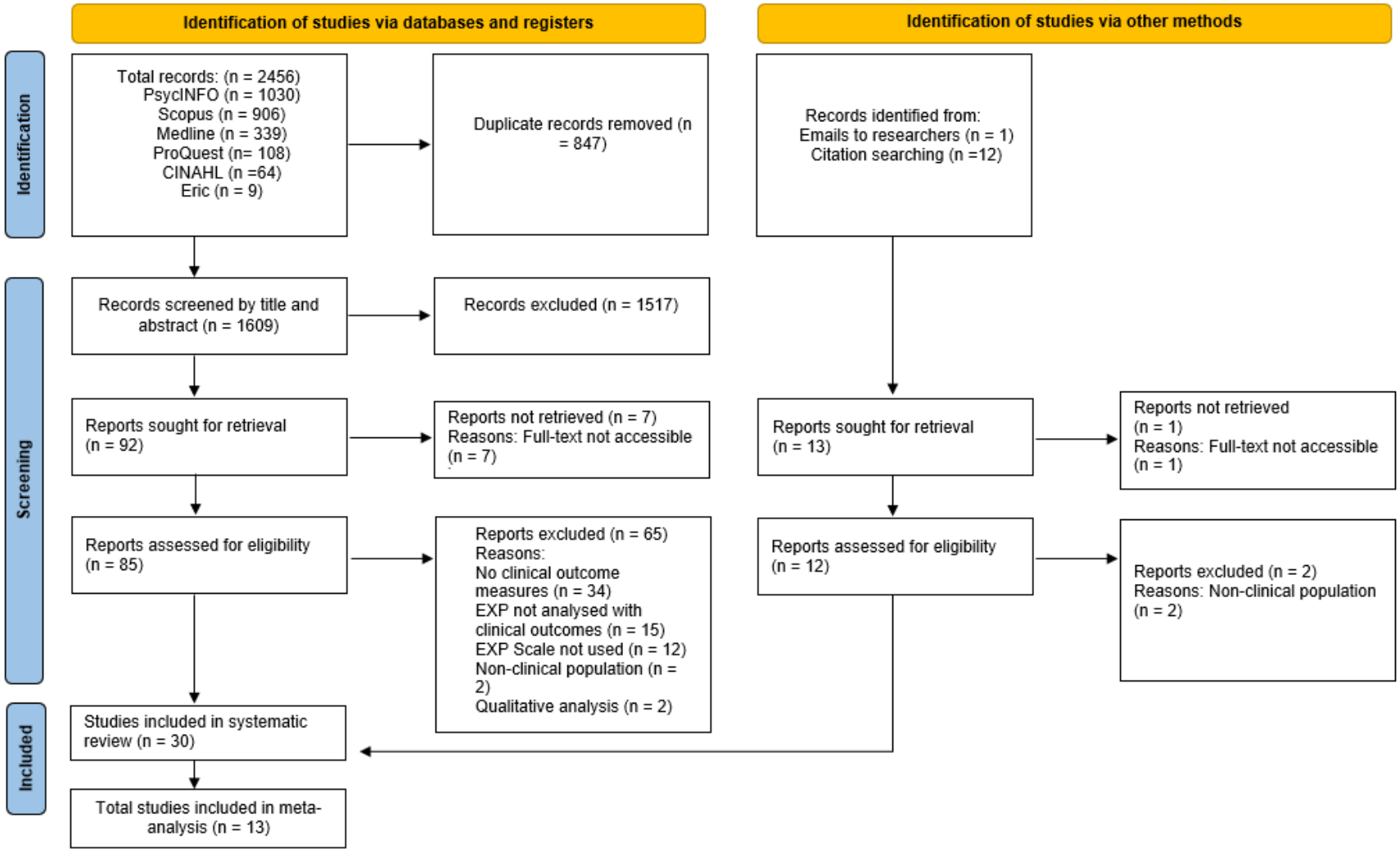
exported all 2456 identified papers to the desktop reference manager software Rayyan (Ouzzani et al., 2016). The duplicator detector was used to highlight potential duplicates, and all papers highlighted were reviewed. Following this process, 847 duplicates were deleted.

The remaining titles and abstracts (n = 1609) were reviewed by the Lead Reviewer, and a subsample by a collaborating researcher, to determine if they met the inclusion criteria. All articles that met inclusion criteria (92) underwent a full-text screen, of which 72 were excluded, yielding 20 papers remaining for inclusion in the systematic review. The most common reasons for papers being removed at this stage were: the paper had not included a clinical treatment outcome measure; EXP had not been analysed in relation to the outcome measures; an alternative measure to EXP had been used; the full text could not be accessed; and the sample was not of a clinical population (See Appendix D for justification of removal at full text). The authors of the seven papers that could not be accessed were contacted, with only one researcher replying but did not provide the paper requested.

To identify any further articles, the 20 papers were reviewed using forward and backward citation searching. In addition, researchers who had contributed more than one article to the research area were contacted for any further research that could be included in the study. From these two processes another 10 papers were included, yielding a total of 30 papers in the systematic review.

Figure 1

PRISMA Diagram.



General Study Characteristics

Table 4 provides a summary of the general characteristics of the studies included in the systematic review.

Of the 30 studies, 21 had been published in peer-reviewed journals. A total of 18 studies were conducted in Canada, 4 in the United States, four in Portugal, two studies from the United Kingdom, and a single study from each from Israel and Germany. According to the EPHPP definitions, 16 studies used secondary data analysis of Controlled Clinical Trials, seven studies used secondary data analysis of Randomised Control Trials, and four studies used secondary data analysis of cohort designs. There were three studies that did not use secondary data comprising a single Randomised Control Trial, a cohort design, and a single case experimental design. All but three studies took place in outpatient settings with the remaining studies taking place in university settings.

The number of participants in each study ranged from 1 to 85. The total number of participants across all studies was 1037, however 19 of the papers used overlapping datasets, and therefore once each participant had only been counted once there were 548 unique participants across all studies. All studies used adult samples only. Three studies had female only participants.

Intervention Characteristics

The majority of interventions offered were for the treatment of depression. Only five studies did not include the treatment of depression: three aimed to treat trauma, and single studies for treating grief and to resolve interpersonal difficulties.

There was a range of interventions delivered across the studies. The most commonly used intervention was emotion focussed therapy (k = 10), followed by cognitive behaviour therapy (k = 8), process experiential therapy (k = 8) and client-centred psychotherapy (k = 5). There were three that had been studied twice: interpersonal psychotherapy, process experiential-emotion focused psychotherapy and psychodynamic psychotherapy. A number of therapies were studied only once: Gestalt therapy, brief dynamic psychotherapy, person-centred experiential therapy, psychoanalytic

therapy, grief constructivist therapy, psychotherapy (undefined), emotion focused therapy for trauma, psychodynamic interpersonal therapy, and perception focused experiential therapy.

Study Characteristics – Overlapping Datasets

Of the included studies, 19 used overlapping datasets (Table 5). Of the 30 included studies, 11 used data collected from the York University Psychotherapy Depression Projects, (Goldman et al., 2006; Greenberg & Watson, 1998), with six papers using data from the 'York I' depression study (Greenberg & Watson, 1998) and five papers combining data from the 'York I' and 'York II' (Goldman et al., 2006) depression studies. The combined sample from the two trials was 74. The 11 papers often studied a sub-sample of these data, albeit with overlapping participants, as sample sizes ranged from two participants to all 74 participants. On a number of occasions studies used the same sessions to explore experiencing and on a couple of occasions studies used experiencing recordings from earlier papers. The number of papers using the York data is underreported here, as initial searches highlighted two papers, Missirilan (2011) and Pos (2006), which used that data, but unfortunately full texts could not be retrieved.

An additional three studies used data from the ISMAI Depression Project (Salgado, 2014, 2019). This project recruited 50 participants. The three papers also had overlapping participants, with sample sizes ranging from two to 50.

The dataset from the Watson et al. (2003) study was also used in three papers, two of which used all 66 participants data whilst the remaining paper used data from 40 participants. There were also two studies that used overlapping data from the Elkin et al. (1989) dataset and another two studies that used overlapping data from the Paivio et al. (2004) dataset.

In total, there were only 11 studies that did not have overlapping datasets.

Table 4

Characteristics of the Studies Included in the Systematic Review.

Authors	Publication Status	Country	Study Design	Clinical Setting	Treatment Focus	Therapy Types	Sample Size	Sample Gender (Men)	Sample Age (STD)	Subsumed Dataset
Anderson et al. (2022)	Published	USA	RCT	University	Depression	Any chosen method	45	18	Mean 19.1	No
Fisher et al. (2019)	Published	Israel	Secondary analysis of Cohort	Outpatients	Depression	PP	18	6	Mean 42.6 (13.71)	No
Goldman et al. (2005)	Published	Canada	Secondary analysis of CCT	Outpatients	Depression	CCP and PET	35	10	Mean 40.74 (11.9)	Yes
Greenberg (1983)	Published	Canada	Cohort	University and outpatients	Interpersonal	GT	28	7	Range of 21 - 52	No
Grooh (1993)	Unpublished	USA	Secondary analysis of Cohort	Outpatients	Depression	BDP	34	12	Range of 20 - 80	No
Hakim (2010)	Unpublished	Canada	Secondary analysis of CCT	Outpatients	Depression	IPT, PET, CBT	85	20	Mean 35.72	Yes
Harrington et al. (2021)	Published	Canada	Secondary analysis of Cohort	Outpatients	Trauma	EFT	45	18	Mean 46 (13)	Yes
Isgar (2024)	Unpublished	UK	Secondary analysis of RCT	Outpatients	Depression	PCET and CBT	40	22	Mean 41.23 (12.15)	No
Jackson (2013)	Unpublished	Canada	Secondary analysis of CCT	Outpatients	Depression	EFT and IPT	56	16	Mean 39.25 (10.29)	Yes

Authors	Publication Status	Country	Study Design	Clinical Setting	Treatment Focus	Therapy Types	Sample Size	Sample Gender (Men)	Sample Age (STD)	Subsumed Dataset
Klug et al. (2021)	Published	Germany	Secondary analysis of RCT	Outpatients	Depression	PaT, PP and CBT	67	18	Details not known	No
Levitt et al. (2000)	Published	USA	Secondary analysis of CCT	Outpatients	Depression	PET	2	1	Range of early 30's to early 60's	Yes
Malin and Pos (2015)	Published	Canada	Secondary analysis of CCT	Outpatients	Depression	CCP and EFT	30	8	Range of 26 - 63	Yes
Pereira et al. (2018)	Published	Portugal	Secondary analysis of CCT	Outpatients	Depression	EFT	21	Unknown	Unknown	Yes
Pinheiro et al. (2022)	Published	Portugal	Secondary analysis of two single cases from a cohort study	Outpatients	Grief	GCT	2	0	20 - 50	No
Pinheiro et al. (2021)	Published	Portugal	Secondary analysis of CCT	Outpatients	Depression	EFT and CBT	50	8	Mean 36.18 (9.7)	Yes
Pinheiro et al. (2018)	Published	Portugal	Secondary analysis of a single case from a CCT	Outpatients	Depression	EFT	1	0	Early 40's	Yes
Pole (1999)	Unpublished	USA	Single case design	Outpatients	Depression	Psychotherapy	3	0	21, 30 and 41	No

Authors	Publication Status	Country	Study Design	Clinical Setting	Treatment Focus	Therapy Types	Sample Size	Sample Gender (Men)	Sample Age (STD)	Subsumed Dataset
Pos et al. (2017)	Published	Canada	Secondary analysis of CCT	Outpatients	Depression	CCP and PET/EFT	32	11	Mean 37 (8.9)	Yes
Pos et al. (2009)	Published	Canada	Secondary analysis of CCT	Outpatients	Depression	PET	74	25	Mean 39.93 (10.96)	Yes
Pos et al, (2003)	Published	Canada	Secondary analysis of CCT	Outpatients	Depression	PET	34	9	Mean 39.64 (11.97)	Yes
Ralston (2006)	Unpublished	Canada	Secondary analysis of RCT	Outpatients	Trauma	EFTT	30	13	Mean 43.6 (12)	Yes
Robichaud (2004)	Unpublished	Canada	Secondary analysis of CCT	Outpatients	Trauma	EFT	37	8	Mean 38 (11.32)	No
Rudkin et al. (2007)	Published	UK	Secondary analysis of RCT	Outpatients	Depression	CBT and PIT	8	2	Range 22 - 57	No
Singh et al. (2021)	Published	Canada	Secondary analysis of CCT	Outpatients	Depression	EFT	14	3	Mean 39.9 (11.5)	Yes
Toukmanian et al. (2010)	Published	Canada	Secondary analysis of CCT	University	Depression	PFET	19	2	Mean 23.5, range 19 - 28	No
Watson and Bedard (2006)	Published	Canada	Secondary analysis of RCT	Outpatients	Depression	CBT and PET	40	13	Mean 41.13 (9.82)	Yes
Watson and Greenberg (1996)	Published	Canada	Secondary analysis of CCT	Outpatients	Depression	CCP and PET	36	11	Mean 40.24 (11.10)	Yes

Authors	Publication Status	Country	Study Design	Clinical Setting	Treatment Focus	Therapy Types	Sample Size	Sample Gender (Men)	Sample Age (STD)	Subsumed Dataset
Watson et al. (2011)	Published	Canada	Secondary analysis of RCT	Outpatients	Depression	PE-EFT and CBT	66	22	Mean 41.52 (10.82)	Yes
Wong (2016)	Unpublished	Canada	Secondary analysis of CCT	Outpatients	Depression	CCP and EFT	55	20	Mean of experientially distant group 41.5 (10.59). Mean of experientially engaged 33.9 (7.63)	Yes
Wong (2023)	Unpublished	Canada	Secondary analysis of RCT	Outpatients	Depression	PE-EFT and CBT	66	22	Mean 42.05 (11.22)	Yes

Note. STD = Standard Deviation; RCT = Randomised Clinical Trial; CCT = Controlled Clinical Trial; EFT = Emotion Focused Therapy; PP = Psychodynamic Psychotherapy; CCP = Client-Centred Psychotherapy; GT = Gestalt Therapy; BDP = Brief Dynamic Psychotherapy; IPT = Interpersonal Psychotherapy; PET = Process Experiential Therapy; CBT = Cognitive Behavioural Therapy; PCET = Person-Centred Experiential Therapy; PaT = Psychoanalytic Therapy; GCT = Grief Constructivist Therapy; EFTT = Emotion Focused Therapy for Trauma; PIT = Psychodynamic Interpersonal Therapy; PFET = Perception Focused Experiential Therapy; PE-EFT = Process Experiential-Emotion Focused Psychotherapy; ICC = intraclass correlation.

Table 5

Summary of Studies with Overlapping Datasets.

Original Study	Overlapping Datasets	Timing of EXP Measurement
York I study (Greenberg & Watson, 1998)		
	Goldman et al. (2005) + 1 additional participant	Middle 20-minute segment from session 2 and three sessions from the second half of therapy.
	Levitt et al. (2000)	Every time a metaphor was spoken.
	Pos et al. (2017)	Emotion episodes from sessions 2 and 3, and emotion episodes from two sessions between session 4 and the fourth from last session.
	Pos et al, (2003)	Emotion episodes in the second session and penultimate session.
	Singh et al. (2021)	The two sessions before and one session after the sudden gain session.
	Watson and Greenberg (1996)	Randomly selected transcripts from sessions 6 - 15.
York I and II studies (Goldman et al., 2006 and Greenberg & Watson, 1998)		
	Hakim (2010) – Also used Elkin et al. (1989) data	Sessions 2 and 3, and two sessions from mid-therapy.
	Jackson (2013) – Also used Elkin et al. (1989) data	Session 2 and 3, and two sessions from mid-therapy.
	Malin and Pos (2015)	Archival recordings from Pos et al. (2009), which used emotion episodes from the fourth session to the fourth from last session.
	Pos et al. (2009)	Emotion episodes from session 2, two mid-therapy sessions and two late therapy sessions.
	Wong (2016)	Archival recordings from Pos et al. (2009) using working therapy session data.
Paivio et al. (2004)		

Original Study	Overlapping Datasets	Timing of EXP Measurement
	Harrington et al. (2021)	Selection of sessions from 3 - 6 and from 7 – 11.
	Ralston (2006)	Emotion episodes in session 4, one session from sessions 7 - 11 and one session from sessions 12 – 16.
Salgado et al. (2014)		
	Pereira et al. (2018)	Each emotion episode in session 1 and session 16.
	Pinheiro et al. (2021)	Each emotion episode in sessions 1, 4, 8, 12, 16.
	Pinheiro et al. (2018)	Each emotion episode from session 1 to session 16.
Watson et al. (2003)		
	Watson and Bedard (2006)	Middle 20-minute segment of session 3, one session from sessions 6 –10 and one session from sessions 11–15.
	Watson et al. (2011)	Midde 20-minute segment from one session between sessions 2 – 4 and two sessions from sessions 5 – 15.
	Wong (2023)	20-minute segments from sessions 3, 9 and 15.

Methodological Quality

Hakim et al. (2010) and Jackson et al. (2013) used secondary data from multiple sources. On those occasions, the papers have been scored using the lowest score across datasets. Of the 30 papers, 9 were rated as weak overall, 19 were rated moderate and only 2 were established as strong quality papers (Appendix E).

The component that was most often rated as weak was selection bias (n = 24). Most participants were self-referrals, with some participants being referred from clinics. Important differences between groups were not routinely reported and it was not always clear what confounders had been controlled for, and as a result 13 papers received a weak rating for this component. The process of blinding participants was not clear in 21 papers, which resulted in their scoring moderate for that component, whilst a further 4 papers were also not transparent regarding

researcher blinding, resulting in 4 articles being scored as weak for this component. General areas of strength across papers were data collection methods, with all papers using reliable and valid outcome measures. Low rates of withdrawals and dropouts were a general strength, with 22 papers having strong retention rates, and the other eight having moderate retention rates.

Meta-Analysis Results

Table 6 explains why papers were excluded from the meta-analysis. Of the 17 papers excluded, 13 had overlapping datasets, three were single-case experimental designs, and one did not have the required type of data to be included in the analyses.

Table 6

Reason for exclusion of papers from the meta-analyses.

Author (Date)	Reason for Exclusion from Meta-Analysis
Goldman et al. (2005)	Overlapping data from York studies.
Jackson (2013)	Overlapping data from York studies.
Klug et al. (2021)	Data in its published form could not be used for meta-analysis. The authors of the paper were unable to provide the necessary information when contacted.
Levitt et al. (2000)	Overlapping data from York studies.
Malin and Pos (2015)	Overlapping data from York studies.
Pereira et al. (2018)	Overlapping data from ISMAI study.
Pinheiro et al. (2022)	Single case design.
Pinheiro et al. (2018)	Single case design.
Pole (1999)	Single case design.
Pos et al. (2017)	Overlapping data from York studies.
Pos et al, (2003)	Overlapping data from York studies.
Ralston (2006)	Overlapping data from Paivio et al. (2004) study.
Singh et al. (2021)	Overlapping data from York studies.

Author (Date)	Reason for Exclusion from Meta-Analysis
Watson & Bedard (2006)	Overlapping data from Watson et al. (2003) study.
Watson & Greenberg (1996)	Overlapping data from York studies.
Wong (2016)	Overlapping data from York studies.
Wong (2023)	Overlapping data from Watson et al. (2003) study.

There were 13 papers deemed suitable for the meta-analyses (Table 7). It should be noted that although the full Hakim (2010) paper was included in the systematic review, only sub-samples were included in the meta-analyses. This is because, of the three therapies originally reviewed, the PET data came from the York studies, therefore only the IPT and CBT data were reviewed in the current study. These have been treated as separate datasets for the purpose of the current study as they had been analysed separately in the original paper. 'Hakim (a) (2010)' refers to the IPT dataset, while 'Hakim (b) (2010)' refers to the CBT dataset. Additionally, the Isgar (2024) paper included experiencing data for CBT and separately for PCET, which allowed for them to be treated as two separate datasets. 'Isgar (a) (2024)' refers to the CBT dataset, while 'Isgar (b) (2024)' refers to the PCET dataset. As a result of the process described above, there were 15 datasets (from 13 papers) which were deemed appropriate for the meta-analysis.

Table 7

Characteristics of the Studies Included in the Meta-Analyses.

Authors	Modal EXP	Peak EXP	Averaged Modal and Peak EXP	Early Therapy Coding	Middle Therapy Coding	Late Therapy Coding	Primary Outcome Measure and Administration	Reliability	Percent of Study Data Checked
Anderson et al. (2022)	X	X			X		OQ-45 at the end of therapy.	ICC = .73 for EXP Modal and ICC = .79 for EXP Peak.	0%. Raters only checked reliability on practice tapes.
Fisher et al. (2019)	X			Averaged across therapy.			BDI-II at the end of therapy.	ICC = .93.	Details not included.
Greenberg (1983)	X			Averaged across therapy.			CRBS and TC each session.	Not applicable.	Not applicable.
Grooh (1993)	X			Averaged across therapy.			Global composite: TC, SCL-90, GAS, OCR and BPRS. Administered at the end of therapy and at 6-month follow-up.	ICC 0.64	100%
Hakim (a) (2010)	X			X	X		BDI at the end of therapy.	Inter-rater reliability of 0.74.	33%
Hakim (b) (2010)	X			X	X		BDI at the end of therapy.	Inter-rater reliability of 0.80.	33%
Harrington et al. (2021)		X			X		IIP at the end of therapy.	Kappa 0.77	33%

Authors	Modal EXP	Peak EXP	Averaged Modal and Peak EXP	Early Therapy Coding	Middle Therapy Coding	Late Therapy Coding	Primary Outcome Measure and Administration	Reliability	Percent of Study Data Checked
Isgar (a) (2024)	X	X		X	X		PHQ-9 at the end of therapy.	ICC = 0.94 for EXP Modal and ICC = 0.93 for EXP Peak.	0%. Raters only checked reliability on practice tapes.
Isgar (b) (2024)	X	X		X	X		PHQ-9 at the end of therapy.	ICC = 0.94 for EXP Modal and ICC = 0.93 for EXP Peak.	0%. Raters only checked reliability on practice tapes.
Pinheiro et al. (2021)		X		Averaged across therapy.			BDI-II at the end of therapy.	Kappa 0.8 to 0.88 depending on pairings of raters.	100%
Pos et al. (2009)			X	X	X	X	BDI at the end of therapy.	Kappa 0.79.	33%
Robichaud (2004)	X	X		X			SCL-90 at the end of therapy.	Kappa 0.85.	33%
Rudkin et al. (2007)	X	X		X	X		BDI at the end of therapy.	ICC 0.85.	33%
Toukmanian et al. (2010)		X			X	X	BDI at the end of therapy.	Inter-rater reliability 0.89.	52.6%
Watson et al. (2011)	X	X		X	X		BDI at the end of therapy.	ICC 0.83.	69%

Note. ICC = Intra Class Correlation; OQ-45 = Outcome Questionnaire-45; BDI-II = Beck Depression Inventory-II; BDI = Beck Depression Inventory; CRBS = Conflict Resolution Box Scale; TC = The Target Complaints; SCL-90 = Symptom Checklist-90; GAS = Global Assessment Scale; OCR = Overall Change Rating; BPRS = The Brief Psychiatric Rating Scale; IIP = Inventory of Interpersonal Problems; PHQ-9 = Patient Health Questionnaire 9-Item.

Of the 15 datasets, 11 analysed modal EXP scores in relation to outcomes, 9 used peak EXP scores and a single paper used an average of combined modal and peak EXP scores. Most datasets (8) used the BDI or BDI-II as the primary outcome measure. Of the 15 datasets, eight coded depths of experiencing in early therapy, 10 used middle/working therapy, while two used late therapy sessions. There were four papers yielding codings at early, middle and late therapy sessions and averaged EXP across all stages.

Of the 15 datasets, 14 administered outcome measures at the end of therapy and a single paper administered outcome measures during each session. The Grooh study (1993) administered outcome measures at the end of therapy and also at 6-month follow-up. For the purpose of the current study, only the end of treatment outcome measures were used.

All but one of the studies completed reliability checks between raters. There was a range in percentage of data checked. Ten of the papers completed reliability checks of the data included in their studies, with the percent of their data checked ranging from 33% to 100%. Three datasets only completed reliability checks using practice tapes. Two papers did not include details on percent of reliability checks completed.

Association between Modal EXP and Outcomes

Figure 2

Forest Plot for Relationship between Modal EXP and Treatment Outcomes.

Study name	Correlation	CI Lower limit	CI Upper limit
Hakim CBT dataset (2010)	0.06	-0.33	0.43
Anderson et al. (2022)	0.07	-0.24	0.36
Watson et al. (2011)	0.13	-0.12	0.36
Rudkin et al. (2007)	0.14	-0.73	0.83
Fisher et al. (2019)	0.14	-0.39	0.59
Isgar PCET dataset (2024)	0.22	-0.28	0.62
Hakim IPT dataset (2010)	0.29	-0.11	0.61
Isgar CBT dataset (2024)	0.29	-0.21	0.67
Grooh (1993)	0.33	-0.02	0.61
Robichaud (2004)	0.35	0.02	0.61
Pos et al. (2009)	0.45	0.24	0.62
Greenberg (1983)	0.89	0.77	0.95
Combined Effect Size	0.33	0.10	0.52

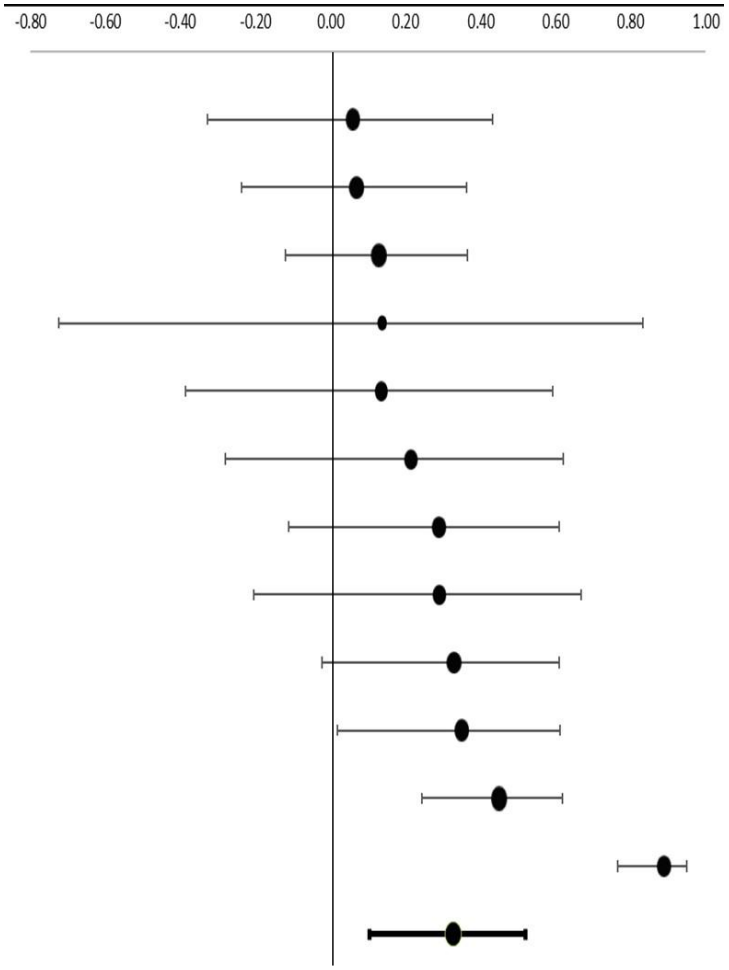


Figure 2 shows the effect sizes of modal EXP scores and therapeutic outcomes from 12 datasets comprising 407 participants, ordered according to increasing effect sizes. There were wide confidence intervals across studies, with only three studies' confidence intervals not crossing the 'no effect' line. Collectively the results suggested a moderate association between level of modal experiencing and treatment outcomes (r = 0.33, 95% CI [0.10, 0.52], p = 0.001), although the 'combined effect size' confidence intervals suggest the association could range from small to large.

Visual inspection of the forest plot identified wide confidence intervals, indicating within-study variability. Heterogeneity was also assessed using the Q-statistic, I^2, and Tau. The Q-statistic suggested heterogeneity, Q = 39.77, $p < 0.001$, although this is likely impacted by the small number of studies. The I^2 analysis suggested a high proportion (72.3%) of variability between studies not due to sampling error. Tau analysis suggested low variability between studies (Tau^2 = 0.09, Tau = 0.30).

Moderator Analysis

Moderator analyses were not conducted for two reasons: 1) There were not enough new papers with the required variables to repeat moderator analyses completed previously, for example, no new papers used therapist rated outcomes and 2) there were insufficient numbers of papers in categories for novel analyses (for example presenting problem could not be compared as there were only two papers that did not study depression).

Sensitivity Analysis

To explore the robustness of the findings, by altering methodological decisions, three sensitivity analyses were conducted. One of the studies (Greenberg, 1983) used both client and therapist outcome measures, which were correlated with EXP. This was different to the other papers, which only used client outcome measures. After removing this study, the association between modal EXP and therapeutic outcomes reduced (r = 0.25, 95% CI [0.15, 0.35], p = 0.001). Heterogeneity reduced following this process, Q = 8.18, p = 0.612, I^2 (0%) and Tau^2 = 0.001, Tau = 0.001.

A second sensitivity analysis was conducted by removing any papers which did not study working therapy, of which there was one (Robichaud, 2004). Following the exclusion of this dataset, the strength of association only minimally changed ($r = 0.33$, 95% CI [0.08, 0.54], $p = 0.002$).

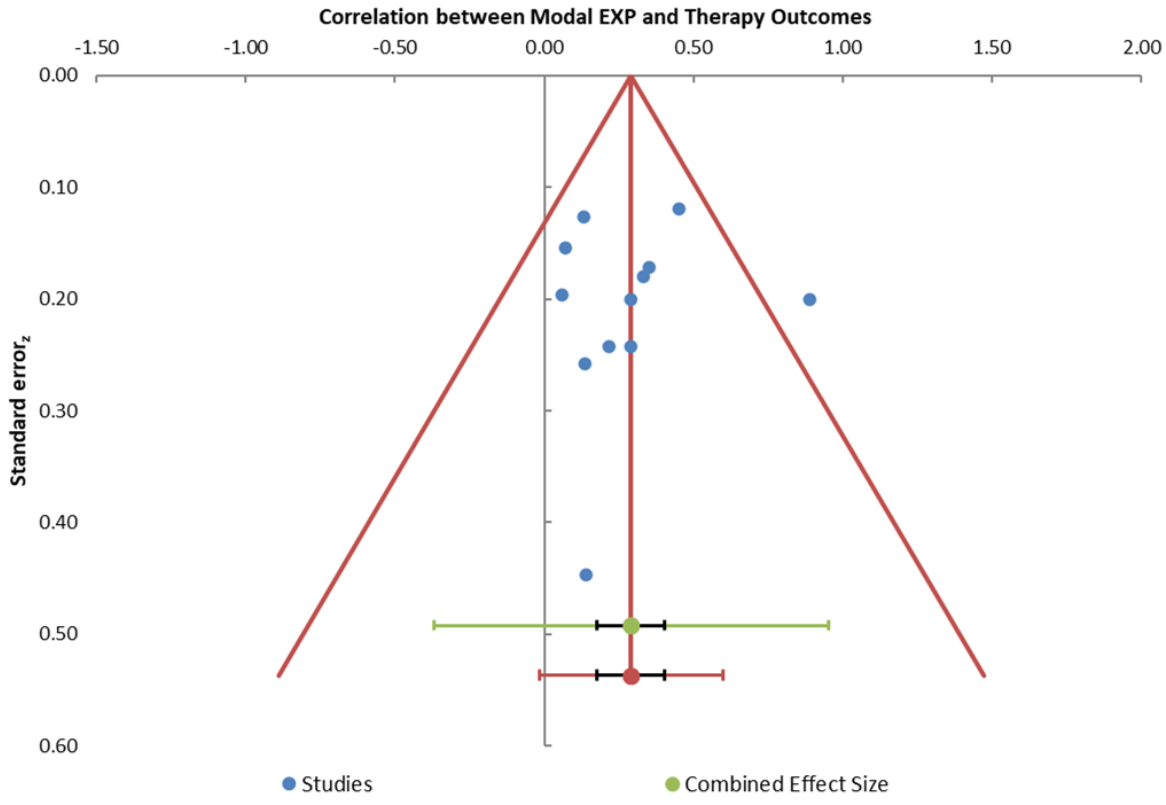
A third sensitivity analysis was conducted by removing papers which did not study depression (Greenberg, 1983; Robichaud, 2004). Following the exclusion of these datasets, the strength of association reduced ($r = 0.24$, 95% CI [0.13, 0.35], $p < 0.0001$).

Publication Bias

Visual analysis of the funnel plot (Figure 3) suggested some asymmetry, indicating potential publication bias. However, the trim-and-fill method did not support this interpretation as zero studies were imputed. Egger's Regression Test (Egger et al., 1997) was non-significant ($t = 0.04$, $p = 0.97$). The result of the Rosenthal (1979) fail-safe N analysis suggested a further 159 additional studies would be required to overturn the significant result.

Figure 3

Level of Publication Bias Assessed Through the Funnel Plot.



Association between Peak EXP and Outcomes

Figure 4

Forest Plot for Relationship between Peak EXP and Treatment Outcomes.

Study name	Correlation	CI Lower limit	CI Upper limit
Isgar PCET dataset (2024)	0.01	-0.46	0.48
Rudkin et al. (2007)	0.07	-0.76	0.81
Anderson et al. (2022)	0.18	-0.13	0.46
Harrington et al. (2021)	0.23	-0.09	0.50
Toukmanian et al. (2010)	0.30	-0.21	0.68
Robichaud (2004)	0.34	0.01	0.61
Isgar CBT dataset (2024)	0.38	-0.11	0.72
Watson et al. (2011)	0.39	0.16	0.58
Pinheiro et al. (2021)	0.44	0.17	0.64
Pos et al. (2009)	0.45	0.24	0.62
Combined Effect Size	0.34	0.24	0.42

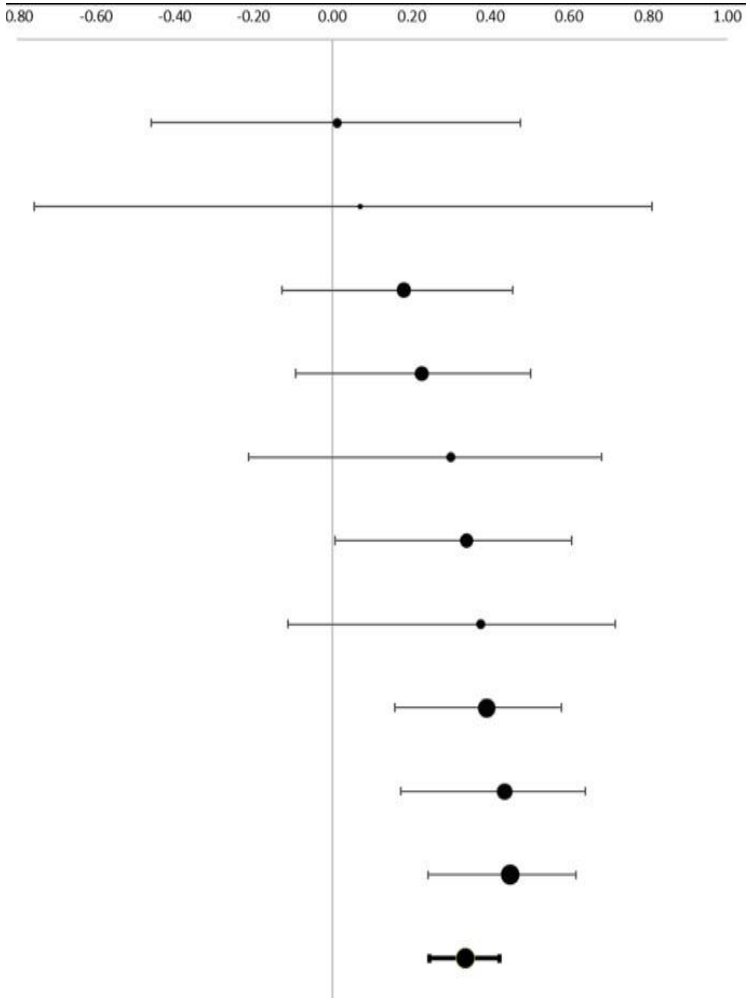


Figure 4 provides a visual of the effect sizes of peak EXP scores and therapeutic outcomes (ordered according to increasing effect sizes) from 10 datasets comprising 381 participants. Individually, there were wide confidence intervals across studies and over half of the studies' confidence intervals were within the 'no effect' range. However, the results suggested an overall moderate association between level of peak experiencing and treatment outcomes (r = 0.34, 95% CI [0.24, 0.42] $p < 0.0001$).

Visual inspection of the forest plot identified wide confidence intervals, indicating within-study variability. Heterogeneity was also assessed using the Q-statistic, I^2 and Tau. The Q-statistic did not suggest heterogeneity, $Q = 6.32$, $p = 0.71$, although this is likely impacted by the small number of studies. The I^2 analysis suggested a low proportion (0%) of variability between studies not due to sampling error. Tau analysis suggested low variability between studies ($Tau^2 = 0.001$, Tau = 0.001).

Moderator Analysis

Moderator analyses were not conducted, due to the same reasons as outlined in the first meta-analysis and also because analyses suggested low variability between studies.

Sensitivity Analysis

To assess the robustness of the findings to methodological variances, two sensitivity analyses were conducted. Similar to the sensitivity analysis with modal data, the Robichaud (2004) paper was removed as that paper solely explored early therapy. The effect size remained the same (r = 0.34, 95% CI [0.23, 0.43], $p = 0.0001$). One of the studies (Pinheiro et al., 2021) took an average of EXP from early, middle and late sessions whereas the other papers, excluding Robichaud (2004), all used working sessions. After removing the study, the association reduced by 0.02 (r = 0.32, 95% CI [0.22, 0.42], $p = 0.0001$).

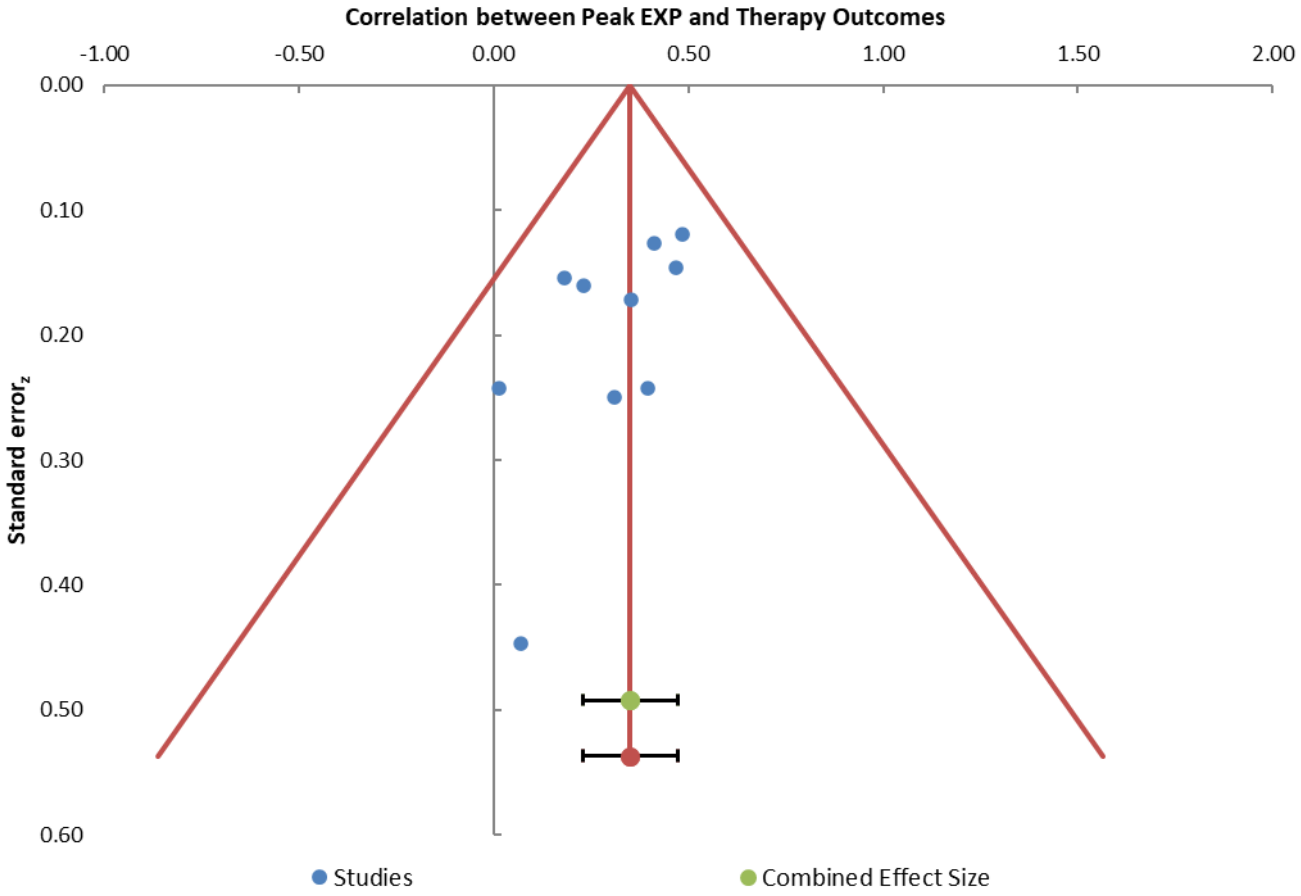
Publication Bias

Visual inspection of the funnel plot (Figure 5) suggested symmetry, thus suggesting there was not publication bias. The trim-and-fill method corroborated this interpretation as zero studies

were imputed. Egger's Regression Test (Egger et al., 1997) was also conducted, which was non-significant (t = -1.99, p = .08). Rosenthal's (1979) fail-safe N suggested 114 additional studies would be required to overturn the significant result.

Figure 5

Level of Publication Bias Assessed Through the Funnel Plot.



Discussion

The current review aimed to update and refine the Pascual-Leone and Yeryomenko's (2016) meta-analysis by conducting a more thorough and transparent search, completing a systematic review, including quality appraisal of individual studies, together with a meta-analysis. The study also aimed to understand the strength of the relationship between depth of experiencing (using the EXP Scale) and therapeutic outcomes, and variables which may influence this relationship.

The current review offered a much more comprehensive search of suitable papers. The systematic review included a total of 30 studies, while 15 datasets from 13 papers were deemed

suitable to include in the meta-analysis. A feature of the current review, not included previously, was an assessment of the quality of studies. The results suggested the vast majority of papers (19) were of moderate quality. Only two papers were rates as strong quality, while nine were weak. Selection bias and transparency of blinding procedures were the elements which most commonly scored lowest across studies.

Meta-Analyses

The current study found a moderate association between both modal and peak EXP and therapeutic outcomes, which is larger than the association found by Pascual-Leone and Yeryomenko (2016). Publication bias was non-significant and the fail-safe N was 159 studies for modal analysis and 114 for peak analysis, which suggests a highly unlikely undiscovered number of papers would be required to overturn the findings.

Results suggested heterogeneity was high within the modal EXP meta-analysis, which was the same finding as the previous meta-analysis, and low within the peak EXP analysis which is different to the previous review. The high heterogeneity in the modal analysis appeared to be largely influenced by the Greenberg (1983) study. The low heterogeneity in the peak analysis was possibly due to the accuracy of the current study applying the inclusion and exclusion criteria, and because of the effect sizes used (the previous study pooled correlations between modal and peak EXP ratings, across stages of therapy and across different outcome measures).

Sensitivity Analyses

Sensitivity analyses were conducted to assess the robustness of findings with changes to methodological decisions. Regarding modal data, the strength of the association between EXP and therapy outcomes remained significant even after the removal of papers that used therapist outcomes (Greenberg, 1983), explored early sessions (Robichaud, 2004), or that did not explore depression (Greenberg, 1983; Robichaud, 2004). Heterogeneity significantly reduced with the removal of the Greenberg paper.

Regarding peak data, the strength of association remained after removing papers that explored early sessions (Robichaud, 2004) and papers that pooled EXP scores across early, middle, and late sessions (Pinheiro et al., 2021).

Given there were no significant changes in strength of association, it can be concluded that the findings were robust to methodological decisions.

Strengths and Limitations

The overall number of studies in the meta-analyses is a limitation due to the lower power in the analyses and the subsequent restrictions on the follow-up analyses that could be conducted, including being unable to conduct moderator analyses planned a priori. This was significantly influenced by the removal of 17 papers following the systematic review, due to the approach taken to managing overlapping datasets (selecting the primary paper with the greatest number of participants, selecting the most relevant outcome measure to the presenting problem and using working therapy data where possible). The advantage of this approach is that participants data is not 'double counted' meaning the meta-analysis conforms to the independence of observations requirement, however, by using only one effect size, there is a risk of introducing potential bias when deciding which effect size to include. Future research could include a greater number of effect sizes, thus enriching the dataset and supporting the investigation of heterogeneity. This could be achieved by combining multiple effects into one effect size per study by pooling the effect and standard error from a mini-meta-analysis of effects within a study or by conducting multiple meta-analyses so only one effect size is included from a study in any one comparison.

Although the removal of 17 papers is a limitation, Cumming (2014) has argued for the importance and usefulness of conducting small scale meta-analyses as a method of improving psychological research and for building a cumulative research base in an understudied area. The current paper only overlapped on seven of the studies included in the previous meta-analysis. So, although the overall paper number was only three greater in number in the current analysis than the previous analysis, the included papers in the current analysis were different in that there were six

papers in the current analysis that were not included in the previous analysis, and three papers from the old analysis were not included in the current analysis. By removing some papers which were previously included, but did not necessarily fit the research question, the quality of the review has improved.

Although 20 of 30 included papers were identified through the agreed search terms and six databases, a high proportion of included studies (10 of 30) were identified through hand searching. This may suggest that the search criteria were not sensitive enough to identify all relevant papers. The search terms were limited by not including the term 'EXP' due to the substantial increase in number of results which would have made the paper review stage incredibly timely. Of the studies identified through hand searching, six of 10 were unpublished, while one further paper had only been published as a poster presentation. Future research may benefit from the inclusion of more than one grey literature database and the inclusion of the search term 'EXP'.

According to the results of the EPHPP, the general quality of papers was poor, with only 2 rated strong, 19 moderate and 9 weak. The main category that was rated weakly was 'selection bias'. Most papers did not use randomised sampling procedures, or recruit participants that were generalisable to the general population. However, the samples appeared appropriate for the research questions being asked as they contained people who were accessing therapy. It may be possible that the EPHPP was not the most suitable measure of quality given the number of papers that used secondary data (which is not appraised on the EPHPP) and the number of unpublished papers (also not appraised on the EPHPP). Given the repeated use of the York datasets for doctoral theses and the number of these papers which have remained unpublished, a tool that was specifically designed to measure quality in secondary data may have been more appropriate. However, it should be recognised that the completion of quality appraisal was an important part of this review, given the last review did not complete this. This process highlighted general areas of strengths and weaknesses across all papers.

Analyses suggested that there was no publication bias in the current review. The author has taken all steps possible to locate all relevant papers, including searching the grey literature, contacting researchers, and completing forward and backward citation searching.

Research and Clinical Implications

The review highlights areas in both research and clinical practice that could be enhanced. With regards to research, the current paper highlights the overuse of data derived from the York studies. From the literature searches, there were a minimum of 11 papers that had used subsamples or the full sample of participants data from the York studies. This is an underrepresentation of the total number of papers using the York data as two papers (Missirilan, 2011; Pos, 2006) that used the York data had to be excluded as their full text could not be retrieved. On some occasions, the same EXP codings from the York data had been used across published papers. Unfortunately, this trend appears to have transferred to other datasets, including Watson et al. (2003), Paivio et al. (2004) and Salgado et al. (2014), which have now all been explored multiple times using the EXP Scale. Continued use of relatively small datasets by the same research groups investigating the same phenomenon to conceptually replicate findings are not the best strategy for advancing knowledge as they are likely to be more vulnerable to a range of biases (Wiggins & Christopherson, 2019). Simons (2014) has suggested that the only method of testing the reliability of findings and thus advancing science is through direct replications of methodology using independent participant samples and researchers. Therefore, the implications of such findings are two-fold; it is advised that there should be no further exploration of depth of experiencing using the participants from the York studies unless there was a clear rationale, and secondly, there is a need for further research using independent data in the field. Given all but one study were conducted in Europe or North America, it would be advantageous if new studies used datasets from other regions to enhance the generalisability of findings.

Future research could explore EXP alongside other measures of experiencing. Sonderland et al. (2023) recently studied emotional changes processes in psychotherapy, and within that research

explored depth of experiencing. They identified a number of measures used to assess experiencing. The paper provided an overall effect size, for all papers. Future research could initially assess whether the different measures are exploring the same concepts, and if so, they could then explore whether the different measures result in different effect sizes. Given the number of additional papers and participants that this method would utilise, the conclusions from any analyses would be much more robust than the current paper.

With regards to clinical implications, given the repeated finding of an association between experiencing and therapeutic outcomes, clinicians should find a way of monitoring this process within therapy. Additionally, training of clinicians should incorporate the idea of developing exploration of feelings in sessions. Furthermore, supervisors should be encouraged to explore depth of experiencing of clients within supervision, upskilling the clinicians to work in the most effective ways for their clients.

Conclusion

A significant number of studies within the EXP literature have repeatedly used the same datasets. This finding suggests that there is a need for researchers to cease re-analysing the data from the York studies and to find novel datasets to test the same theories. The results of a thorough review of the literature examining the relationship between depth of experiencing (as measured using the EXP Scale) and therapy outcomes provides support for the view that there is a relationship between both. This suggests there is a need for clinicians and clients to work together to strive to increase the depth of experiencing of the client when delivering therapy.

References

Amir-Behghadami, M., & Janati A. (2020). Population, intervention, comparison, outcomes and study (PICOS) design as a framework to formulate eligibility criteria in systematic reviews. *Emergency Medicine Journal*, *37*(6), 386-387. http://dx.doi.org/10.1136/emermed-2020-209567

*Anderson, T., Stone, S. J., Angus, L., & Weibel, D. T. (2022). Double trouble: Therapists with low facilitative interpersonal skills and without training have low in-session experiential processes. *Psychotherapy Research*, *32*(1), 65-77. https://doi.org/10.1080/10503307.2021.1913293

Angus, L. E., Boritz, T., Bryntwick, E., Carpenter, N., Macaulay, C., & Khattra, J. (2017). The Narrative-Emotion Process Coding System 2.0: A multi-methodological approach to identifying and assessing narrative-emotion process markers in psychotherapy. *Psychotherapy Research*, *27*(3), 253-269. https://doi.org/10.1080/10503307.2016.1238525

Armijo-Olivo, S., Stiles, C. R., Hagen, N. A., Biondo, P. D., & Cummings, G. G. (2012). Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *Journal of Evaluation in Clinical Practice, 18,* 12-18. https://doi.org/10.1111/j.1365-2753.2010.01516.x

Boland, A., Dickson, R., & Cherry, G. (2017). *Doing a systematic review: A student's guide* (2nd ed). London: SAGE Publications Ltd.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. John Wiley & Sons.

Burgess Moser, M. (2012). *The cognitive-affective and behavioural impact of emotionally focused couple therapy* [Doctoral dissertation, Université d'Ottawa/University of Ottawa]. University of Ottawa Online Repository. https://doi.org/10.20381/ruor-6474

Card N. A. (2012). *Applied meta-analysis for social science research*. Guilford.

Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, *10*(1), 101-129. https://doi.org/10.2307/3001666

Cohen, J. (1988). *Statistical power for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cox, N. J. (2008). Speaking Stata: Correlation with confidence, or Fisher's z revisited. *The Stata Journal*, *8*(3), 413-439. https://doi.org/10.1177/1536867X0800800

Cumming, G. (2014). "The new statistics: Why and how." *Psychological Science, 25*(1), 7-29. https://doi.org/10.1177/0956797613504966

Deeks, J. J., Higgins, J. P., Altman, D. G., & Cochrane Statistical Methods Group. (2019). Analysing data and undertaking meta-analyses. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page & V. A. Welch (2nd ed.), *Cochrane handbook for systematic reviews of interventions* (pp. 241-284). The Cochrane Collaboration.

Dreier, M. (2013). Quality assessment in Meta-Analysis. In S. A. R. Doi. & G. M. Williams (Eds.), *Methods of clinical epidemiology* (pp. 213-228). Springer.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629-634. https://doi.org/10.1136/bmj.315.7109.629

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., Docherty, J. P., Fiester, S. J., & Parloff, M. B. (1989). National Institute of Mental Health treatment of depression collaborative research program: General effectiveness of treatments. *Archives of General Psychiatry*, *46*(11), 971-982. https://doi.org/10.1001/archpsyc.1989.01810110013002

Elliott, R., Greenberg, L. S., Watson, J. C., Timulak, L., & Freire, E. (2013). Research on humanistic-experiential psychotherapies. In M.J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.) (pp. 495-538). Wiley.

Falkenström, F., Finkel, S., Sandell, R., Rubel, J. A., & Holmqvist, R. (2017). Dynamic models of individual change in psychotherapy process research. *Journal of Consulting and Clinical Psychology*, *85*(6), 537. https://doi.org/10.1037/ccp0000203

*Fisher, H., Hacohen, N., Shimshi, S., Rand-Lakritz, S., Shapira, K., & Tuval-Mashiach, R. (2020). Ability to move between self-states and emotional experiencing and processing as predictors

of symptomatic change. *Psychology and Psychotherapy: Theory, Research and Practice*, *93*(4), 723-738. https://doi.org/10.1111/papt.12256

Goldman, R. N., Greenberg, L. S., & Angus, L. (2006). The effects of adding emotion-focused interventions to the client-centered relationship conditions in the treatment of depression. *Psychotherapy Research*, *16*(5), 537-549. https://doi.org/10.1080/10503300600589456

*Goldman, R. N., Greenberg, L. S., & Pos, A. E. (2005). Depth of emotional experience and outcome. *Psychotherapy Research*, *15*(3), 248-260. https://doi.org/10.1080/10503300512331385188

*Greenberg, L. S. (1983). Toward a task analysis of conflict resolution in Gestalt therapy. *Psychotherapy: Theory, Research & Practice*, *20*(2), 190-201. https://doi.org/10.1037/h0088490

Greenberg, L., & Watson, J. (1998). Experiential therapy of depression: Differential effects of client-centered relationship conditions and process experiential interventions. *Psychotherapy Research*, *8*(2), 210-224. https://doi.org/10.1080/10503309812331332317

*Grooh, H. R. (1993). *The role of insight and patient involvement in brief psychodynamic psychotherapy outcomes* [Unpublished Doctoral Dissertation, Pacific Graduate School of Psychology]. University Microfilms International. https://www.proquest.com/docview/304113849?pq-origsite=gscholar&fromopenview=true&sourcetype=Dissertations%20&%20Theses

*Hakim, L. Z. (2010). *A comparative analysis of the qualities of therapists' communications in three empirically-supported psychotherapies* [Unpublished Doctoral Dissertation, York University]. Library and Archives Canada. https://www.collectionscanada.gc.ca/obj/thesescanada/vol2/002/NR80596.PDF?is_thesis=1&oclc_number=879940060

*Harrington, S., Pascual-Leone, A., Paivio, S., Edmondstone, C., & Baher, T. (2021). Depth of experiencing and therapeutic alliance: What predicts outcome for whom in emotion-focused therapy for trauma? *Psychology and Psychotherapy: Theory, Research and Practice, 94*(4), 895-914. https://doi.org/10.1111/papt.12342

Harrison, J. K., Reid, J., Quinn, T. J., & Shenkin, S. D. (2017). Using quality assessment tools to critically appraise ageing research: A guide for clinicians. *Age and Ageing, 46*(3), 359-365. https://doi.org/10.1093/ageing/afw223

Hendricks, M. (2009). Experiencing level: An instance of developing a variable from a first-person process so it can be reliably measured and taught. *Journal of Consciousness Studies, 16*(10–12), 129–155. https://www-ingentaconnect-com.sheffield.idm.oclc.org/content/imp/jcs/2009/00000016/f0030010/art00006#

Hendricks, M. N. (2002). Focusing-oriented/experiential psychotherapy. In D. J. Cain (Eds.), *Humanistic psychotherapies: Handbook of research and practice* (pp. 221–251). American Psychological Association. https://doi.org/10.1037/10439-007

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*(7414), 557-560. https://doi.org/10.1136/bmj.327.7414.557

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist*, *41*(2), 159. https://doi.org/10.1037/0003-066X.41.2.159

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological Methods*, *11*(2), 193-206. https://doi.org/10.1037/1082-989X.11.2.193

*Isgar, J. (2024). Clients *emotional processing in psychotherapy: a comparison between CBT & PCET* [Unpublished DClinPsy Dissertation]. University of Sheffield, England, United Kingdom.

*Jackson, S. L. (2013). *Change processes in emotion-focused and interpersonal psychotherapies for depression: A comparative study* [Unpublished Doctoral Dissertation, York University]. Library and Archives Canada. https://central.bac-lac.gc.ca/.item?id=NR88705&op=pdf&app=Library&oclc_number=910984285

Kiesler, D. J., Mathieu, P. L., & Klein, M. H. (1964). Sampling from the recorded therapy interview: A comparative study of different segment lengths. *Journal of Consulting Psychology*, *28*(4), 349. https://doi.org/10.1037/h0047981

Klein, M. H., Mathieu, P. L., Gendlin, E. T, & Kiesler, D. J. (1969). *The Experiencing Scale: A research and training manual* (Vols. 1 and 2). Madison: University of Wisconsin Extension Bureau of Audiovisual Instruction.

*Klug, G., Seybert, C., Ratzek, M., Grimm, I., Zimmermann, J., & Huber, D. (2021). Insight and Outcome in Long-Term Psychotherapies of Depression. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, *68*(1), 54-73. https://doi.org/10.13109/zptm.2021.67.oa10

*Levitt, H., Korman, Y., & Angus, L. (2000). A metaphor analysis in treatments of depression: Metaphor as a marker of change. *Counselling Psychology Quarterly*, *13*(1), 23-35. https://doi.org/10.1080/09515070050011042

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *Annals of Internal Medicine, 151*(4), W-65. https://doi.org/10.7326/0003-4819-151-4-200908180-00136

Luborsky, L., Crits-Christoph, P., Mintz, J., & Auerbach, A. (1988). *Who will benefit from psychotherapy? Predicting therapeutic outcomes.* Basic Books.

Makinen, J. A., & Johnson, S. M. (2006). Resolving attachment injuries in couples using emotionally focused therapy: Steps toward forgiveness and reconciliation. *Journal of Consulting and Clinical Psychology*, *74*(6), 1055. https://doi.org/10.1037/0022-006X.74.6.1055

*Malin, A. J., & Pos, A. E. (2015). The impact of early empathy on alliance building, emotional processing, and outcome during experiential treatment of depression. *Psychotherapy Research*, *25*(4), 445-459. https://doi.org/10.1080/10503307.2014.901572

Marín-Martínez, F., & Sánchez-Meca, J. (2010). Weighting by inverse variance or by sample size in random-effects meta-analysis. *Educational and Psychological Measurement*, *70*(1), 56-73. https://doi.org/10.1177/0013164409344534

Missirilan, T. M. (2011). *A comparative study of the nature of change processes in emotion focused and cognitive-behavioural psychotherapies for depression* [Unpublished Doctoral Dissertation]. York University, Toronto, Canada.

Nair, A. S. (2019). Publication bias-Importance of studies with negative results! *Indian Journal of Anaesthesia*, *63*(6), 505-507. https://doi.org/10.4103/ija.IJA_142_19

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(210), 1-10. https://doi.org/10.1186/s13643-016-0384-4

Paivio, S. C., Holowaty, K. A., & Hall, I. E. (2004). The influence of therapist adherence and competence on client reprocessing of child abuse memories. *Psychotherapy: Theory, Research, Practice, Training*, *41*(1), 56-68. https://doi.org/10.1037/0033-3204.41.1.56

Pascual-Leone, A., & Greenberg, L. S. (2005). Classification of affective-meaning states. In A. Pascual-Leone (Ed.), *Emotional processing in the therapeutic hour: "Why the only way out is through"* (pp. 289–367). Unpublished doctoral thesis. York University.

Pascual-Leone, A., Paivio, S., & Harrington, S. (2016). Emotion in psychotherapy: An experiential–humanistic perspective. In D. J. Cain, K. Keenan, & S. Rubin (Eds.), *Humanistic*

psychotherapies: Handbook of research and practice (pp. 147–181). American Psychological Association. https://doi.org/10.1037/14775-006

Pascual-Leone, A., & Yeryomenko, N. (2016). The client "experiencing" scale as a predictor of treatment outcomes: A meta-analysis on psychotherapy process. *Psychotherapy Research*, *27*(6), 653-665. https://doi.org/10.1080/10503307.2016.1152409

*Pereira, R., Pinheiro, P., & Salgado, J. (2018). *Emotional processing on good and poor outcome cases of eft for depression* [Poster Presentation]. https://doi.org/10.13140/RG.2.2.24381.77284

Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *Journal of Applied Psychology, 90*(1), 175–181. https://doi.org/10.1037/0021-9010.90.1.175

*Pinheiro, P., Gonçalves, M. M., Nogueira, D., Pereira, R., Basto, I., Alves, D., & Salgado, J. (2022). Emotional processing during the therapy for complicated grief. *Psychotherapy Research*, *32*(5), 678-693. https://doi.org/10.1080/10503307.2021.1985183

*Pinheiro, P., Gonçalves, M. M., Sousa, I., & Salgado, J. (2021). What is the effect of emotional processing on depression? A longitudinal study. *Psychotherapy Research*, *31*(4), 507-519. https://doi.org/10.1080/10503307.2020.1781951

*Pinheiro, P., Mendes, I., Silva, S., Gonçalves, M. M., & Salgado, J. (2018). Emotional processing and therapeutic change in depression: A case study. *Psychotherapy*, *55*(3), 263-274. https://doi.org/10.1037/pst0000190

*Pole, N. (1999). *Client appraisals of danger and safety in psychotherapy and its physiological, facial, and subjective correlates* [Unpublished Doctoral Dissertation, University of California]. Bell & Howell Information and Learning. https://www.proquest.com/docview/304496121?pq-origsite=gscholar&fromopenview=true&sourcetype=Dissertations%20&%20Theses

Pos, A. E. (2006). *Experiential treatment for depression: A test of the experiential theory of change, differential effectiveness, and predictors of maintenance of gains* [Unpublished Doctoral

Dissertation, York University]. ProQuest Dissertations & Theses Global. https://proquest.com/dissertations-theses/experiential-treatment-depression-test-theory/docview/304984208/se-2

*Pos, A. E., Greenberg, L. S., Goldman, R. N., & Korman, L. M. (2003). Emotional processing during experiential treatment of depression. *Journal of Consulting and Clinical Psychology*, *71*(6), 1007-1016. https://doi.org/10.1037/0022-006X.71.6.1007

*Pos, A. E., Greenberg, L. S., & Warwar, S. H. (2009). Testing a model of change in the experiential treatment of depression. *Journal of Consulting and Clinical Psychology*, *77*(6), 1055-1066. https://doi.org/10.1037/a0017059

*Pos, A. E., Paolone, D. A., Smith, C. E., & Warwar, S. H. (2017). How does client expressed emotional arousal relate to outcome in experiential therapy for depression? *Person-Centered & Experiential Psychotherapies*, *16*(2), 173-190. https://doi.org/10.1080/14779757.2017.1323666

*Ralston, M. (2006). *Emotional arousal and depth of experiencing in imaginal confrontation versus evocative empathy* [Unpublished Doctoral Dissertation, University of Windsor]. Electronic Theses and Dissertations. https://scholar.uwindsor.ca/etd/2382

*Robichaud, L. K. (2004). *Depth of experiencing as a client prognostic variable in emotion-focused therapy for adult survivors of childhood abuse* [Unpublished Doctoral Dissertation, University of Windsor]. Electronic Theses and Dissertations. https://scholar.uwindsor.ca/etd/1642

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

*Rudkin, A., Llewelyn, S., Hardy, G., Stiles, W. B., & Barkham, M. (2007). Therapist and client processes affecting assimilation and outcome in brief psychotherapy. *Psychotherapy Research*, *17*(5), 613-621. https://doi.org/10.1080/10503300701216298

Schoeninger, D. W., Klein, M. H., & Mathieu, P. L. (1967). Sampling from recorded therapy interview: patient experiencing ratings made with and without therapist speech cues. *Psychological Reports*, *20*(1), 250-250. https://doi.org/10.2466/pr0.1967.20.1.250

Simons, D. J. (2014). The value of direct replication. *Perspectives on Psychological Science*, *9*(1), 76-80. https://doi.org/10.1177/1745691613514755

*Singh, T., Pascual-Leone, A., Morrison, O. P., & Greenberg, L. (2021). Working with emotion predicts sudden gains during experiential therapy for depression. *Psychotherapy Research*, *31*(7), 895-908. https://doi.org/10.1080/10503307.2020.1866784

Sønderland, N. M., Solbakken, O. A., Eilertsen, D. E., Nordmo, M., & Monsen, J. T. (2023). Emotional changes and outcomes in psychotherapy: A systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology*. Advance online publication. https://doi.org/10.1037/ccp0000814

Stiles, W. B. (1996). When more of a good thing is better: Reply to Hayes et al. (1996). *Journal of Consulting and Clinical Psychology, 64*(5), 915–918. https://doi.org/10.1037/0022-006X.64.5.915

Suurmond, R., van Rhee, H., & Hak, T. (2017). Introduction, comparison and validation of Meta-Essentials: A free and simple tool for meta-analysis. *Research Synthesis Methods, 8*(4), 537–553. https://doi.org/10.1002/jrsm.1260

Thomas, B. H., Ciliska, D., Dobbins, M., & Micucci, S. (2004). A process for systematically reviewing the literature: Providing the research evidence for public health nursing interventions. *Worldviews on Evidence-Based Nursing, 1*(3), 176-184. https://doi.10.1111/j.1524-475X.2004.04006.x

Thorlund, K., Imberger, G., Johnston, B. C., Walsh, M., Awad, T., Thabane, L., Gludd, C., Devereaux, P.J., & Wetterslev, J. (2012). Evolution of heterogeneity (I^2) estimates and their 95% confidence intervals in large meta-analyses. *PloS One*, *7*(7), e39471. https://doi.10.1371/journal.pone.0039471

*Toukmanian, S. G., Jadaa, D. A., & Armstrong, M. S. (2010). Change processes in clients' self-perceptions in experiential psychotherapy. *Person-Centered & Experiential Psychotherapies*, *9*(1), 37-51. https://doi.org/10.1080/14779757.2010.9688503

Wang, R., & Ware, J. H. (2013). Detecting moderator effects using subgroup analyses. *Prevention Science*, *14*, 111-120. https://doi.org/10.1007/s11121-011-0221-x

*Watson, J. C., & Bedard, D. L. (2006). Clients' emotional processing in psychotherapy: A comparison between cognitive-behavioral and process-experiential therapies. *Journal of Consulting and Clinical Psychology*, *74(1),* 152-159. https://doi.org/10.1037/0022-006X.74.1.152

Watson, J. C., Gordon, L. B., Stermac, L., Kalogerakos, F., & Steckley, P. (2003). Comparing the effectiveness of process-experiential with cognitive-behavioral psychotherapy in the treatment of depression. *Journal of Consulting and Clinical Psychology*, *71*(4), 773-781. https://doi.org/10.1037/0022-006X.71.4.773

*Watson, J. C., & Greenberg, L. S. (1996). Pathways to change in the psychotherapy of depression: Relating process to session change and outcome. *Psychotherapy: Theory, Research, Practice, Training*, *33*(2), 262-274. https://doi.org/10.1037/0033-3204.33.2.262

*Watson, J. C., McMullen, E. J., Prosser, M. C., & Bedard, D. L. (2011). An examination of the relationships among clients' affect regulation, in-session emotional processing, the working alliance, and outcome. *Psychotherapy Research*, *21*(1), 86-96. https://doi.org/10.1080/10503307.2010.518637

Wiggins, B. J., & Christopherson, C. D. (2019). The replication crisis in psychology: An overview for theoretical and philosophical psychology. *Journal of Theoretical and Philosophical Psychology*, *39*(4), 202-217. https://doi.org/10.1037/teo0000137

Wilson, D. B. (2023). *Practical meta-analysis effect size calculator* (Version 2023.11.27) [Computer Software]. https://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php

*Wong, K. K. Y. (2016). *Examining processes of change for experientially distant and experientially engaged client subgroups in experiential therapy for depression* [Unpublished Doctoral Dissertation, York University]. Electronic Theses and Dissertations. http://hdl.handle.net/10315/32344

*Wong, E. (2023). *The Role of Treatment-Seeking Stigma on Treatment Engagement and Outcome* [Unpublished Doctoral Dissertation, University of Toronto]. ProQuest Dissertations & Theses Global. https://www.proquest.com/docview/2889809628?pq-origsite=gscholar&fromopenview=true&sourcetype=Dissertations%20&%20Theses

Wood, J. A. (2008). Methodology for dealing with duplicate study effects in a meta-analysis. *Organizational Research Methods*, *11*(1), 79-95. https://doi.org/10.1177/1094428106296638

Yeryomenko, N. (2012). *Does the depth of client experiencing predict good psychotherapy outcomes? A meta-analysis of treatment outcomes* [Unpublished Doctoral Dissertation, University of Windsor]. Electronic Theses and Dissertations. https://scholar.uwindsor.ca/etd/4847

Appendices

Appendix A – The Experiencing Scale

the experiencing scale

The Experiencing Scale (EXP) describes 7 levels of emotional and cognitive involvement with one's ongoing experience. Feeling safe enough to & knowing how to work sometimes at deeper levels typically leads to better progress. This improved effectiveness is probably due to several factors that include a.) clearer understanding at head, heart & gut levels of why one thinks, feels & reacts as one does. b.) more contact with feelings allows them to be worked on, processed & integrated to produce better balance in thoughts, emotions and behaviours. c.) stronger connection with healthy, appropriate emotions can energise us to act to change outer circumstances in helpful ways.

1.) the client simply talks about events, ideas or others

2.) refers to self but without expressing emotions.

3.) expresses emotions but only as they relate to external circumstances.

4.) the client focuses directly on emotions and thoughts about self

5.) engages in an exploration of his or her inner experience

6.) gains awareness of previously implicit feelings & meanings

7.) on-going process of in-depth self-understanding, which provides new perspectives to solve significant problems

---

The scale has been shown to predict client change, especially in client-centred therapy, but it is relevant for cognitive therapy and for group work too.

Klein M H, Mathieu P L, Gendlin E T & Kiesler D J
The experiencing scale: a research and training manual (vol 1)
Madison: University of Wisconsin, 1969

Appendix B – PRISMA Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	6
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	7
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	9
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	8
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	8
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	10
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	10
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	11
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	11
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	12
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	12-14
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	12-14
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	13
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	12, 15
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	12-15
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	15-16
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	15
Reporting bias	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	16

Section and Topic	Item #	Checklist item	Location where item is reported
assessment			
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Not applicable
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	18
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	27 and 62
Study characteristics	17	Cite each included study and present its characteristics.	21 and 29
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	71
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	29, 32-37
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	32-37, 71
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	32-37
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	33
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	33, 36
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	34, 36
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Not applicable
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	38
	23b	Discuss any limitations of the evidence included in the review.	39-41
	23c	Discuss any limitations of the review processes used.	37-38
	23d	Discuss implications of the results for practice, policy, and future research.	41-42
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	8
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	8
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Not applicable
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Not applicable
Competing interests	26	Declare any competing interests of review authors.	Not applicable
Availability of	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included	Not

Section and Topic	Item #	Checklist item	Location where item is reported
data, code and other materials		studies; data used for all analyses; analytic code; any other materials used in the review.	applicable

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: http://www.prisma-statement.org/

Appendix C – Effective Public Health Practice Project Quality Appraisal Tool

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?

1 Very likely
2 Somewhat likely
3 Not likely
4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?

1 80 - 100% agreement
2 60 – 79% agreement
3 less than 60% agreement
4 Not applicable
5 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design

1 Randomized controlled trial
2 Controlled clinical trial
3 Cohort analytic (two group pre + post)
4 Case-control
5 Cohort (one group pre + post (before and after))
6 Interrupted time series
7 Other specify _____
8 Can't tell

Was the study described as randomized? If NO, go to Component C.

No Yes

If Yes, was the method of randomization described? (See dictionary)

No Yes

If Yes, was the method appropriate? (See dictionary)

No Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?

1 Yes
2 No
3 Can't tell

The following are examples of confounders:

1 Race
2 Sex
3 Marital status/family
4 Age
5 SES (income or class)
6 Education
7 Health status
8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?

1 80 – 100% (most)
2 60 – 79% (some)
3 Less than 60% (few or none)
4 Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?

1 Yes
2 No
3 Can't tell

(Q2) Were the study participants aware of the research question?

1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?

1 Yes
2 No
3 Can't tell

(Q2) Were data collection tools shown to be reliable?

1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?

1 Yes
2 No
3 Can't tell
4 Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell
5 Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell

(Q2) Was the consistency of the intervention measured?

1 Yes
2 No
3 Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?

4 Yes
5 No
6 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)

community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)

community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?

1 Yes
2 No
3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?

1 Yes
2 No
3 Can't tell

GLOBAL RATING

COMPONENT RATINGS

Please transcribe the information from the gray boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

A	**SELECTION BIAS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
B	**STUDY DESIGN**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
C	**CONFOUNDERS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
D	**BLINDING**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
E	**DATA COLLECTION METHOD**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
F	**WITHDRAWALS AND DROPOUTS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one):

1	STRONG	(no WEAK ratings)
2	MODERATE	(one WEAK rating)
3	WEAK	(two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?

No Yes

If yes, indicate the reason for the discrepancy

1	Oversight
2	Differences in interpretation of criteria
3	Differences in interpretation of study

Final decision of both reviewers (circle one):

1	**STRONG**
2	**MODERATE**
3	**WEAK**

Appendix D – Justification of Removal of Papers at Full Text

First Author (Date)	Title	Reason for Exclusion
Adams et al. (2011)	Therapist influence on depressed clients' therapeutic experiencing and outcome.	EXP not analysed with treatment outcomes.
Arimura et al. (1990)	An analysis of process movement measured on the experiencing scales in a family therapy session.	No treatment outcomes.
Berthoud et al. (2015)	Emotional processing in a ten-session general psychiatric treatment for borderline personality disorder: A case study.	Not using EXP Scale.
Brintzinger et al. (2021)	Patients' Style of Emotional Processing Moderates the Impact of Common Factors in Psychotherapy.	Not using EXP Scale.
Brown (1984)	Counsellor and client in-therapy verbal behaviour.	No treatment outcomes.
Crisostomo (2013)	Emotion Processes in Cognitive Behavioral Therapy for Adolescent Depression.	Not using EXP Scale.
Cummings et al. (2014)	Expressive writing in psychotherapy: A tool to promote and track therapeutic change.	Not using EXP Scale.
Custers (1973)	Experiencing in the therapeutic process: Study of the relation between experiencing change and personality change in client-centered psychotherapy.	Full text could not be accessed.

First Author (Date)	Title	Reason for Exclusion
Douglass (1989)	The effect of therapist interpretations on patient level of anxiety.	No treatment outcomes.
Drapeau et al. (2018)	A process study of long-term treatment: Comparing a successful and a less successful outcome.	No treatment outcomes.
Feldman et al. (2009)	Change in emotional processing during a dialectical behavior therapy-based skills group for major depressive disorder.	Not using EXP Scale.
Fitzpatrick et al. (1999)	Client emotional involvement and occurrence of in-session therapeutic phenomena.	No treatment outcomes.
Friedlander et al. (2018)	"If those tears could talk, what would they say?" multi-method analysis of a corrective experience in brief dynamic therapy.	Not using EXP Scale.
Garson (1983)	Field independence and productive client behaviour in psychotherapy.	No treatment outcomes.
Gazzola (1997)	An investigation of counselor interpretations in client-centered therapy.	No treatment outcomes.
Gill (2002)	The contribution of client-identified therapeutically valuable events to the working alliance.	No treatment outcomes.
Gleiser (2003)	Emotional arousal during therapy for posttraumatic stress disorder with childhood sexual abuse survivors.	Not using EXP Scale.

First Author (Date)	Title	Reason for Exclusion
Gordon et al. (2002)	Is how it is said important? The association between quality of therapist interventions and client processing.	No treatment outcomes.
Greenberg et al. (1979)	Differential effects of the two-chair experiment and empathic reflections at a conflict marker.	Non-clinical population.
Greenberg et al. (1981)	Specific effects of Gestalt two-chair dialogue on intrapsychic conflict in counseling.	No treatment outcomes.
Greenberg et al. (1996)	Task analysis exemplified: The process of resolving unfinished business.	No treatment outcomes.
Greenberg et al. (1981)	The specific effects of a Gestalt intervention.	No treatment outcomes.
Hager (1987)	Experiencing scale discrimination between more and less productive psychotherapy sessions.	No treatment outcomes.
Harte et al. (2020)	Processing emotional pain using the expanded Emotion Focused Therapy task of Focusing: A single-session case study.	No treatment outcomes.
Holowaty (2005)	Process characteristics of client-identified helpful events in emotion-focused therapy for adult survivors of childhood abuse (EFT-AS).	EXP not analysed with treatment outcomes.

First Author (Date)	Title	Reason for Exclusion
Holtz (2004)	The self- and interactive regulation and coordination of vocal rhythms, interpretive accuracy, and progress in brief psychodynamic psychotherapy.	EXP not analysed with treatment outcomes.
Horowitz et al. (1993)	Elaboration and dyselaboration: Measures of expression and defense in discourse.	EXP not analysed with treatment outcomes.
Hu et al. (2010)	Relationship between client exploration, counselor direction and session outcome.	No treatment outcomes.
Johnson et al. (1988)	Relating process to outcome in marital therapy.	EXP not analysed with treatment outcomes.
Kailanko et al. (2022)	Impact of repeating somatic cues on the depth of experiencing for withdrawers and pursuers in emotionally focused couple therapy.	No treatment outcomes.
Katz et al. (2017)	The creation of the Client Reflexivity Scale: A measure of minute fluctuations in self-awareness and exploration.	EXP not analysed with treatment outcomes.
Kiesler (1971)	Patient experiencing and successful outcome in individual psychotherapy of schizophrenics and psychoneurotics.	No treatment outcomes.
Kiesler (1969)	Refinement of the Experiencing Scale as a Counseling Tool. Final Report.	No treatment outcomes.

First Author (Date)	Title	Reason for Exclusion
Kiesler et al. (1970)	Comparison of Experiencing scale ratings of naive versus clinically sophisticated judges.	No treatment outcomes.
Klein et al. (2006)	Client accounts of personal change in process-experiential psychotherapy: A methodologically pluralistic approach.	Not using EXP Scale.
Kray (2010)	Evaluation of the therapeutic alliance and patient-therapist emotional exploration in time-limited therapy.	Qualitative data.
Levitt (1994)	A comparative analysis of the narrative process coding system and three standardized psychotherapy process measures: A multimodal analysis.	EXP not analysed with treatment outcomes.
Lewin (2011)	The importance of emotional-reflexive patterns for productive therapy: A narrative process analysis of emotion-focused and client-centered psychotherapy.	Full text could not be accessed.
Lewis et al. (1983)	Experiencing level in the process of group development.	Not using EXP Scale.
Macauley (2011)	A comparison of narrative process sequences in cognitive behavioural and emotion focused therapies for depression.	EXP not analysed with treatment outcomes.

First Author (Date)	Title	Reason for Exclusion
Macauley et al. (2007)	Attunement as the core of therapist-expressed empathy.	No treatment outcomes.
Manne et al. (2007)	Social-Cognitive Processes as Moderators of a Couple-Focused Group Intervention for Women With Early Stage Breast Cancer.	Not using EXP Scale.
Manne et al. (2017)	Emotional processing during psychotherapy among women newly diagnosed with a gynecological cancer.	EXP not analysed with treatment outcomes.
Missirilan (2011)	A comparative study of the nature of change processes in emotion focused and cognitive-behavioural psychotherapies for depression.	Full text could not be accessed.
Muran et al. (2001)	A cognitive-interpersonal case study of a self.	No treatment outcomes.
Naaman et al. (2005)	Treating Attachment Injured Couples with Emotionally Focused Therapy: A Case Study Approach.	No treatment outcomes.
Newton et al. (2004)	Implosive therapy in alcoholism: Comparison with brief psychotherapy.	No treatment outcomes.
Nichols (1977)	The delayed impact of group therapists' interventions.	Full text could not be accessed.
O'Driscoll et al. (2016)	Process analysis of trauma-focused cognitive behavioural therapy for individuals with schizophrenia.	EXP not analysed with treatment outcomes.

First Author (Date)	Title	Reason for Exclusion
Pascual-Leone (2009)	Dynamic Emotional Processing in Experiential Therapy: Two Steps Forward, One Step Back.	EXP not analysed with treatment outcomes.
Pascual-Leone (2007)	Emotional processing in the therapeutic hour: Why "the only way out is through".	Not using EXP Scale.
Pinheiro (2022)	Emotional processing during the therapy for complicated grief.	Qualitative data.
Pos (2006)	Experiential treatment for depression: A test of the experiential theory of change, differential effectiveness, and predictors of maintenance of gains.	Full text could not be accessed.
Rhodes (1991)	The influence of the therapist's process on the patient's experience of sadness.	No treatment outcomes.
Rogan (2001)	Experiencing and Emotional Expression in Psychotherapy: An Investigation of Two In-Session Client Processes.	No treatment outcomes.
Romano (2007)	Attachment in psychotherapy—the secure base hypothesis and the role of the therapist.	No treatment outcomes.
Rubinstein (1971)	The effects of different kinds of therapist responses made within the context of high and low levels of facilitation on client experiencing.	Full text could not be accessed.

First Author (Date)	Title	Reason for Exclusion
Schaeffer (1977)	Client attraction and distress: Unexpected impact on psychotherapeutic process.	EXP not analysed with treatment outcomes.
Sherman (1988)	Client language and clinical process: A cognitive-semantic analysis.	No treatment outcomes.
Singh (2021)	More to gain: Sudden gains in experiential therapy for depression	EXP not analysed with treatment outcomes.
Stalikas et al. (1995)	Client good moments: An intensive analysis of a single session.	No treatment outcomes.
Stiegler et al. (2018)	Does an emotion-focused two-chair dialogue add to the therapeutic effect of the empathic attunement to affect?	Not using EXP Scale.
Virtue et al. (2019)	Levels of emotional awareness during psychotherapy among gynecologic cancer patients.	Non-clinical population.
Voigt et al. (1992)	Intervention style and client progress in time-limited group psychotherapy for adults sexually abused as children.	No treatment outcomes.
Walden (2013)	The real relationship, therapist immediacy, and client experiencing level: A dyad study of psychotherapy process and connection.	No treatment outcomes.
Warwar (1996)	The relationship between level of experiencing and session outcome in	No treatment outcomes.

First Author (Date)	Title	Reason for Exclusion
	client-centered and process-experiential therapies.	
Warwar (2005)	Relating emotional processes to outcome in experiential psychotherapy of depression.	Full text could not be accessed.
Watson (1992)	The process of change when exploring problematic reactions: An inquiry into self.	EXP not analysed with treatment outcomes.
Wiseman (1986)	Single-case studies of the resolution of problematic reactions in short-term client-centered therapy: A task-focused approach.	EXP not analysed with treatment outcomes.
Wiser et al. (1993)	Comparative Study of Emotional Experiencing in Psychodynamic-Interpersonal and Cognitive-Behavioral Therapies.	No treatment outcomes.
Zuccarini et al. (2013)	Forgiveness and Reconciliation in Emotionally Focused Therapy for Couples: The Client Change Process and Therapist Interventions.	No treatment outcomes.
Zuccarini (2012)	The Attachment Injury Resolution Model in emotionally focused couple therapy: A psychotherapy process study of in-session client performances and therapist behaviors.	No treatment outcomes.

Appendix E – Effective Public Health Practice Project Quality Appraisal Tool Scoring.

Paper	Selection Bias	Study Design	Confounders	Blinding	Data Collection Method	Withdrawals and Dropouts	Overall Rating
Anderson et al., (2022)	3	1	1	2	1	2	2
Fisher et al., (2020)	2	2	3	1	1	1	2
Goldman et al., (2005)	3	1	1	2	1	1	2
Greenberg et al., (1983)	3	2	3	3	1	1	3
Grooh et al., (1993)	2	2	1	2	1	1	1
Hakim et al., (2010)	3	1	3	2	1	1	3
Harrington et al., (2021)	3	2	1	2	1	1	2
Isgar (2024)	2	1	1	1	1	1	1
Jackson et al., (2013)	3	1	3	2	1	1	3
Klug et al., (2021)	2	1	3	2	1	2	2
Levitt et al., (2000)	3	2	3	3	1	1	3
Malin et al., (2015)	3	2	1	2	1	1	2
Pereira et al., (2018)	3	2	3	3	1	1	3
Pinheiro et al., (2018)	3	2	N/A	2	1	1	2
Pinheiro et al., (2021)	3	1	3	2	1	1	3

Paper	Selection Bias	Study Design	Confounders	Blinding	Data Collection Method	Withdrawals and Dropouts	Overall Rating
Pinheiro et al., (2022)	2	2	3	2	1	2	2
Pole et al., (1999)	3	2	3	3	1	1	3
Pos et al., (2003)	3	1	1	2	1	1	2
Pos et al., (2009)	3	1	1	2	1	1	2
Pos et al., (2017)	3	1	3	2	1	1	3
Ralston et al., (2006)	3	1	1	2	1	1	2
Robichaud et al., (2004)	3	2	1	2	1	1	2
Rudkin et al., (2007)	2	1	3	1	1	1	2
Singh et al., (2021)	3	1	1	2	1	2	2
Toukmanian et al., (2010)	3	1	3	1	1	1	3
Watson et al., (1996)	3	1	1	1	1	1	2
Watson et al., (2006)	3	1	1	2	1	2	2
Watson et al., (2011)	3	1	1	2	1	2	2
Wong (2016)	3	1	1	2	1	2	2
Wong (2023)	3	1	1	2	1	2	2

Note. A score of '1' with a green background indicates the item has been scored as 'strong. A score of '2' with an amber background indicates the item has been scored as 'moderate'. A score of '3' in a red background indicates the item has been scored as 'weak'.

Section Two: Empirical Study

Clients' Emotional Processing in Psychotherapy: A Comparison Between Cognitive-Behavioural and Person-Centred Experiential Therapies Utilising the Archived PRaCTICED Trial Data Set

Abstract

Objectives: To conduct a partial replication and refinement of Watson and Bedard (2006), testing the hypothesis that levels of client depth of experiencing will be higher in person-centred experiential therapy (PCET) than in cognitive behavioural therapy (CBT).

Design: A mixed-method secondary data analysis design was employed, using data from a pragmatic non-inferiority randomised trial of the clinical and cost effectiveness of PCET vs. CBT.

Methods: Forty participants were selected, 20 receiving PCET and 20 receiving CBT, who recorded the greatest and least change at the end of therapy according to the Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001). The middle 20 minutes of early and middle therapy (sessions 2 and 6) were coded using the Experiencing Scale (EXP Scale; Klein et al., 1969), generating data for the modal and peak EXP scores for each participant. EXP scores were subjected to a 2 X 2 split-plot analysis of variance (ANOVA). Post-hoc analyses were also conducted. The study was pre-registered.

Results: For both modal and peak EXP, PCET scores were significantly higher than those for CBT. Stage of therapy did not yield a significant effect.

Conclusions: The results provide mixed support for the current hypotheses and previous literature.

Key Words: Experiencing, Emotional Processing, Depression, Cognitive Behavioural Therapy, Person Centred Experiential Therapy.

Practitioner Points:

- Therapists should be mindful of clients' capacity to process emotions and where possible allocate them to the most appropriate therapy.

- Training therapists in different therapies should be specific. For example, while both therapies should focus on developing the emotional processing skills of clients, PCET training should place greater emphasis on this in comparison to CBT.
- Themes of conversation within supervision of therapists across therapies could differ. In the earlier stages of therapy, supervision should focus on emotional processing across both therapies, whereas towards mid-therapy, there could be greater focus on emotional processing within PCET supervision and less so in CBT supervision.

Introduction

Psychological therapy process research is the exploration of in-session changes, achieved through the observation and analysis of moment-to-moment interactions within sessions (Watson, 2023). Of the people who access mental health care, it is estimated that over sixty percent do not meaningfully benefit from the support offered (Kraus et al., 2011). Therefore, one of the aims of psychological therapy process research is to identify the differing roles of processes within therapy, and their relations to outcomes, so that clinical practice can be further refined. Watson (2023) described process research as 'the bedrock of psychotherapy'. She reflected the development of theories and observations have been, in part, due to process research. Increased understanding may improve therapist responsiveness (Stiles, 2021), by adapting interventions to meet the needs of individuals (Crits-Christoph & Connolly Gibbons, 2021). Process research is also beneficial for therapists in training as it provides a valuable source of information for theory development and the evolution of training (Westerman & de Roten, 2017).

The Experiencing Scale

Rogers (1959) described the goal of successful therapy as the person becoming open to their feelings, 'from remoteness to living meaningfully'. Rogers (1951) emphasised the process of reaching the end of the continuum (described as fully-functioning person) required higher levels of emotional processing.

The most extensively used measure of emotional processing is the Experiencing Scale (EXP; Klein et al., 1969). Psychotherapy process researchers have labelled it as the "gold standard" regarding measuring this process (Pascual-Leone et al., 2017), resulting in it being used to study a variety of therapeutic approaches (Hendricks, 2009).

The EXP was created to measure the quality and nature of a client's active participation in therapy (Klein et al., 1986). The scale measures from impersonal and superficial accounts (stage 1) to content which is expansive in nature, exposing a client's expanding awareness of feelings and processes (stage 7).

Emotional Processing Research

Castonguay et al. (1996) explored emotional processing in cognitive therapy. They observed the level of emotional processing was higher for the cognitive-therapy group than for the group that received cognitive therapy alongside pharmacotherapy for depression. One of their primary conclusions was more research needs to be conducted to gain a better understanding of the role of emotional processing in the change process of cognitive therapy for depression. Emotional processing has since been analysed in other therapies, including emotion-focussed therapy (EFT; Burgess-Moser, 2012) and interpersonal therapy (IPT; Hakim, 2010).

There have been a few studies that have compared emotional processing between different therapies. Watson and Bedard (2006) compared depressed clients' emotional processing in good and bad outcome cases in CBT and process-experiential therapy (PET). They also explored how clients' emotional processing changed over the course of therapy. They concluded that clients in the PET group showed significantly higher levels of emotional processing than those in the CBT group. They explained this by suggesting CBT clients are more separated from their emotional experience than clients in PET and therefore their EXP scores are significantly lower. They also concluded that there was a significant increase in level of emotional processing from the early stages of therapy to the middle stages. The latter outcome corroborates with results from other studies and therapies, including experiential therapy (Goldman et al., 2005; Pos et al., 2009), CBT, IPT and process experiential therapy (Hakim, 2010), which all found depth of emotional processing increasing from early to middle therapy.

Rudkin et al. (2007) compared EXP scores between clients with depression who were accessing CBT and psychodynamic-interpersonal therapy (PIT). Clients who accessed PIT had significantly higher EXP scores compared to clients accessing CBT. Thus, the findings from this study and Watson and Bedard's (2006) suggest that depth of emotional processing differs between therapies.

More recently, additional papers have been published that have compared emotional processing between therapies. Klug et al. (2021) found emotional processing was higher in participants accessing psychoanalytic therapy compared to CBT, whilst Pinheiro et al. (2021) found emotional processing was higher for people receiving EFT compared to CBT.

Emotional Processing and its Relation to Outcomes

Research studying clients' emotional processing in psychotherapy using the EXP has found that experiencing levels vary from session to session. However, emotional processing has consistently been one therapeutic variable which has been associated with better therapeutic outcomes in a range of therapies (Ulvenes et al., 2014) but in particular experiential therapy (Pos et al. 2003).

Within Pascual-Leone and Yeryomenko's (2016) meta-analysis they studied therapy type as a moderator variable. The analysis revealed that EXP was related to outcomes in all therapies analysed. In addition, there was no significant difference in the effect sizes for the therapies analysed. The authors cautiously concluded that the relationship between EXP and outcomes holds across treatment approaches. They also explored depth of experiencing throughout therapy and found depth of experiencing increased from early to middle therapy sessions before reducing as therapy terminates.

Current Study

The aim of the current study was to conduct a partial replication of Watson and Bedard's study by comparing emotional processing in clients who received either CBT or person-centred experiential therapy (PCET). PCET is an approach which developed from the competencies for humanistic psychological therapies (Roth et al., 2009). It combines principles and processes from person-centred therapy, emotion theory and EFT (Elliott et al., 2004). PCET requires the clinician to actively work with clients' emotions, more so than standard person-centred therapy (Murphy, 2019), with the aim of helping clients to access and interpret underlying feelings (Hill, 2011).

The current study is connected to a second study being conducted by a fellow trainee who is exploring emotional processing in good and poor therapy outcomes using the same dataset. Given a goal of psychotherapy is to explore the processes of change and their relations to outcomes (Watson, 2018), it was deemed appropriate to collaborate but retain distinctly independent research questions and analytic strategies.

Clinical Rationale and Implications

The general volume of literature in the field of emotional processing is sparse (Pascual-Leone & Yeryomenko, 2016). Recommendations from previous researchers have consistently included the need for further research in the field, both on therapies already explored (e.g., CBT; Castonguay et al., 1996) and therapies that haven't been explored (Pascual-Leone & Yeryomenko, 2016).

Watson (2023) argues that given emotional processing has been identified as a variable associated with positive outcomes across different therapies, understanding its role across different therapies is important. There have been few studies that have explored the differences in level of experiencing across therapy type, and there have been no studies that have explored this process in PCET. In addition, identifying mechanisms influencing outcomes in National Health Service (NHS) Talking Therapies (previously known as Improving Access to Psychological Therapies (IAPT; Clark, 2018) has been minimal, in part because of methodological difficulties (Hubble et al., 1999).

In a recent pragmatic, non-inferiority trial, PCET and CBT have been shown to have similar outcomes for depression at the end of therapy within Talking Therapies services, although results favoured CBT at 12-month post-randomisation with PCET gains not being as well maintained (Barkham et al., 2021). The results of the current study may identify differential processes within the therapies that contributed to these outcomes. Emotional processing appears a sensible avenue to explore given previous researchers have titled emotional processing as an 'important variable' in its contribution to good therapeutic outcomes across a number of therapies (Town et al., 2017).

Lutz et al. (2021) reflected that process researchers need to utilise 'fine-grained' analyses of psychotherapy to identify, describe, and measure in-session processes to improve practice. Understanding the role of emotional processing in both therapies would help to guide clinicians within and between sessions, meaning they can work in the most optimal ways, which in turn should improve treatment outcomes (Watson & Wiseman, 2024). Understanding the role of emotional processing would not only be helpful in improving clinical practice but also in relation to guiding further research and in developing training.

Research Aims

The aims of the present study were to:

1. Compare depressed clients' levels of emotional processing between clients receiving CBT and clients receiving PCET.
2. Determine whether depressed clients' emotional processing increases over the course of therapy in each therapy.

Hypotheses

1. It is hypothesised that clients in the PCET group would have higher levels on the EXP Scale for both modal and peak scores than those in CBT.
2. It is hypothesised that clients' levels of EXP would be higher at the midpoint of therapy compared to the beginning for both therapies.

Methods

Design

The PRaCTICED trial (Barkham et al., 2021) was conducted within the Sheffield Talking Therapies service and its primary aim was to test the non-inferiority of PCET compared with CBT for clients experiencing moderately severe or severe depression. Given the trial found that PCET was non inferior to CBT at 6-months post-randomisation, the current study focused on exploring and understanding a specific therapeutic process, namely emotional processing, and its influence

within both therapies across early and middle sessions. A sequential mixed-methods, between-subjects design of secondary data was utilised.

A selected subsample of data from the PRaCTICED trial was used for the current study. The PHQ-9 pre and post change scores were used to identify participants to be sampled.

Ethical Considerations and Open Science

The Health Research Authority (Research Ethics Committee 14/YH/0001) granted the original study NHS ethical approval that included studies aimed at furthering understanding of patients' outcomes and differential treatment effects.

Ethical approval from the University of Sheffield was sought and approved (050463 - Appendix A) for the current study. For the purposes of open science, the current study was pre-registered with AsPredicted (140203 - Appendix B). To aid transparency and replicability, the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist was completed (Appendix C).

Participants

Participants who took part in the PRaCTICED trial were originally screened for eligibility and were required to consent to participate. The inclusion criteria for participants were that: they were aged 18 years or older, had a score of 12 or above on the PHQ-9, indicated depression was their primary concern, had a diagnosis of moderate or severe depression on the Clinical Interview Schedule-Revised (Lewis et al., 1992), and did not have a strong treatment preference for either of the two therapies being offered.

Participants were excluded if they had a long-term physical health condition, were dependent on drugs or alcohol, had an elevated risk of suicide, had a previous diagnosis of personality disorder, schizophrenia and / or bipolar disorder. Individuals who did not meet the criteria for the trial received treatment as usual within the local NHS Talking Therapies service.

A total of 510 participants (293 female) were recruited and met the criteria for the trial, with 254 randomly assigned to receive PCET and 256 randomly assigned to receive CBT. Of the total,

102 participants did not attend any therapy sessions and therefore a total of 408 participants started treatment. Up to 20 weekly sessions of one-to-one CBT or PCET therapy could be offered, which is in accordance with recommendations made by the National Institute for Health and Clinical Excellence (NICE, 2022) regarding the treatment of depression.

Therapists and Counsellors

In total there were 18 PCET counsellors and 32 CBT therapists, with similar levels of training according to the therapy modality they were trained in. Only those who successfully completed the pre-requisite training and were eligible to be Talking Therapies practitioners were included in the trial.

To increase consistency in the delivery of the therapies, treatment manuals were developed for both modalities. In addition, the 12-item Cognitive Therapy Scale-Revised (Blackburn et al., 2001) and the 10-item Person-Centred and Experiential Psychotherapy Scale (Freire et al., 2014) were administered at three time points (sessions 2, 6 and 12) for a selected subsample of recordings to assess adherence to CBT and PCET respectively. All therapists also received supervision in a combination of individual and group settings.

Procedure

Prior to the beginning of the PRaCTICED trial, participants provided demographic information and they also completed a range of measures which were collected at intervals during the study, the frequency of which depended on the measure used. The PHQ-9, Work and Social Adjustment Scale (WSAS; Mundt et al, 2002) and Generalised Anxiety Disorder-7 (GAD-7; Spitzer et al., 2006) were collected before each therapy session. Measures were also collected at 6-month and 12-month post-randomisation and comprised the; PHQ-9, GAD-7, Clinical Outcomes in Routine Evaluation- Outcome Measure (Evans et al., 2002), EQ-5D-5L (Herdman et al., 2011), WSAS, Beck Depression Inventory (BDI-II; Beck et al., 1996), Quality of Life Scale (Flanagan 1982), and the Client Satisfaction Questionnaire (Larsen et al., 1979).

Following randomisation, the participants were offered weekly one-to-one therapy sessions of either CBT or PCET.

Primary Measures in the Current Study

PHQ-9 (Appendix D)

The PHQ-9 is the primary measure that the current study used to investigate differences in symptoms. It is a standard measure for depression screening (Kroenke et al., 2010). The PHQ-9 has nine questions that cover a range of topics including eating, sleeping, concentration and sleep. Informers answer each question based on their experiences of the previous two weeks, which aligns with the criteria in the Diagnostic and Statistical Manual of Mental Disorders 5th Edition (2013). Response options are on a four-point scale ranging from "not at all" (0) to "nearly every day" (3). Higher scores are indicative of increased severity of depression. Scores of 5, 10, 15 and 20 represent the criteria for mild, moderate, moderately severe and severe depression. Internal reliability has been shown to have a Cronbach's α of 0.86, test re-test reliability has also been reported to be excellent (Kroenke et al., 2001).

EXP (Appendix E)

The EXP is a measure of the depth of a person's emotional processing. Each client's speaking turn (of more than three words) is coded on a 7-point scale (Table 1). All speech is rated, leading to a modal score (most frequently occurring score in a passage) and a peak score (the highest level reached in the excerpt). Interrater reliabilities range from .63 to .93 for modal ratings and from .61 to .93 for peak ratings (Klein et al., 1986).

Table 1

The Stages of The EXP Scale and a Descriptor of Each Stage.

Stage 1	The speaker will talk about external events that are impersonal in nature. The content is not directly about the speaker and their association to the event is unclear.

Stage 2	The speaker will discuss events that are more personal, describing themselves as a participant within the narrative. The description may be impersonal, but the speaker's role is clear. Feelings may be revealed implicitly, but not explicitly.
Stage 3	The speaker will describe a narrative of an event, which could be from the past, present or future. They will provide some reaction to the event, although this will be limited in nature, for example by attaching a behavioural description of the feeling.
Stage 4	The speaker will describe a personal and specific situation, completely from their point of view. They will describe their reactions to the events in much more depth than Stage 3, for example sharing several feelings. There will be increased depth to the description of a feeling or multiple feelings and personal experiences.
Stage 5	The speaker begins to explore potential problems, the problem must be about their feelings or reactions to personal experiences. The speaker must explore their problems in a personal way.
Stage 6	The speaker progresses to a state of resolving issues related to experiences and feelings. Feelings must be vividly presented.
Stage 7	The content is expansive in nature, exposing the client's expanding awareness of feelings and processes.

Data Security

The manager of the original PRaCTICED trial only provided access to data that was relevant to this study. All data was anonymised by a wider member of the PRaCTICED trial team. They stored the recordings on an encrypted file within a shared drive. Permissions to access this file were only granted to the researchers and research supervisors.

Analysis

Power Analysis

Given the aim of this study was to explore emotional processing from an existing dataset, by partially replicating the methodology used by Watson and Bedard (2006), including the same sample size, a purposive sampling strategy was used. Therefore, a priori power analysis did not fit with the research design. However, retrospective power analyses were conducted to provide context of whether the study had the power to detect the desired effects and to indicate what sample size future studies would require for robust conclusions from their results. Retrospective power analyses were conducted using G*Power 3.1.9.7 (Faul et al., 2007).

Data extraction

The baseline PHQ-9 data, the weekly PHQ-9 data and end of therapy PHQ-9 data for all participants was initially made available, to determine selection for the study. Once the sample had been selected, only demographic information for those people was made available.

Stage 1 – Selection of Participants

Initially researchers had to assess whether participants were eligible for the current study by applying inclusion/exclusion criteria. This included: clients had to have had a minimum of seven sessions of therapy (as session 2 and 6 were being coded); they needed to have outcome measures taken at the first and last session; they could have not switched treatment groups; and their sessions needed to have been recorded. Once these criteria were applied the sample comprised 215 clients, 110 in the PCET group and 105 in the CBT group.

Data from 40 of the above 215 participants in the main trial were used in the current study. The study sample comprised 40 participants as per the Watson and Bedard (2006) study: 10 CBT good outcome, 10 CBT poor outcome, 10 PCET good outcome and 10 PCET poor outcome. This allowed comparisons between therapies (the current study) and between good and poor outcomes (a study being conducted by a fellow researcher – Appendix F).

The participants were systemically selected using their PHQ-9 data. Specifically, the 10 participants in each therapy who showed the largest change in PHQ-9 scores and those who showed the least change, as measured in their first session of therapy and last session of therapy, were selected. The rationale for using the PHQ-9 rather than replicating Watson and Bedard's (2006) sampling method was because the current study was focussing on pre-post therapy change and only the PHQ-9 was used at the end of treatment (the BDI-II was defined by being 6-months post-randomisation and not end of therapy).

Stage 2 - Qualitative analysis

From the 40 participants selected at stage one, extracts of session recordings were interpreted with the EXP scale, a validated measure of depth of experiencing, using basic deductive content analysis (Berelson, 1952).

Basic content analysis is an eight-stage technique (Appendix G) which identifies words or themes and assigns them to pre-specified codes which are then analysed using quantitative analyses (Neuendorf, 2002). Other qualitative analyses were not considered, given that data was to be converted into numerical form. Basic content analysis has been widely used in healthcare research (Hsieh & Shannon, 2005) and has predominantly been used for analysis of secondary data, making it suitable for the current study (Drisko & Maschi, 2016). Epistemology has been unexplored within the content analysis literature, however most basic content analysis studies have positioned themselves in a positivist position (Drisko & Maschi, 2016). Giddings (2006) has suggested that mixed-methods approaches have tended to take a positivist viewpoint. This position requires a realist and objective stance (Park et al., 2020).

Two therapy sessions were used for the qualitative ratings. Two time points were chosen because Klein et al., (1969) outlined at least two time points are needed to be reviewed in therapy, especially if EXP is being considered in relation to outcome criteria.

The middle twenty minutes (twenty minutes following the initial twenty minutes) of the session were coded. The middle section was selected because this typically represents the "working

segment" of the session. Sessions were chosen from the early (session 2) and middle (session 6) stages of therapy. Early and middle sessions were chosen as Watson and Bedard (2006) found this was where most change in the EXP scale occurred. Sessions 2 and 6 were specifically chosen because those sessions had been assessed for adherence to CBT and PCET.

Each participant response occurring between one therapist response, and another was rated and categorised on the EXP scale. This generated a datum. The most frequently coded number was used as the modal score, whilst the highest coded number was used as the peak score. This was repeated for all 80 recordings, giving the researchers 80 modal scores and 80 peak scores.

The author of the current study and a collaborating researcher (KN) from a parallel study, collaborated in securing the ratings that were used in both studies, to address different questions. To reduce researcher bias, where possible, both researchers were blind to selected recordings, with no identifiable information provided alongside the recordings. Each researcher coded an equal mix of session 2 and session 6 recordings (therefore blinding the researchers to session number) and an equal mix of CBT and PCET recordings (blinding researchers to therapy type, although due to the researcher's backgrounds, it was possible to identify therapy type). The two researchers were also blind as to whether the participants had a good or poor outcome.

Stage 3 – Quantitative analysis

To assess whether there were any significant differences in level of emotional processing between PCET and CBT and between early and middle stages of therapy, the mean modal EXP scores were subjected to a 2 X 2 split-plot analysis of variance (ANOVA), with therapy type (PCET or CBT) as the between-subjects factors, and the stage of therapy (early or middle) as the repeated within-subjects factor. The ANOVA was also conducted a second time, using the mean peak EXP scores. The primary aim of the analyses was to understand the interaction effect between clients emotional processing between the two therapies at the two time points.

If significant interaction effects were found, it would suggest that the contrasting groups responded differently across the stages of the investigation and that they therefore should be

examined separately in simple effects analyses. Therefore, simple effect analyses were planned if interaction effects were significant.

Post-hoc comparisons were also conducted to increase understanding of the differences between the two therapies. The aim of these analyses was to further explore and explain the qualitative differences between PCET and CBT. To do this, Chi-square or Fisher's exact test were used to explore whether there were significantly different numbers of clients in each therapy exploring their emotions at each level of the EXP scale. As per the recommendations, Chi-square was used when expected counts were 5 and over, while Fisher's exact test was used when expected counts were less than 5 (Kim, 2016). Given the multiple tests being conducted, a method of correction was planned. However, the literature suggests that the smallest sample sizes required for corrections were larger than the sample size for the current study. The literature suggested 65 participants were required to complete 6 comparisons using the Bonferroni Correction (Campelo & Wanner, 2020). Therefore, the results are interpreted tentatively, with an awareness of an increased risk of Type 1 error.

Quality control

The Experiencing Scale: A Research and Training Manual was purchased (Klein et al., 1969). It provides an overview of the scale, offers procedural suggestions and provides formal training (Klein et al., 1969). Excellent rater reliabilities have been found using this manual. The manual comprises a training manual and original transcripts with accompanying audio recordings. These were used to calibrate initial training on the EXP Scale.

The author of the current study and the collaborating researcher (KN) were trained in securing the ratings that were used in both studies. Initially they used training materials and transcripts to become proficient in using the scale. They independently coded all 90 training recordings from the manual. In addition, 5 PRaCTICED session recordings (that were not drawn from the study sample) were provided to both researchers to practice coding and check for reliability. Therefore, a total of 95 recordings were coded for practice and reliability purposes.

To reduce the chance of therapeutic drift, the procedure required each researcher to code in batches of 10 recordings, after which one of the researchers would code the tenth recording of the other researcher for calibration purposes. Both researchers would meet to discuss their scores for that recording and explain how they came to that conclusion. If there were disagreements in scoring, both researchers would discuss the case until an agreement was reached. This process was repeated after each batch of 10 recordings, with raters alternating rating the additional calibrated coding, until all recordings were complete. By the end of all 40 recording's, each researcher had coded a total of 42 recordings comprising the dataset of 40 sessions and two additional calibration sessions each within the dataset.

Service User Involvement

The PRaCTICED trial was informed by a patient and public group who provided feedback on the initial proposal and were involved in the design of the original study. Further involvement was not possible for this study as the group had disbanded.

Results

Demographic Information

Table 2 provides an overview of the demographics of participants. Participants in both therapies were relatively evenly matched on all four variables, although there were more males in the PCET group. The PCET clients also lived in more deprived areas than the CBT group, although this difference was small. There were no significant differences in age $t(38) = -.97$, $p = .32$, gender X^2 (1, N = 40) = 0.40, $p = 0.53$, ethnicity X^2 (1, N = 40) = 4.00, $p = 0.41$ or Index of Multiple Deprivation X^2 (1, N = 40) = 3.98, $p = 0.91$ between the two groups.

Table 2

Demographic Information Across Each Therapeutic Modality

	PCET ($N = 20$)	CBT ($N = 20$)
Mean Age (SD)	43.10 (13.15)	39.35 (11.09)
Gender ratio Female:Male	8:12	10:10

Ethnicity	18 White British, 2 Global Majority	18 White British, 2 Global Majority
Index of Multiple Deprivation	4.6 (2.96)	5 (3.29)

Note. SD = Standard Deviation; PCET = Person Centred Experiential Therapy; CBT = Cognitive Behavioural Therapy.

Descriptive Statistics

Table 3 provides an overview of the descriptive statistics for both therapies. Participants in both therapies were relatively evenly matched on the three variables. However, those in the PCET group had higher scores on the PHQ-9 and GAD-7 (indicating more severe baseline symptoms of depression and anxiety). PCET clients also received two sessions less on average. There were no significant differences in baseline PHQ-9 X^2 (1, N = 40) = 13.20, p = 0.51, baseline GAD-7 X^2 (1, N = 40) = 14.81, p = 0.25 or mean number of sessions t(38) = -.1.60, p = .80.

Table 3

Descriptive Statistics Across Each Therapeutic Modality.

	PCET (N = 20)	CBT (N = 20)
Baseline Severity PHQ-9 (SD)	20.2 (4.26)	18.5 (4.92)
Baseline Severity GAD-7 (SD)	15.35 (4.41)	13.2 (4.88)
Mean Number of Sessions (SD)	14.4 (4.59)	16.7 (4.26)

Note. SD = Standard Deviation; PCET = Person Centred Experiential Therapy; CBT = Cognitive Behavioural Therapy.

Reliabilities

Interrater reliability of the 95 practice recordings between the two collaborating researchers was assessed based on a mean-rating (k = 2), absolute-agreement, 2-way mixed-effects model of intraclass correlation (3, 2), using 95% confidence intervals. The analysis was conducted twice,

with modal scores and separately with peak scores. The correlations were significant for modal (r = .938, p <.001) and peak ratings (r = .931, p <.001).

Mean Modal Ratings

Table 4 presents the means and standard deviations for clients modal EXP scores.

Initially, the assumptions of the ANOVA were checked. Modal scores were mildly positively skewed; however, the parametric test was used as ANOVA's are robust to small violations of normality (Harwell et al., 1992). Levene's test was not significant at session 2 (p = 0.075) or session 6 (p = 0.668), however Box's M test was significant F(3,259920) = 4.08, p = .01. The significant result means the variances and covariances are non-homogeneous (unequal) across the groups, so the assumption of homogeneity was not satisfied, and an ANOVA carried out in its standard form would not be valid. If Box's M test is significant, it is recommended that Pillai's trace criterion should be used as it is more robust to departures from assumptions (Fidell & Tabachnick, 2003). The profile plot is shown in Figure 1.

Table 4

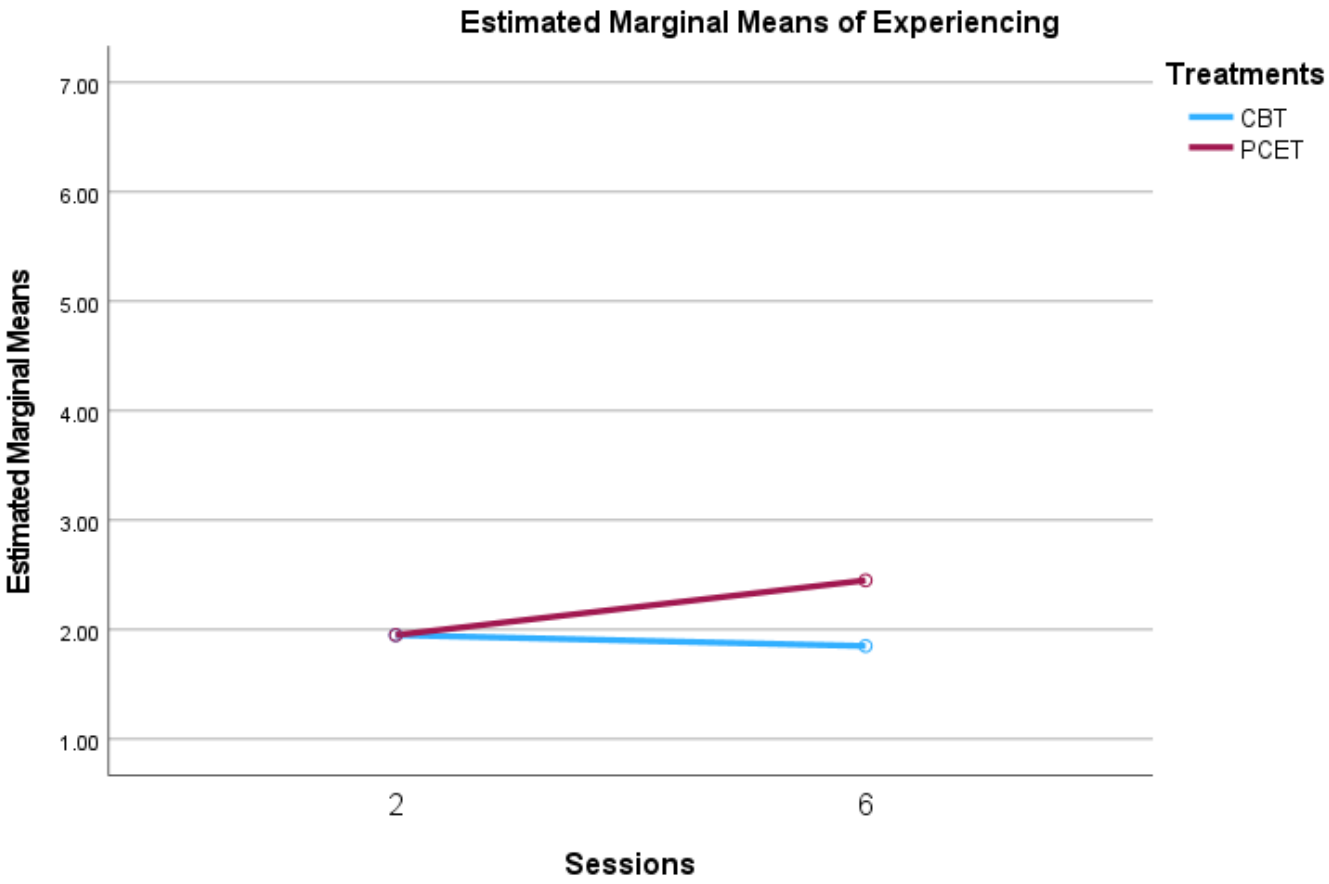
Means (and Standard Deviations) of the Modal EXP Ratings Across Therapy Stage and Therapeutic Modality.

	Therapy Stage	
Therapy Group	Early	Middle
PCET (N = 20)	1.95 (0.22)	2.45 (0.69)
CBT (N = 20)	1.95 (0.51)	1.85 (0.88)
Combined Therapies	1.95 (0.39)	2.15 (0.83)

Note. EXP = Experiencing Scale; PCET = Process Centred Experiential Therapy; CBT = Cognitive-Behavioural Therapy.

Figure 1

Profile Plot for Modal EXP Scores.



There was a significant interaction between therapy type and stage of therapy $F(1,38) = 5.10$, $p = .03$. Whereas the PCET group showed an increase in modal scores from session 2 to 6, the CBT group showed a decline in modal experiencing scores. Simple effects analyses showed that the change in the modal experiencing scores in the PCET group across therapy was significant, $t(38) = -3.10$ $p = .004$, with increased levels of experiencing from early to middle sessions. Whereas, for the CBT group the result was not significant, $t(38) = 0.44$ $p = 0.66$, meaning experiencing levels did not increase from early to middle sessions.

Post-hoc Power Analyses

A sample size calculation was conducted to provide context for whether the study had power to detect desired interaction effects. The effect size for this analysis was calculated using G Power, with a Cohen's f = 0.118 (converted from the ANOVA F score), alpha error probability set at 0.05 and power set at 0.80. A post-hoc analysis of achieved power suggested the analysis was underpowered, while a retrospective power analysis indicated that it would be necessary to recruit

at least 283 participants per group to reliably detect this effect size, with alpha set at .05, at 80% power.

Post-hoc Analyses for Modal Emotional Processing Scores

As there was a significant difference between PCET and CBT in modal EXP scores, it was deemed appropriate to partially replicate the post hoc analyses conducted in the Watson and Bedard (2006) study.

For level 1, there were no significant differences at session 2 between CBT and PCET clients (expected counts were less than 5, so Fisher's exact test was used, $p = 0.30$). At session 6 there was a significant difference ($p = 0.01$), with significantly more CBT than PCET clients being rated at level 1 on the EXP scale.

For level 2, there was no significant differences at session 2 ($p = 0.09$) or 6 (expected counts were over 5 so Chi-square was used) X^2 (1, N = 40) = 0.00, $p = 1.00$.

For level 3, there was no significant difference at session 2 ($p = 0.24$) however there was at session 6 ($p = 0.02$). PCET clients were significantly more likely to be rated at level 3 compared to CBT clients.

For levels 4 and 5 there were no significant differences at session 6 ($p = 0.24$ and $p = 0.50$, respectively).

In summary there were no significant differences at session 2 but significant differences were observed at session 6. Specifically, CBT clients were significantly more likely to process emotions at level 1 and PCET clients at level 3.

Mean Peak Ratings

Table 5 presents the means and standard deviations for clients' peak EXP scores.

Table 5

Means (and Standard Deviations) of the Peak EXP Ratings Across Therapy Stage and Therapeutic Modality.

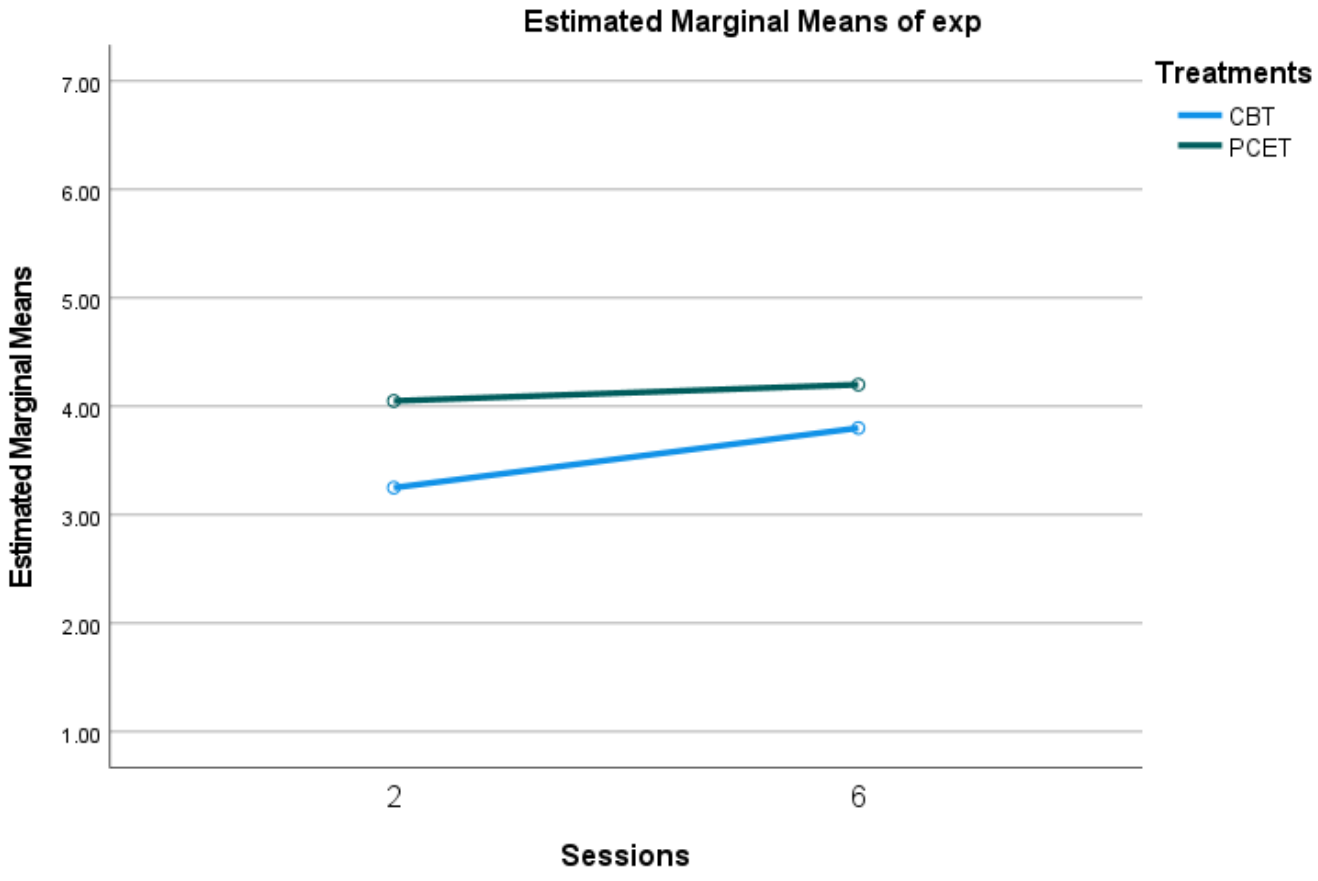
Therapy Group	Therapy Stage	
	Early	Middle
PCET ($N = 20$)	4.05 (0.94)	4.20 (1.11)
CBT ($N = 20$)	3.25 (0.85)	3.80 (1.40)
Combined Therapies	3.65 (0.98)	4.00 (1.26)

Note. EXP = Experiencing Scale; PCET = Process Centred Experiential Therapy; CBT = Cognitive-Behavioural Therapy.

The assumptions of the ANOVA were checked. Scores appeared normally distributed. Both Levene's test and Box's M test were non-significant $F(3,259920) = 0.59$, $p = 0.62$. This means that the assumptions of the ANOVA were met. Figure 2 contains the profile plot.

Figure 2

Profile Plot for Peak EXP Scores.



There was a non-significant interaction between therapy type and stage of therapy $F(1,38) = 0.61$, $p = .44$. Both therapy groups showed a similar non-significant increase in peak experiencing score from early session to middle session.

A main effect of session stage was non-significant, $F(1,38) = 1.87$, $p = .18$, meaning the mean peak levels of reported experiencing did not significantly differ across sessions 2 and 6. However, the mean peak experiencing score was non-significantly higher in session 6 compared to session 2.

There was a significant main effect of therapy type $F(1,38) = 6.62$, $p < .01$. The PCET group had the higher mean peak score and the CBT group had the lower mean peak score.

Post-hoc Power Analyses

A sample size calculation was conducted to provide context for whether the study had power to detect desired effects. The effect size for this analysis was calculated using G Power, with a Cohen's f = 0.016 (converted from the ANOVA F score), alpha error probability set at 0.05 and power set at 0.80. A post-hoc analysis of achieved power suggested the analysis was underpowered, while a retrospective power analysis indicated that it would be necessary to recruit at least 394 participants per group to reliably detect a small effect size, with alpha set at .05, at 80% power.

Post-hoc Analyses for Peak Emotional Processing Scores

As there was a significant difference between PCET and CBT in peak EXP scores, the same post-hoc analyses were conducted, as those carried out for the modal scores.

Firstly, level 2 of the EXP scale was analysed. There were no significant differences between the numbers of PCET and CBT clients rated for level 2 at session 2 ($p = 0.12$) or session 6 ($p = 0.50$).

For level 3, there was no significant difference at session 2 X^2 (1, N = 40) = 1.62, $p = 0.20$, but there was a significant difference at session 6, X^2 (1, N = 40) = 4.80, $p = 0.03$. CBT clients were significantly more likely to process emotions at level 3 compared to PCET clients.

For level 4, there was no significant difference at session 2, X^2 (1, N = 40) = 0.53, p = 0.47, nor at session 6, X^2 (1, N = 40) = 3.14, p = 0.08. Although, PCET clients were non-significantly more likely to process emotions at level 4.

There were no differences in level 5 emotional processing at session 2 (p = 0.12) or session 6 (p = 0.12). There were also no significant differences in level 6 emotional processing at session 2 (p = 0.50) or session 6 (p = 0.33).

In summary there were no differences in session 2 but differences were observed at session 6. Specifically, CBT clients were significantly more likely to process emotions at level 3. There was also a non-significant trend for more PCET than CBT clients to be working at level 4.

Discussion

PCET and CBT have been shown to have similar outcomes for depression at the end of therapy within NHS Talking Therapies services, although results from the recent PRaCTICED trial have reported results favouring CBT compared with PCET at 12-month post-randomisation (Barkham et al., 2021). The primary aim of the current study was to conduct a partial replication of Watson and Bedard's (2006) study by comparing emotional processing in clients accessing either PCET or CBT in a subsample of patients from the PRaCTICED trial. The purpose was to gain a better understanding of the role of emotional processing in both therapies. It was hoped the results of the current study would identify differential processes within the therapies that may have contributed to these outcomes.

Regarding the first hypothesis which explored whether there would be significant differences in depths of emotional processing between the two therapy groups, clients who accessed PCET had significantly higher modal levels of emotional processing at session 6 and higher peak levels of emotional processing at session 2 and session 6 compared to clients accessing CBT. Although the scores were higher in PCET, the highest levels of emotional processing, according to the EXP scale, were not evident in the study for either therapy.

The secondary aim of the study was to understand whether depressed clients' emotional processing would increase over the course of the therapeutic journey in each therapy. The findings from this analysis are mixed. For modal scores, PCET modal EXP scores increased from early to middle therapy, while CBT modal EXP scores decreased from early to middle therapy. For peak scores, emotional processing did not significantly differ across sessions 2 and 6. However, the mean peak experiencing score was non-significantly higher in session 6 compared to session 2 for both therapies.

Comparison to Current Literature

The results support the previous literature in that clients who engage in experiential approaches process their emotions in greater depth than CBT (Klug et al., 2021; Pinheiro et al., 2021; Rudkin et al., 2007; and Watson & Bedard, 2006). During follow-up analyses Watson and Bedard found that CBT clients were significantly more likely to process their emotions at level 2, whilst PET clients were significantly more likely to process their emotions at level 3. Within the current study, significant differences were also found for the two therapies studied. It was found that at the middle stage of therapy, CBT clients were significantly more likely to process emotions for the majority of the time at level 1 and PCET clients at level 3.

Analyses do not support the hypotheses, that depth of emotional processing would change from early to middle therapy. The hypotheses were predominantly based on the Watson and Bedard (2006) study findings and Pascual-Leone and Yeryomenko's meta-analysis (2016). Both these found that depth of experiencing increased from early to middle sessions. As a result of the discrepancy, the current paper explored the findings of the meta-analysis in more depth to review their conclusions. Only five studies were included in that analysis and the authors noted that due to such a low number of studies included in the analyses, the findings should be considered 'tenuous at best'. Four of the studies saw consistent increases in EXP from early to middle sessions (Burgess-Moser, 2012; Hakim, 2010; Pos, Greenberg & Warwar, 2009; and Watson & Bedard, 2006). However, the Ralston study (2006) did not find a significant difference from early to middle

sessions for people accessing EFT. Given the limited literature and the differing results of the previous studies, the current study could have been more tentative in their hypotheses related to this question.

A further reason why a relationship was possibly not found was because the study was underpowered. It is recognised that by the nature of process research, the designs are often limited by small samples as researchers are exploring small elements of interactions in therapy (Llewelyn & Hardy, 2001). Power is influenced by the size of the sample; the size of the effect being detected and by the noise (uncontrolled factors) (Dix, 2020). Given that process research is typically limited in sample size numbers, one approach to increase power is therefore to reduce noise (Dix, 2020). By developing and following a clear methodology, removing extraneous variables where possible, and having relative homogeneity between the two groups characteristics, the current study has made attempts to maximise power whilst acknowledging the limited sample size. Rosnow and Rosenthal (2003) have also argued that a further way of increasing power is by maximising contrasts between the two groups. Given the aim of the current study was to compare the phenomenon of experiencing in two contrasting therapy modalities, the sampling strategy used was designed to maximise differential effects between both groups. However, future research could increase their sample size as a third method of maximising power.

Theory Development

Within the limitations of the Pascual-Leone and Yeryomenko (2016) meta-analysis which explored studies using the EXP Scale, they highlighted the general lack of studies in the field. They went on to propose that there is a need for further research in the field which makes use of a greater range of therapies. Therefore, the current study adds valuable and crucial information to the current field with regards to the role of emotional processing in an alternative therapy, PCET. It also increases evidence that suggests the role of emotional processing differs between experiential therapies and CBT, with emotions being explored in greater depth in experiential therapies. Nevertheless, emotional processing still appears an important variable in CBT.

Pascual-Leone and Yeryomenko (2016) acknowledged the sampling within previous studies was an issue. This is because they found of the studies they identified; the majority were re-analysing data from only several datasets. In addition, previous studies (Watson & Bedard, 2006; Watson et al., 2011) highlighted that their samples would benefit from greater numbers of men given the high proportion of female clients. Therefore, the current study is valuable in that it has analysed data on participants not studied before and it has a better representation of both sexes (18 females and 22 males), albeit this was due to chance rather than study design.

A novel finding from this paper was in relation to emotional processing and stage of therapy. The current study found that PCET EXP modal scores increased from early to middle sessions, but CBT EXP modal scores reduced. This finding has only been reported in one other paper, Rudkin et al., (2007). This finding suggests that both therapies may explore emotions to a similar extent in early sessions, however as therapy progresses, emotional content may be processed in different ways, with PCET clients consistently processing emotion in greater depth. This may suggest that other mechanisms are potentially more influential in the working stages of CBT.

Limitations and Future Directions

Given the current study was interested in exploring mechanisms of change in both therapies, the methodology allowed for the exploration of change in emotional processing across both therapies in early and middle sessions. However, conclusions about mechanisms of change based on the results of this study alone should be tentative, given the context of the relationship with outcome was not directly tested. Future research would benefit from using a methodology which would facilitate the exploration of emotional processing in good and poor outcome cases from the two therapies at different time points.

The current study planned to look at early and middle segments of therapy by studying sessions 2 and 6. Given that one of the inclusion criteria was a minimum of 7 sessions, it was possible that late sessions rather than middle were being investigated for clients if they only received 7 sessions. This may have influenced the analysis of level of EXP at different time points, as previous research

has found that early and late sessions have more similar EXP scores than early and middle sessions (Watson & Bedard, 2006). Coincidentally, there were no clients in the current sample who had less than 8 sessions. Future research could avoid this potential limitation by studying session 6 but changing the inclusion criteria so participants had at least 8 sessions.

Given the finding that there is a greater focus on emotional processing over time, with some indication of an increase in depth for PCET, future research could explore EXP in later sessions for PCET clients. A similar methodology could be used to the Watson and Bedard (2006) study.

At times the researchers deviated from the advice in the EXP manual due to replicating the Watson and Bedard study. For example, the manual suggests that recordings of 5 to 8 minutes provide enough material to identify levels of emotional processing without becoming 'overly complex or tedious'. Given that the literature suggests there is no significant difference in EXP scores when altering the duration of codings (Klein et al., 1969), future research could code shorter durations of therapy segments. A second deviation from the manual was the time within sessions that were coded. The current study coded 20 minutes after the first 20 minutes of therapy whereas the manual recommends that when length of session durations differ, to sample proportionately (from the same relative point in time). Future research could follow the recommendations from the manual and code the same relative point in time.

There are some limitations which are typically associated with process research. Given both raters knowledge of therapy it was impossible to blind them to the therapy they were coding. Future research would benefit from coders who have no background knowledge of the therapies. If this is not feasible, there could be an increase in the number of reliabilities checks between the two raters (this was not conducted in the current study due to the excellent reliability in the practice tapes). Alternatively, future research could replicate Watson and Bedard's process of employing a third rater to code a percentage of recordings.

The sample was limited to depressed clients, the vast majority of whom were White British and middle aged. The generalisability of the results is affected due to the homogenous nature of the

sample. In addition, conclusions are tentative in nature due to the small sample. Future research should aim to study a more diverse sample including a variety of presenting problems.

Qualitatively, the researchers observed that some recordings had either much longer durations and/or frequencies at peak EXP whereas other recordings only reached their peak EXP on one occasion for a short duration of time. Specifically, it was observed that PCET had longer durations and higher number of frequencies at peak EXP compared to CBT. This is unsurprising given Wiser and Goldfried (1998) had previously explored some factors that influence emotional processing and found that more controlled therapies (CBT) were likely to dampen emotional processing. It would be interesting for future research to investigate the duration and/or frequency of peaks and analyse whether the length and frequency of depth of emotional processing influences outcomes. Wiser and Goldfried also found that therapists' approaches influenced emotional processing, specifically when CBT therapists intervened from an affiliative and controlling interpersonal stance and switched between emotional and non-emotional material, clients emotional processing reduced. Future research could explore the role of therapist effects in client emotional processing.

Clinical Implications

There have not been any studies which have explored levels of emotional processing in PCET. Therefore, this study provides the first details regarding the role of emotional processing in PCET which has clinical implications such as understanding mechanisms of change within the therapy. However, conclusions about mechanisms of change based on the results of this study alone should be tentative, given the context of the relationship with outcome was not directly tested.

This study builds on Castonguay et al.'s (1996) recommendations with regards to conducting research which aims to gain a better understanding of the role of emotional processing in the change process of CBT.

Having a clearer understanding of the role of experiencing in both therapies could improve therapy as experienced by clients. We know that some people have greater capacity to process emotions (Watson & Greenberg, 2017) and we know that some therapies require greater levels of

emotional processing to lead to better outcomes. Therefore, over time we may move to personalised therapy, which is the idea that some therapies may result in better outcomes for certain people with certain features. Studies such as Delgadillo and Gonzalez Salas Duhne (2020) found that clients who received treatments which best matched their features had better outcomes than those who receive suboptimal treatment.

Having a stronger understanding of the role of emotional processing in both therapies could also influence therapists by informing clinical practice through the training of clinicians (to work optimally). For example, it was evident that the higher levels of emotional processing were not regularly achieved for either therapy, therefore, PCET training could focus more on encouraging greater depth in emotional processing. Additionally, supervision of PCET therapists in comparison to CBT therapists may differ, with greater focus on emotional processing for PCET therapists. Also, the emphasis on emotional processing in sessions and supervision may differ at different stages of the therapeutic process in CBT and PCET.

Conclusions

In conclusion, there were significant differences in level of experiencing between CBT and PCET. Previous studies have found CBT clients are more separated from their emotional experience and experiential therapies have been associated with greater focus on emotional processing. The current study corroborates those conclusions.

Contrary to the majority of previous findings, there was not a significant difference in depth of experiencing across early or middle stages of therapy.

Given this was the first research exploring experiencing in PCET it is recommended that further research is conducted to further understand the role of emotional processing in PCET.

References

American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental health disorders: DSM-5* (5th ed.). Washington, DC: American psychiatric association. https://doi.org/10.1176/appi.books.9780890425596

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). Person-centred experiential therapy versus cognitive behavioural therapy delivered in the English Improving Access to Psychological Therapies service for the treatment of moderate or severe depression (PRaCTICED): a pragmatic, randomised, non-inferiority trial. *The Lancet Psychiatry, 8 (6),* 487-499. https://doi.org/10.1016/S2215-0366(21)00083-3

Beck, A. T., Steer, R. A., & Brown, G. (1996). *Beck Depression Inventory–II (BDI-II) [Database record]*. APA PsycTests. https://doi.org/10.1037/t00742-000

Berelson, B. (1952). *Content analysis in communication research*. New York: Hafner Press.

Blackburn, I. M., James, I. A., Milne, D. L., Reichelt, F. K., Garland, A., Baker, C., Standard, A., & Claydon, A. (2001). Cognitive therapy scale—revised (CTS-R). *Newcastle-upon-Tyne: Newcastle Cognitive and Behavioural Therapies Centre*. https://doi.org/10.1017/S1352465801004040

Bommert, R., & Dahlhoff, H. D. (1978). Das Selbsterleben (Experiencing) *in der Psychotherapie [Self experiencing in psychotherapy]*. Munich, Germany: Urban and Schwarzenberg. https://doi.org/10.1037/0022-006X.74.1.152

Brydges, C. R. (2019). Effect size guidelines, sample size calculations, and statistical power in gerontology. *Innovation in Aging*, *3*(4), igz036. https://doi.org/10.1093/geroni/igz036

Burgess Moser, M. (2012). *The cognitive-affective and behavioural impact of emotionally focused couple therapy* [Doctoral dissertation, Université d'Ottawa/University of Ottawa]. University of Ottawa Online Repository. https://doi.org/10.20381/ruor-6474

Campelo, F., & Wanner, E. F. (2020). Sample size calculations for the experimental comparison of multiple algorithms on multiple problem instances. *Journal of Heuristics*, *26*(6), 851-883. https://doi.org/10.1007/s10732-020-09454-w

Carapetis, J. R., & Dadi, A. F. (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet, 390*(10100), 1211-1259. https://doi.org/10.1016/S0140-6736(17)32154-2

Castonguay, L. G., Goldfried, M. R., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. *Journal of Consulting and Clinical Psychology, 64*(3), 497–504. https://doi.org/10.1037/0022-006X.64.3.497

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//0033-2909.112.1.1555

Connor, K. M., & Davidson, J. R. (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and Anxiety, 18*(2), 76-82. https://doi.org/10.1002/da.10113

Crits-Christoph, P., Gibbons, M. B. C., & Mukherjee, D. (2013). Psychotherapy Process-Outcome Research. In M. J. Lambert (6th Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 298-340). John Wiley and Sons Inc.

Crits-Christoph, P., & Connolly Gibbons, M. B. (2021). Psychotherapy process-outcome research: Advances in understanding causal connections. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 263–296). John Wiley and Sons Inc.

Delgadillo, J., & Gonzalez Salas Duhne, P. (2020). Targeted prescription of cognitive–behavioral therapy versus person-centered counseling for depression using a machine learning approach. *Journal of Consulting and Clinical Psychology, 88*(1), 14-24. https://doi.org/10.1037/ccp0000476

di Stephano, J. (2003). How much power is enough? Against the development of an arbitrary convention for statistical power calculations. *Functional Ecology, 17*(5), 707-709. https://www.jstor.org/stable/3599167

Dix, A. (2020). *Statistics for HCI: Making sense of quantitative data*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-02228-9_10

Drisko, J. W., & Maschi, T. (2016). *Content analysis*. Oxford University Press, USA.

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

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE–OM. *The British Journal of Psychiatry, 180*(1), 51-60. https://doi.org/10.1192/bjp.180.1.51

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*(2), 175–191. https://doi.org/10.3758/BF03193146

Fidell, L. S., & Tabachnick, B. G. (2003). Preparatory data analysis. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology (*pp. 115–141). John Wiley & Sons. https://doi.org/10.1002/0471264385.wei0205

Flanagan, J. C. (1982). Measurement of quality of life: current state of the art. *Archives of Physical Medicine and Rehabilitation*, *63*(2), 56-59. https://europepmc.org/article/med/6460487

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

Gendlin, E. T. (1962). *Experiencing and the creation of meaning*. The Free Press of Glencoe.

Gendlin, E. T., Beebe, J., Cassens, J., Klein, M., & Oberlander, M. (1968). Focusing ability in psychotherapy, personality, and creativity. In J. M. Shlein (Eds.), *Research in Psychotherapy* (pp. 217–238). American Psychological Association.

Giddings, L. S. (2006). Mixed-methods research: Positivism dressed in drag?. *Journal of Research in Nursing*, *11*(3), 195-203. https://doi.org/10.1177/1744987106064635

Greenberg, L. S., & Safran, J. D. (1987). *Emotion in psychotherapy: Affect, cognition, and the process of change.* Guilford. https://doi.org/10.1037/0003-066X.44.1.19

Hakim, L. Z. (2010). *A comparative analysis of the qualities of therapists' communications in three empirically-supported psychotherapies* [Unpublished Doctoral Dissertation, York University]. Library and Archives Canada. https://www.collectionscanada.gc.ca/obj/thesescanada/vol2/002/NR80596.PDF?is_thesis=1&oclc_number=879940060

Harwell, M. R., Rubinstein, E. N., Hayes, W. S., & Olds, C. C. (1992). Summarizing Monte Carlo results in methodological research: The one-and two-factor fixed effects ANOVA cases. *Journal of Educational Statistics*, *17*(4), 315-339. https://doi.org/10.3102/10769986017004315

Hendricks, M. (2009). Experiencing level: An instance of developing a variable from a first-person process so it can be reliably measured and taught. *Journal of Consciousness Studies, 16*(10–12), 129–155. https://www-ingentaconnect-com.sheffield.idm.oclc.org/content/imp/jcs/2009/00000016/f0030010/art00006#

Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, *20*(10), 1727–1736. https://doi.org/10.1007/s11136-011-9903-x

Hill, A. (2011). Curriculum for counselling for depression: Continuing professional development for qualified therapists delivering high intensity interventions. *National IAPT Programme Team.*

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Hubble, M. A., Duncan, B. L., & Miller, S. D. (1999). *The heart and soul of change: Delivering what works in therapy.* American Psychological Association. https://doi.org/10.1037/11132-000

Institute of Health Metrics and Evaluation (2021). *Global Health Data Exchange (GHDx).* Retrieved from: http://ghdx.healthdata.org/gbd-results-tool?params=gbd-api-2019-permalink/d780dffbe8a381b25e1416884959e88b

Isgar, J. (2024). Clients *emotional processing in psychotherapy: a comparison between CBT & PCET* [Unpublished DClinPsy Dissertation]. University of Sheffield, England, United Kingdom.

Kim, H. Y. (2016). Statistical notes for clinical researchers: Sample size calculation 2. Comparison of two independent proportions. *Restorative Dentistry and Endodontics*, *41*(2), 154-156. https://doi.org/10.5395/rde.2016.41.2.154

Klein, M. H., Mathieu, P. L., Gendlin, E. T, & Kiesler, D. J. (1969). *The Experiencing Scale: A research and training manual* (Vols. 1 and 2). Madison: University of Wisconsin Extension Bureau of Audiovisual Instruction.

Klein, M. H., Mathieu-Coughlan, P., & Kiesler, D. J. (1986). The Experiencing Scales. In L. S. Greenberg and W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 21–71). Guilford Press.

Klug, G., Seybert, C., Ratzek, M., Grimm, I., Zimmermann, J., & Huber, D. (2021). Insight and outcome in long-term psychotherapies of depression. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, *68*(1), 54-73. https://doi.org/10.13109/zptm.2021.67.oa10

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Kraus, D. R., Castonguay, L., Boswell, J. F., Nordberg, S. S., & Hayes, J. A. (2011). Therapist effectiveness: Implications for accountability and patient care. *Psychotherapy Research*, *21*(3), 267-276. https://doi.org/10.1080/10503307.2011.563249

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

Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2010). The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. *General Hospital Psychiatry*, *32*(4), 345-359. https://doi.org/10.1016/j.genhosppsych.2010.03.006

Larsen, D. L., Attkisson, C. C., Hargreaves, W. A. & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning*, *2*(3), 197-207. https://doi.org/10.1016/0149-7189(79)90094-6

Lewis, G., Pelosi, A., 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

Llewelyn, S., & Hardy, G. (2001). Process research in understanding and applying psychological therapies. *British Journal of Clinical Psychology*, *40*(1), 1-21. https://doi.org/10.1348/014466501163436

Lutz, W., DeJong, K., Rubel, J. A., & Delgadillo, J. (2021). Measuring, predicting and tracking change in psychotherapy. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 89–134). John Wiley and Sons Inc.

Maas, C. J. M. & Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Statistica Neerlandica, 58*(2), 127-137. https://doi.org/10.1046/j.0039-0402.2003.00252.x

Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. H. (2002). The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *The 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: a manual for training and practice*. SAGE Publishing.

NICE: National Institute for Health and Care Excellence. (2022). Depression in adults: treatment and management. Retrieved from: https://www.nice.org.uk/guidance/ng222

Norcross, J. C., & Lambert, M. J. (2019). What works in the psychotherapy relationship: Results, conclusions, and practices. In J. C. Norcross & M. J. Lambert (Eds.), *Psychotherapy relationships that work: Evidence-based therapist contributions* (pp. 631–646). Oxford University Press. https://doi.org/10.1093/med-psych/9780190843953.003.0018

Neuendorf, K. A. (2002). *The content analysis guidebook*. SAGE Publishing. https://doi.org/10.4135/9781071802878

Park, Y. S., Konge, L., & Artino Jr, A. R. (2020). The positivism paradigm of research. *Academic Medicine, 95*(5), 690-694. https://doi.org/10.1097/ACM.0000000000003093

Pascual-Leone, A., Paivio, S., & Harrington, S. (2016). Emotion in psychotherapy: An experiential-humanistic perspective. In D. Cain, S. Rubin, & K. Keenan (Eds.), *Humanistic psychotherapies: Handbook of research and practice* (2nd ed.). (pp. 147–181). American Psychological Association. https://doi.org/10.1037/14775-006

Pascual-Leone, A., & Yeryomenko, N. (2016). The client "experiencing" scale as a predictor of treatment outcomes: A meta-analysis on psychotherapy process. *Psychotherapy Research, 27*(6), 653-665. https://doi.org/10.1080/10503307.2016.1152409

Pinheiro, P., Gonçalves, M. M., Sousa, I., & Salgado, J. (2021). What is the effect of emotional processing on depression? A longitudinal study. *Psychotherapy Research*, *31*(4), 507-519. https://doi.org/10.1080/10503307.2020.1781951

Pos, A. E., Greenberg, L. S., Goldman, R. N., & Korman, L. M. (2003). Emotional processing during experiential treatment of depression. *Journal of Consulting and Clinical Psychology*, *71*(6), 1007-1016. https://doi.org/10.1037/0022-006X.71.6.1007

Pos, A. E., Greenberg, L. S., & Warwar, S. H. (2009). Testing a model of change in the experiential treatment of depression. *Journal of Consulting and Clinical Psychology*, *77*(6), 1055-1066. https://doi.org/10.1037/a0017059

Ralston, M. B. (2006). *Imaginal confrontation versus evocative empathy in emotion-focused trauma therapy* [Unpublished doctoral dissertation, University of Windsor]. Electronic Theses and Dissertations. https://scholar.uwindsor.ca/etd/2382

Rice, L. N. & Greenberg, L. S. (1984). *Patterns of change*. Guilford.

Rogers, C. (1951). *Client-centered therapy: Its current practice, implications and theory*. Constable.

Rogers, C. R. (1959). A theory of therapy, personality, and interpersonal relationships, as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of science* (Vol. 3, pp. 184–256). New York: McGraw-Hill.

Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale, 57*(3), 221–237. https://doi.org/10.1037/h0087427

Roth, A., Hill, A., & Pilling, S. (2009). The competences required to deliver effective Humanistic Psychological Therapies. *London: University College London*.

Rudkin, A., Llewelyn, S., Hardy, G., Stiles, W. B., & Barkham, M. (2007). Therapist and client processes affecting assimilation and outcome in brief psychotherapy. *Psychotherapy Research, 17*(5), 613-621. https://doi.org/10.1080/10503300701216298

Ryan, R. (1966). *The role of the experiencing variable in the psychotherapeutic process* [Unpublished Doctoral Dissertation, University of Illinois]. ProQuest Dissertations & Theses Global. https://www.proquest.com/docview/302210368?pq-origsite=gscholar&fromopenview=true&sourcetype=Dissertations%20&%20Theses

Salgado, J. (2014). ISMAI depression project: Results of a comparative clinical trial of EFT and CBT. In *45th Society for Psychotherapy Research Annual Meeting, Copenhagen, Denmark*.

Saxon, D., & Barkham, M. (2012). Patterns of therapist variability: therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology, 80*(4), 535-546. https://doi.org/10.1037/a0028898

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R. & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59*(20), 22-33. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9881538

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). SAGE Publishing.

Sønderland, N. M., Solbakken, O. A., Eilertsen, D. E., Nordmo, M., & Monsen, J. T. (2023). Emotional changes and outcomes in psychotherapy: A systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology*. Advance online publication. https://doi.org/10.1037/ccp0000814

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archive of Internal Medicine, 166*(10), 1092-1097. https://doi.org/10.1001/archinte.166.10.1092

Stiles, W. B. (2021). Responsiveness in psychotherapy research: Problems and ways forward. In J. C. Watson, & H. Wiseman (Eds.), *The responsive psychotherapist: Attuning to clients in the*

moment (pp. 15–35). American Psychological Association. https://doi.org/10.1037/0000240-002

Town, J. M., Salvadori, A., Falkenström, F., Bradley, S., & Hardy, G. (2017). Is affect experiencing therapeutic in major depressive disorder? Examining associations between affect experiencing and changes to the alliance and outcome in intensive short-term dynamic psychotherapy. *Psychotherapy*, *54*(2), 148–158. https://doi.org/10.1037/pst0000108

Tripathy, J. P. (2013). Secondary data analysis: Ethical issues and challenges. *Iranian Journal of Public Health, 42(12)*, 1478-1479. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4441947/

Ulvenes, P. G., Berggraf, L., Wampold, B. E., Hoffart, A., Stiles, T., & McCullough, L. (2014). Orienting patient to affect, sense of self, and the activation of affect over the course of psychotherapy with cluster C patients. *Journal of Counseling Psychology*, *61*(3), 315-324. https://doi.org/10.1037/cou0000028

Wang, Y. P., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. *Brazilian Journal of Psychiatry*, *35*(4), 416-431. https://doi.org/10.1590/1516-4446-2012-1048

Watson, J. C., & Bedard, D. L. (2006). Clients' emotional processing in psychotherapy: A comparison between cognitive-behavioral and process-experiential therapies. *Journal of Consulting and Clinical Psychology*, *74(1)*, 152-159. https://doi.org/10.1037/0022-006X.74.1.152

Watson, J. C., & Greenberg, L. S. (2017). *Emotion-focused therapy for generalized anxiety.* American Psychological Association. https://doi.org/10.1037/0000018-000

Watson, J. C., McMullen, E. J., Prosser, M. C., & Bedard, D. L. (2011). An examination of the relationships among clients' affect regulation, in-session emotional processing, the working alliance, and outcome. *Psychotherapy Research*, *21*(1), 86-96. https://doi.org/10.1080/10503307.2010.518637

Watson, J. C. (2018). Mapping patterns of change in emotion-focused psychotherapy: Implications for theory, research, practice, and training. *Psychotherapy Research*, *28*(3), 389-405. https://doi.org/10.1080/10503307.2018.1435920

Watson, J. C. (2023). Psychotherapy process research: Identifying productive in-session processes to enhance treatment outcomes and therapist responsiveness. *Psychotherapy Research*, 1-13. Advance online publication. https://doi.org/10.1080/10503307.2023.2252160

Watson, J. C., & Wiseman, H. (2024). Responsiveness: Can psychotherapies and psychotherapists adapt effectively to participant and clinical contexts? In F. T. L. Leong, J. L. Callahan, J. Zimmerman, M. J. Constantino, & C. F. Eubanks (Eds.), *APA handbook of psychotherapy: Evidence-based practice, practice-based evidence, and contextual participant-driven practice* (pp. 95–115). American Psychological Association. https://doi.org/10.1037/0000354-007

Westerman, M. A., & de Roten, Y. (2017). Investigating how interpersonal defense theory can augment understanding of alliance ruptures and resolutions: A theory-building case study. *Psychoanalytic Psychology*, *34*(1), 13–25. https://doi.org/10.1037/pap0000122

Wiser, S. L., & Goldfried, M. R. (1993). Comparative study of emotional experiencing in psychodynamic-interpersonal and cognitive-behavioral therapies. *Journal of Consulting and Clinical Psychology*, *61*(5), 892-895. https://doi.org/10.1037//0022-006x.61.5.892.

Appendices

Appendix A – University Ethics Approval

The University Of Sheffield.

Downloaded: 23/12/2022
Approved: 20/12/2022

Jack Isgar
Registration number: 210154881
Psychology
Programme: Doctorate in Clinical Psychology

Dear Jack

PROJECT TITLE: Clients' Emotional Processing in Psychotherapy: A Comparison Between Cognitive-Behavioural and Person-Centred Experiential Therapies. A secondary analysis of the PRaCTICED trial data set
APPLICATION: Reference Number 050463

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 20/12/2022 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 050463 (form submission date: 06/12/2022); (expected project end date: 31/05/2024).

The reviewers have left the following comments reguarding the application:

Based on the attached DMP, it appears that all data is anonymised. In which case I believe that you are NOT processing special category personal data. (section E, question 2).

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since written approval will be required.

Your responsibilities in delivering this research project are set out at the end of this letter.

Yours sincerely

Department Of Psychology Research Ethics Committee
Ethics Administrator
Psychology

Please note the following responsibilities of the researcher in delivering the research project:

- The project must abide by the University's Research Ethics Policy:https://www.sheffield.ac.uk/research-services/ethics-integrity/policy
- The project must abide by the University's Good Research & Innovation Practices Policy: https://www.sheffield.ac.uk/polopoly_fs/1.671066!/file/GRIPPolicy.pdf
- The researcher must inform their supervisor (in the case of a student) or Ethics Administrator (in the case of a member of staff) of any significant changes to the project or the approved documentation.
- The researcher must comply with the requirements of the law and relevant guidelines relating to security and confidentiality of personal data.
- The researcher is responsible for effectively managing the data collected both during and after the end of the project in line with best practice, and any relevant legislative, regulatory or contractual requirements.

Appendix B – AsPredicted Registration

ASPREDICTED

Wharton CREDIBILITY LAB
University of Pennsylvania

CONFIDENTIAL - FOR PEER-REVIEW ONLY

Clients ' Emotional Processing in Psychotherapy: A Comparison Between CBT & PCET (#140203)

Created: 08/04/2023 03:32 AM (PT)

This is an anonymized copy (without author names) of the pre-registration. It was created by the author(s) to use during peer-review.
A non-anonymized version (containing author names) should be made available by the authors when the work it supports is made public.

---

1) Have any data been collected for this study already?

It's complicated. We have already collected some data but explain in Question 8 why readers may consider this a valid pre-registration nevertheless.

2) What's the main question being asked or hypothesis being tested in this study?

Research aims:

1. To compare depressed clients' levels of emotional processing in cases of Cognitive Behavioural Therapy (CBT) and Person-Centred Experiential Therapy (PCET)
2. To see whether depressed clients' emotional processing increases over the course of therapy in each approach

Hypotheses:

1. CBT involves explaining and reducing emotions in a cognitive framework, whereas PCET has greater focus on emotional processing, therefore clients in the PCET group would have higher levels on the Experiencing Scale (EXP) for both modal and peak scores than those in CBT
2. Watson and Bedard (2006) found clients' level of emotional processing significantly increased from the beginning to the midpoint of therapy. Therefore it is expected clients' EXP scores would be higher at the midpoint of therapy

3) Describe the key dependent variable(s) specifying how they will be measured.

Depth of experiencing measured using the EXP scale at early and mid-therapy for CBT and PCET clients with the best and worst outcomes

4) How many and which conditions will participants be assigned to?

The study sample will consist of 40 patients: 10 CBT good outcome, 10 CBT poor outcome, 10 PCET good outcome and 10 PCET poor outcome. Outcomes are based on participants who show the largest and smallest change in PHQ-9 scores from first to last session.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

Each client response occurring between one therapist response, and another will be rated on the EXP scale, using an approach akin to Content Analysis. Resulting in a modal score (most frequently occurring EXP level in an excerpt) and a peak score (highest EXP level reached in the excerpt). To determine whether there are any differences between the two therapies in clients' emotional processing and stage of therapy, the mean modal and peak scores of the EXP ratings will be analysed using a 2 x 2 ANOVA.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

There will be no outliers. Participants from the original sample will only be excluded if they had less than 7 sessions (due to the requirement of sampling sessions 2 and 6), switched therapies, did not have outcome measures taken at first and last session and their sessions had not been recorded.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

Initially researchers will assess whether patients are eligible for the study by applying inclusion criteria. Once these criteria are applied, remaining participants will be selected using their PHQ-9 data. Specifically, participants who showed the largest and smallest change in PHQ-9 scores as measured in their first session of therapy and last session of therapy.

The study sample consists of 40 patients: 10 CBT good outcome, 10 CBT poor outcome, 10 PCET good outcome and 10 PCET poor outcome. The middle 20-minute segment of sessions 2 and 6 for each of these clients will be coded (total of 80). This will allow comparisons between therapies (the current study) and between good and poor outcomes (a study being conducted by Katie Neal).

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

The current study uses a secondary analysis design of IAPT data collected from the "pragmatic non-inferiority randomised trial of the clinical and cost of counselling for depression versus CBT (PRaCTICED)" (Barkham et al., 2021). This trial used a pragmatic, randomised, non-inferiority design and was conducted within the Sheffield NHS Talking Therapies service. It aimed to explore the outcomes of CBT verses PCET for moderate to severe depression.

Appendix C - The STROBE Checklist

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	73
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	73
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	76-78
Objectives	3	State specific objectives, including any prespecified hypotheses	79
Methods			
Study design	4	Present key elements of study design early in the paper	79-80
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	80-81
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	80-81
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	80-81
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*82, 84 - 85*
Bias	9	Describe any efforts to address potential sources of bias	87-88
Study size	10	Explain how the study size was arrived at	84
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	86-87
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	84-87
		(*b*) Describe any methods used to examine subgroups and interactions	86-87
		(*c*) Explain how missing data were addressed	N/A
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	84-85
		(*e*) Describe any sensitivity analyses	N/A

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	84-85
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	88-90
		(b) Indicate number of participants with missing data for each variable of interest	N/A
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	Report numbers of outcome events or summary measures over time	*88-90*
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	91, 93-94.
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	92, 94.
Discussion			
Key results	18	Summarise key results with reference to study objectives	95.
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	98-100.
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	99-100.
Generalisability	21	Discuss the generalisability (external validity) of the study results	100-101.
Other information			

Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	N/A.

Appendix D - Patient Health Questionnaire 9 (PHQ- 9)

Over the last 2 weeks, how often have you been bothered by any of the following problems? *(Use "✔" to indicate your answer)*	**Not at all**	**Several days**	**More than half the days**	**Nearly every day**
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself — or that you are a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed? Or the opposite — being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead or of hurting yourself in some way	0	1	2	3

FOR OFFICE CODING 0 + _____ + _____ + _____

=Total Score: _____

Appendix E – The Experiencing Scale

the experiencing scale

The Experiencing Scale (EXP) describes 7 levels of emotional and cognitive involvement with one's ongoing experience. Feeling safe enough to & knowing how to work sometimes at deeper levels typically leads to better progress. This improved effectiveness is probably due to several factors that include a.) clearer understanding at head, heart & gut levels of why one thinks, feels & reacts as one does. b.) more contact with feelings allows them to be worked on, processed & integrated to produce better balance in thoughts, emotions and behaviours. c.) stronger connection with healthy, appropriate emotions can energise us to act to change outer circumstances in helpful ways.

1.) the client simply talks about events, ideas or others

2.) refers to self but without expressing emotions.

3.) expresses emotions but only as they relate to external circumstances.

4.) the client focuses directly on emotions and thoughts about self

5.) engages in an exploration of his or her inner experience

6.) gains awareness of previously implicit feelings & meanings

7.) on-going process of in-depth self-understanding, which provides new perspectives to solve significant problems

---

The scale has been shown to predict client change, especially in client-centred therapy, but it is relevant for cognitive therapy and for group work too.

Klein M H, Mathieu P L, Gendlin E T & Kiesler D J
The experiencing scale: a research and training manual (vol 1)
Madison: University of Wisconsin, 1969

Appendix F – Collaboration Statement

This statement outlines the contributions of this thesis that were undertaken jointly by fellow researcher KN and I. These contributions were undertaken equally. All other work in this thesis was undertaken independently.

Work conducted in collaboration:

- Data management plan was written by me and used as a template for KN's plan.
- AsPredicted protocol was written by me and used as a template for KN's protocol.
- Coding of the 95 practice transcripts to check reliabilities.
- Coding of the 80 recordings, each of us coded 42 recordings.

- Descriptive statistics of included participants.

Appendix G – Eight-Stage Content Analysis Process

Stages	Description in current study
1. Preparation of the data	Participants were screened against inclusion criteria. Those who met criteria were systemically selected using their PHQ-9 data.
2. Deciding how the data will be measured	Speech was coded on the EXP Scale.
3. Identifying the coding system	Turn talks of more than three words were coded from the middle twenty minutes of session 2 and session 6 of therapy sessions on the EXP Scale.
4. Testing the coding system	95 practice transcripts were coded.
5. Checking the reliability	Intraclass correlation was coded for the 95 practice transcripts.
6. Coding the data	80 sessions of data were coded. To reduce therapeutic drift, researchers coded in batches of 10 recordings. On the tenth recording, one of the researchers would code the tenth recording of the other researcher for calibration purposes.
7. Inferential statistical analyses	2 X 2 ANOVA's and follow-up post-hoc analyses were conducted.
8. Reporting the results	Results were interpreted and written up.

Appendix H – Demographic Analyses

Age

Group Statistics

	Treatments	N	Mean	Std. Deviation	Std. Error Mean
Age	CBT	20	39.35	11.089	2.480
	PCET	20	43.10	13.151	2.941

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p			Lower	Upper
Age	Equal variances assumed	1.010	.321	-.975	38	.168	.336	-3.750	3.847	-11.537	4.037
	Equal variances not assumed			-.975	36.947	.168	.336	-3.750	3.847	-11.544	4.044

Gender

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.404[a]	1	.525		
Continuity Correction[b]	.101	1	.751		
Likelihood Ratio	.405	1	.525		
Fisher's Exact Test				.751	.376
Linear-by-Linear Association	.394	1	.530		
N of Valid Cases	40				

Ethnicity

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	4.000[a]	4	.406
Likelihood Ratio	5.545	4	.236
Linear-by-Linear Association	.567	1	.451
N of Valid Cases	40		

Index of Multiple Deprivation

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	3.978[a]	9	.913
Likelihood Ratio	4.444	9	.880
Linear-by-Linear Association	.159	1	.690
N of Valid Cases	40		

Appendix I – Analyses of Statistics Across Each Therapeutic Modality

Baseline Severity PHQ-9

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	13.200[a]	14	.511
Likelihood Ratio	17.489	14	.231
Linear-by-Linear Association	1.285	1	.257
N of Valid Cases	40		

Baseline Severity GAD-7

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	14.810[a]	12	.252
Likelihood Ratio	18.671	12	.097
Linear-by-Linear Association	1.977	1	.160
N of Valid Cases	40		

Mean Number of Therapy Sessions

T-Test

Group Statistics

	Treatments	N	Mean	Std. Deviation	Std. Error Mean
Total_Contacts	CBT	20	16.7000	4.36614	.97630
	PCET	20	14.4000	4.70610	1.05232

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Total_Contacts	Equal variances assumed	.059	.809	1.602	38	.059	.117	2.30000	1.43545	-.60592	5.20592
	Equal variances not assumed			1.602	37.788	.059	.117	2.30000	1.43545	-.60646	5.20646

Appendix J - Intraclass Correlation

Reliability

Scale: ALL VARIABLES

Case Processing Summary

		N	%
Cases	Valid	95	100.0
	Excluded[a]	0	.0
	Total	95	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.937	2

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.883[a]	.829	.920	15.905	94	94	<.001
Average Measures	.938[c]	.906	.959	15.905	94	94	<.001

Two-way mixed effects model where people effects are random and measures effects are fixed.

a. The estimator is the same, whether the interaction effect is present or not.

b. Type A intraclass correlation coefficients using an absolute agreement definition.

c. This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Reliability

Scale: ALL VARIABLES

Case Processing Summary

		N	%
Cases	Valid	95	100.0
	Excluded[a]	0	.0
	Total	95	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.930	2

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.870[a]	.812	.912	14.300	94	94	<.001
Average Measures	.931[c]	.896	.954	14.300	94	94	<.001

Two-way mixed effects model where people effects are random and measures effects are fixed.

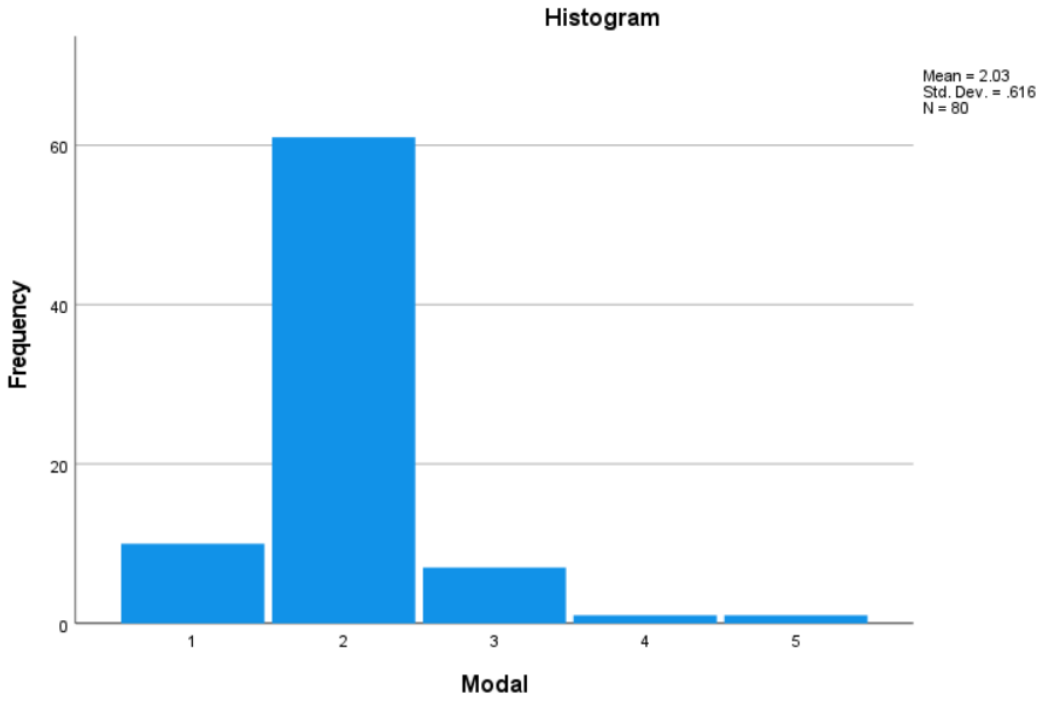
a. The estimator is the same, whether the interaction effect is present or not.

b. Type A intraclass correlation coefficients using an absolute agreement definition.

c. This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Appendix K – Modal ANOVA

Modal



Between-Subjects Factors

		Value Label	N
Treatments	1.00	CBT	20
	2.00	PCET	20

Descriptive Statistics

	Treatments	Mean	Std. Deviation	N
Session_2	CBT	1.9500	.51042	20
	PCET	1.9500	.22361	20
	Total	1.9500	.38895	40
Session_6	CBT	1.8500	.87509	20
	PCET	2.4500	.68633	20
	Total	2.1500	.83359	40

Box's Test of Equality of Covariance Matrices[a]

Box's M	12.989
F	4.083
df1	3
df2	259920.000
Sig.	.007

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Sessions	Pillai's Trace	.056	2.269[b]	1.000	38.000	.140
	Wilks' Lambda	.944	2.269[b]	1.000	38.000	.140
	Hotelling's Trace	.060	2.269[b]	1.000	38.000	.140
	Roy's Largest Root	.060	2.269[b]	1.000	38.000	.140
Sessions * Treatments	Pillai's Trace	.118	5.104[b]	1.000	38.000	.030
	Wilks' Lambda	.882	5.104[b]	1.000	38.000	.030
	Hotelling's Trace	.134	5.104[b]	1.000	38.000	.030
	Roy's Largest Root	.134	5.104[b]	1.000	38.000	.030

a. Design: Intercept + Treatments
Within Subjects Design: Sessions

b. Exact statistic

Mauchly's Test of Sphericity[a]

Measure: Experiencing

					Epsilon[b]		
Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Sessions	1.000	.000	0	.	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + Treatments
Within Subjects Design: Sessions

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Tests of Within-Subjects Effects

Measure: Experiencing

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Sessions	Sphericity Assumed	.800	1	.800	2.269	.140
	Greenhouse-Geisser	.800	1.000	.800	2.269	.140
	Huynh-Feldt	.800	1.000	.800	2.269	.140
	Lower-bound	.800	1.000	.800	2.269	.140
Sessions * Treatments	Sphericity Assumed	1.800	1	1.800	5.104	.030
	Greenhouse-Geisser	1.800	1.000	1.800	5.104	.030
	Huynh-Feldt	1.800	1.000	1.800	5.104	.030
	Lower-bound	1.800	1.000	1.800	5.104	.030
Error(Sessions)	Sphericity Assumed	13.400	38	.353		
	Greenhouse-Geisser	13.400	38.000	.353		
	Huynh-Feldt	13.400	38.000	.353		
	Lower-bound	13.400	38.000	.353		

Tests of Within-Subjects Contrasts

Measure: Experiencing

Source	Sessions	Type III Sum of Squares	df	Mean Square	F	Sig.
Sessions	Linear	.800	1	.800	2.269	.140
Sessions * Treatments	Linear	1.800	1	1.800	5.104	.030
Error(Sessions)	Linear	13.400	38	.353		

Levene's Test of Equality of Error Variances[a]

		Levene Statistic	df1	df2	Sig.
Session_2	Based on Mean	3.350	1	38	.075
	Based on Median	3.234	1	38	.080
	Based on Median and with adjusted df	3.234	1	28.046	.083
	Based on trimmed mean	5.125	1	38	.029
Session_6	Based on Mean	.187	1	38	.668
	Based on Median	.000	1	38	1.000
	Based on Median and with adjusted df	.000	1	37.620	1.000
	Based on trimmed mean	.001	1	38	.974

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Treatments
Within Subjects Design: Sessions

Tests of Between-Subjects Effects

Measure: Experiencing

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	336.200	1	336.200	798.475	<.001
Treatments	1.800	1	1.800	4.275	.046
Error	16.000	38	.421		

Appendix L – Modal EXP T-Tests

T-Test

Group Statistics

	Therapy_time	N	Mean	Std. Deviation	Std. Error Mean
EXP_SCORE	early	20	1.9500	.22361	.05000
	working	20	2.4500	.68633	.15347

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
EXP_SCORE	Equal variances assumed	31.735	<.001	-3.098	38	.002	.004	-.50000	.16141	-.82675	-.17325
	Equal variances not assumed			-3.098	22.989	.003	.005	-.50000	.16141	-.83391	-.16609

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
EXP_SCORE	Cohen's d	.51042	-.980	-1.631	-.316
	Hedges' correction	.52078	-.960	-1.599	-.310
	Glass's delta	.68633	-.729	-1.381	-.059

T-Test

Group Statistics

	THERAPY_TIME	N	Mean	Std. Deviation	Std. Error Mean
EXP_SCORE	EARLY	20	1.9500	.51042	.11413
	WORKING	20	1.8500	.87509	.19568

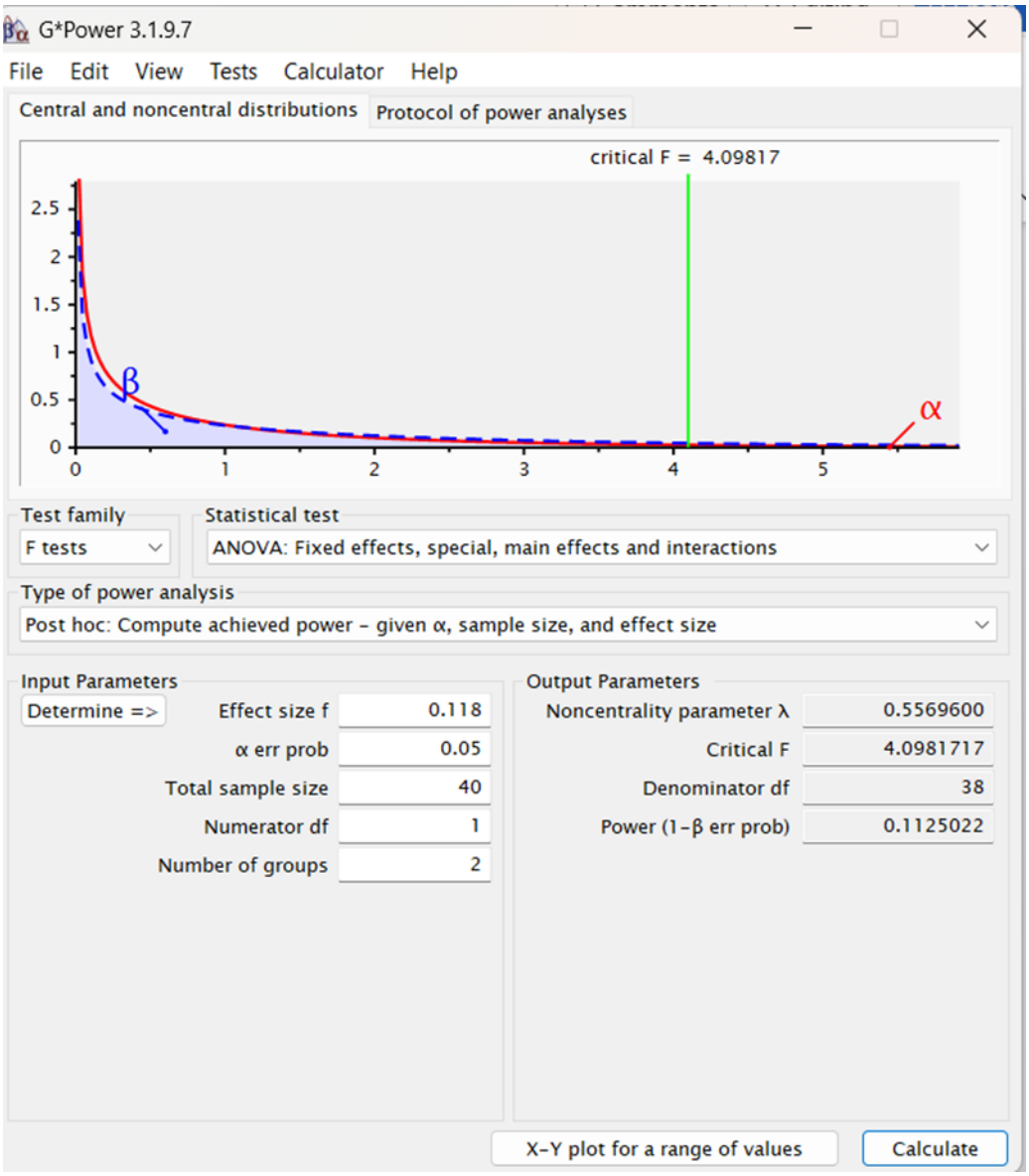
Independent Samples Test

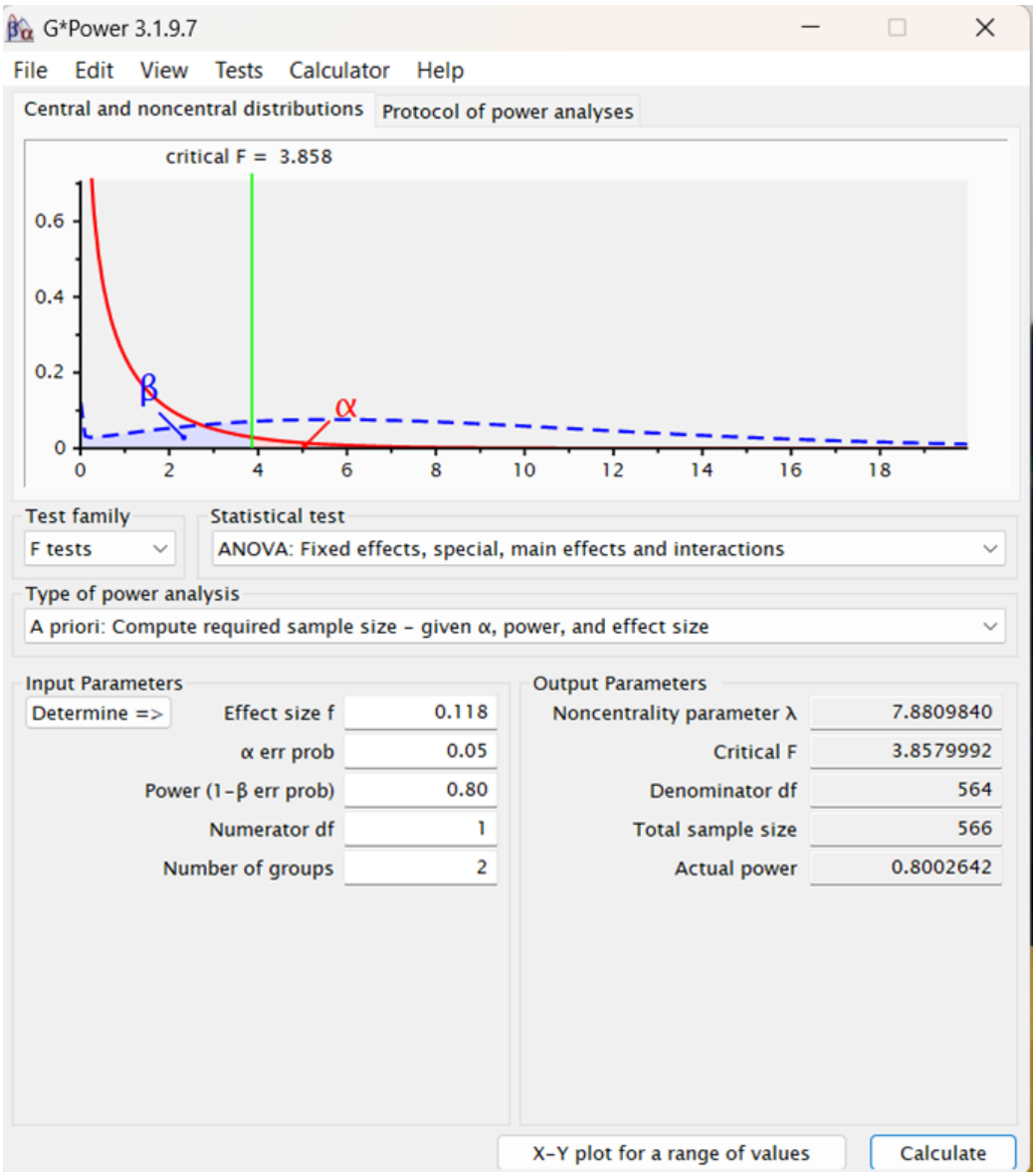
		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
EXP_SCORE	Equal variances assumed	1.518	.226	.441	38	.331	.661	.10000	.22653	-.35859	.55859
	Equal variances not assumed			.441	30.587	.331	.662	.10000	.22653	-.36226	.56226

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
EXP_SCORE	Cohen's d	.71635	.140	-.482	.759
	Hedges' correction	.73089	.137	-.472	.744
	Glass's delta	.87509	.114	-.508	.734

Appendix M – Power Analyses for Modal Data


G*Power 3.1.9.7
File Edit View Tests Calculator Help
Central and noncentral distributions
Protocol of power analyses
critical F = 4.09817
β
α
Test family
F tests
Statistical test
ANOVA: Fixed effects, special, main effects and interactions
Type of power analysis
Post hoc: Compute achieved power – given α, sample size, and effect size
Input Parameters
Determine =>
Effect size f 0.118
α err prob 0.05
Total sample size 40
Numerator df 1
Number of groups 2
Output Parameters
Noncentrality parameter λ 0.5569600
Critical F 4.0981717
Denominator df 38
Power (1-β err prob) 0.1125022
X-Y plot for a range of values
Calculate


G*Power 3.1.9.7
File Edit View Tests Calculator Help
Central and noncentral distributions
Protocol of power analyses
critical F = 3.858
0.6
0.4
0.2
0
0 2 4 6 8 10 12 14 16 18
β
α
Test family
F tests
Statistical test
ANOVA: Fixed effects, special, main effects and interactions
Type of power analysis
A priori: Compute required sample size – given α, power, and effect size
Input Parameters
Determine =>
Effect size f 0.118
α err prob 0.05
Power (1-β err prob) 0.80
Numerator df 1
Number of groups 2
Output Parameters
Noncentrality parameter λ 7.8809840
Critical F 3.8579992
Denominator df 564
Total sample size 566
Actual power 0.8002642
X-Y plot for a range of values
Calculate

Appendix N - Chi Square/Fisher's Exact Test for Modal Scores

Session6Mlevel1 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session6Mlevel1	.00	Count	14	20	34
		Expected Count	17.0	17.0	34.0
		% within Session6Mlevel1	41.2%	58.8%	100.0%
		% within Treatment	70.0%	100.0%	85.0%
		% of Total	35.0%	50.0%	85.0%
	1.00	Count	6	0	6
		Expected Count	3.0	3.0	6.0
		% within Session6Mlevel1	100.0%	0.0%	100.0%
		% within Treatment	30.0%	0.0%	15.0%
		% of Total	15.0%	0.0%	15.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session6Mlevel1	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	7.059[a]	1	.008		
Continuity Correction[b]	4.902	1	.027		
Likelihood Ratio	9.382	1	.002		
Fisher's Exact Test				.020	.010
Linear-by-Linear Association	6.882	1	.009		

Session2Mlevel2 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session2Mlevel2	no	Count	5	1	6
		Expected Count	3.0	3.0	6.0
		% within Session2Mlevel2	83.3%	16.7%	100.0%
		% within Treatment	25.0%	5.0%	15.0%
		% of Total	12.5%	2.5%	15.0%
	yes	Count	15	19	34
		Expected Count	17.0	17.0	34.0
		% within Session2Mlevel2	44.1%	55.9%	100.0%
		% within Treatment	75.0%	95.0%	85.0%
		% of Total	37.5%	47.5%	85.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session2Mlevel2	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.137[a]	1	.077		
Continuity Correction[b]	1.765	1	.184		
Likelihood Ratio	3.383	1	.066		
Fisher's Exact Test				.182	.091
Linear-by-Linear Association	3.059	1	.080		

Session6Mlevel2 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session6Mlevel2	no	Count	7	7	14
		Expected Count	7.0	7.0	14.0
		% within Session6Mlevel2	50.0%	50.0%	100.0%
		% within Treatment	35.0%	35.0%	35.0%
		% of Total	17.5%	17.5%	35.0%
	yes	Count	13	13	26
		Expected Count	13.0	13.0	26.0
		% within Session6Mlevel2	50.0%	50.0%	100.0%
		% within Treatment	65.0%	65.0%	65.0%
		% of Total	32.5%	32.5%	65.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session6Mlevel2	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.000[a]	1	1.000		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.000	1	1.000		
Fisher's Exact Test				1.000	.629
Linear-by-Linear Association	.000	1	1.000		

Session2Mlevel3 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session2Mlevel3	.00	Count	18	20	38
		Expected Count	19.0	19.0	38.0
		% within Session2Mlevel3	47.4%	52.6%	100.0%
		% within Treatment	90.0%	100.0%	95.0%
		% of Total	45.0%	50.0%	95.0%
	1.00	Count	2	0	2
		Expected Count	1.0	1.0	2.0
		% within Session2Mlevel3	100.0%	0.0%	100.0%
		% within Treatment	10.0%	0.0%	5.0%
		% of Total	5.0%	0.0%	5.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session2Mlevel3	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.105[a]	1	.147		
Continuity Correction[b]	.526	1	.468		
Likelihood Ratio	2.878	1	.090		
Fisher's Exact Test				.487	.244
Linear-by-Linear Association	2.053	1	.152		
N of Valid Cases	40				

Session6Mlevel3 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session6Mlevel3	.00	Count	20	15	35
		Expected Count	17.5	17.5	35.0
		% within Session6Mlevel3	57.1%	42.9%	100.0%
		% within Treatment	100.0%	75.0%	87.5%
		% of Total	50.0%	37.5%	87.5%
	1.00	Count	0	5	5
		Expected Count	2.5	2.5	5.0
		% within Session6Mlevel3	0.0%	100.0%	100.0%
		% within Treatment	0.0%	25.0%	12.5%
		% of Total	0.0%	12.5%	12.5%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session6Mlevel3	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5.714[a]	1	.017		
Continuity Correction[b]	3.657	1	.056		
Likelihood Ratio	7.648	1	.006		
Fisher's Exact Test				.047	.024
Linear-by-Linear Association	5.571	1	.018		

Session2Mlevel4 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session2Mlevel4	.00	Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session2Mlevel4	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session2Mlevel4	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value
Pearson Chi-Square	.[a]
N of Valid Cases	40

a. No statistics are computed because Session2Mlevel4 is a constant.

Session6Mlevel4 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session6Mlevel4	.00	Count	20	18	38
		Expected Count	19.0	19.0	38.0
		% within Session6Mlevel4	52.6%	47.4%	100.0%
		% within Treatment	100.0%	90.0%	95.0%
		% of Total	50.0%	45.0%	95.0%
	1.00	Count	0	2	2
		Expected Count	1.0	1.0	2.0
		% within Session6Mlevel4	0.0%	100.0%	100.0%
		% within Treatment	0.0%	10.0%	5.0%
		% of Total	0.0%	5.0%	5.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session6Mlevel4	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.105[a]	1	.147		
Continuity Correction[b]	.526	1	.468		
Likelihood Ratio	2.878	1	.090		
Fisher's Exact Test				.487	.244
Linear-by-Linear Association	2.053	1	.152		

Session6Mlevel5 * Treatment Crosstabulation

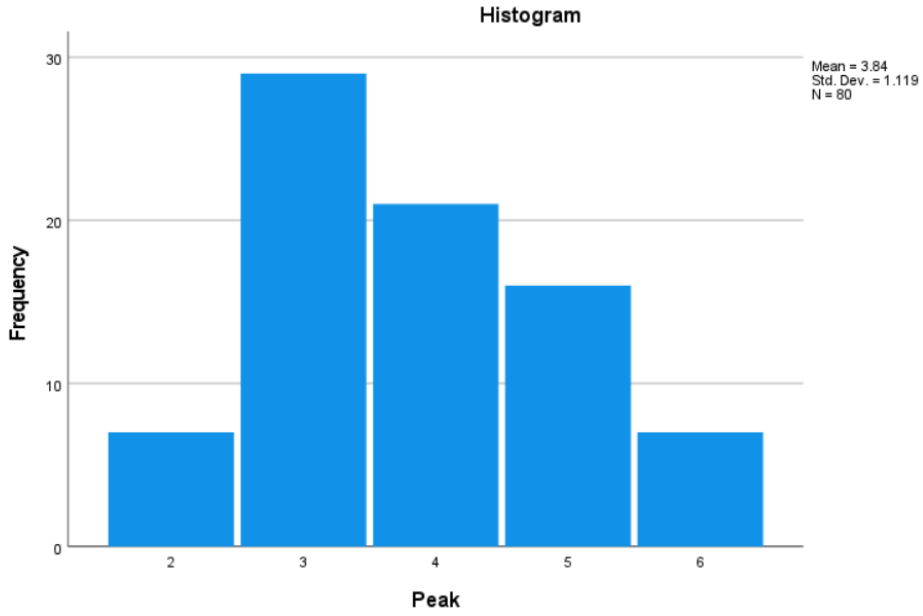
			Treatment		
			cbt	pcet	Total
Session6Mlevel5	.00	Count	19	20	39
		Expected Count	19.5	19.5	39.0
		% within Session6Mlevel5	48.7%	51.3%	100.0%
		% within Treatment	95.0%	100.0%	97.5%
		% of Total	47.5%	50.0%	97.5%
	1.00	Count	1	0	1
		Expected Count	.5	.5	1.0
		% within Session6Mlevel5	100.0%	0.0%	100.0%
		% within Treatment	5.0%	0.0%	2.5%
		% of Total	2.5%	0.0%	2.5%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session6Mlevel5	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.026[a]	1	.311		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	1.412	1	.235		
Fisher's Exact Test				1.000	.500
Linear-by-Linear Association	1.000	1	.317		

Appendix O – Peak EXP ANOVA

Peak



Between-Subjects Factors

		Value Label	N
Treatments	1.00	CBT	20
	2.00	PCET	20

Descriptive Statistics

	Treatments	Mean	Std. Deviation	N
Session_2_P	CBT	3.2500	.85070	20
	PCET	4.0500	.94451	20
	Total	3.6500	.97534	40
Session_6_P	CBT	3.8000	1.39925	20
	PCET	4.2000	1.10501	20
	Total	4.0000	1.26085	40

Box's Test of Equality of Covariance Matrices[a]

Box's M	1.877
F	.590
df1	3
df2	259920.000
Sig.	.622

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
sessions	Pillai's Trace	.047	1.871[b]	1.000	38.000	.179	.047
	Wilks' Lambda	.953	1.871[b]	1.000	38.000	.179	.047
	Hotelling's Trace	.049	1.871[b]	1.000	38.000	.179	.047
	Roy's Largest Root	.049	1.871[b]	1.000	38.000	.179	.047
sessions * Treatments	Pillai's Trace	.016	.611[b]	1.000	38.000	.439	.016
	Wilks' Lambda	.984	.611[b]	1.000	38.000	.439	.016
	Hotelling's Trace	.016	.611[b]	1.000	38.000	.439	.016
	Roy's Largest Root	.016	.611[b]	1.000	38.000	.439	.016

a. Design: Intercept + Treatments
Within Subjects Design: sessions

b. Exact statistic

Mauchly's Test of Sphericity[a]

Measure: exp

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
sessions	1.000	.000	0	.	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + Treatments
Within Subjects Design: sessions

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Tests of Within-Subjects Effects

Measure: exp

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
sessions	Sphericity Assumed	2.450	1	2.450	1.871	.179	.047
	Greenhouse-Geisser	2.450	1.000	2.450	1.871	.179	.047
	Huynh-Feldt	2.450	1.000	2.450	1.871	.179	.047
	Lower-bound	2.450	1.000	2.450	1.871	.179	.047
sessions * Treatments	Sphericity Assumed	.800	1	.800	.611	.439	.016
	Greenhouse-Geisser	.800	1.000	.800	.611	.439	.016
	Huynh-Feldt	.800	1.000	.800	.611	.439	.016
	Lower-bound	.800	1.000	.800	.611	.439	.016
Error(sessions)	Sphericity Assumed	49.750	38	1.309			
	Greenhouse-Geisser	49.750	38.000	1.309			
	Huynh-Feldt	49.750	38.000	1.309			
	Lower-bound	49.750	38.000	1.309			

Tests of Within-Subjects Contrasts

Measure: exp

Source	sessions	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
sessions	Linear	2.450	1	2.450	1.871	.179	.047
sessions * Treatments	Linear	.800	1	.800	.611	.439	.016
Error(sessions)	Linear	49.750	38	1.309			

Levene's Test of Equality of Error Variances[a]

		Levene Statistic	df1	df2	Sig.
Session_2_P	Based on Mean	.477	1	38	.494
	Based on Median	1.034	1	38	.316
	Based on Median and with adjusted df	1.034	1	36.281	.316
	Based on trimmed mean	.413	1	38	.524
Session_6_P	Based on Mean	2.382	1	38	.131
	Based on Median	.924	1	38	.342
	Based on Median and with adjusted df	.924	1	32.880	.343
	Based on trimmed mean	2.293	1	38	.138

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Treatments
Within Subjects Design: sessions

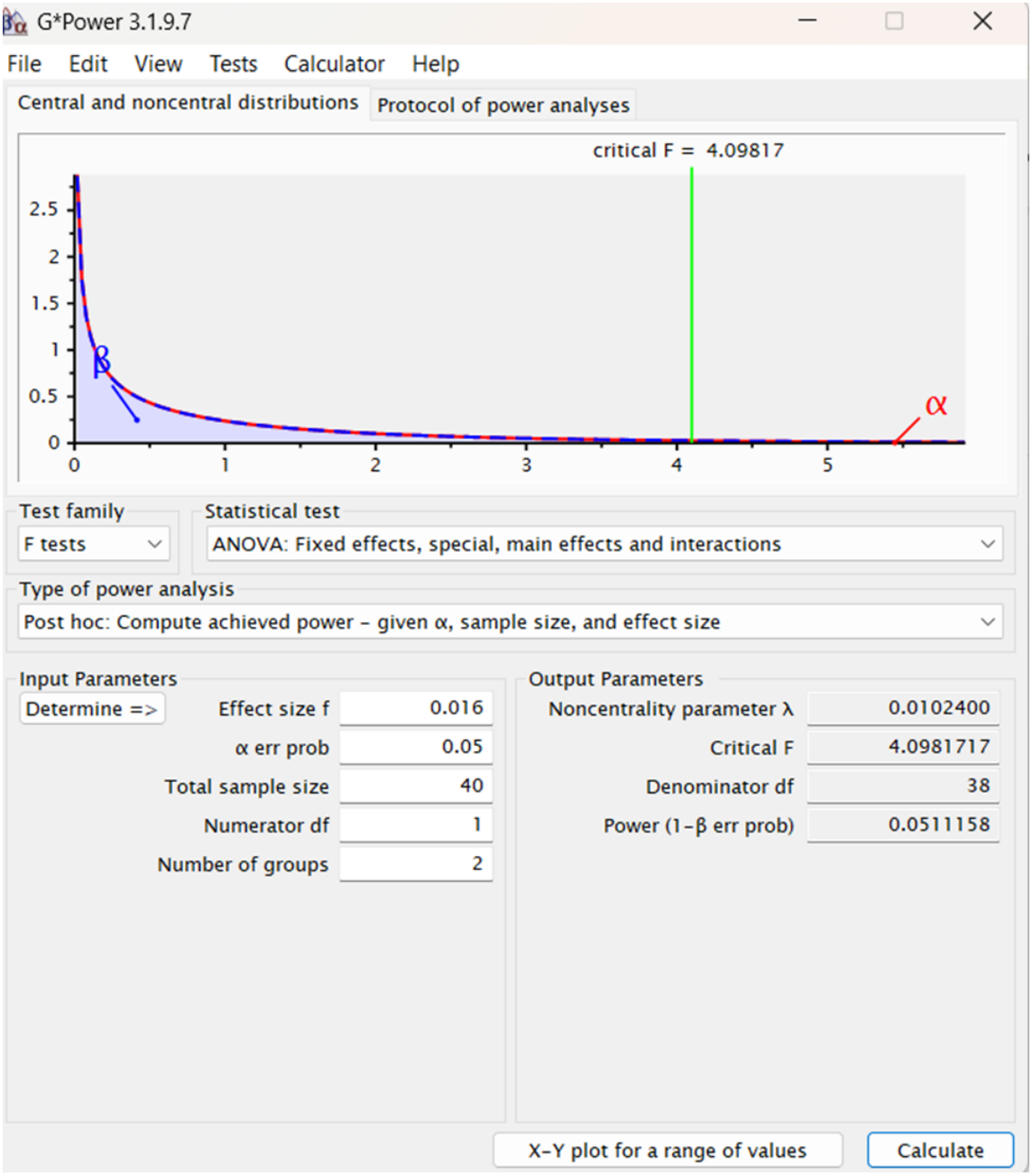
Tests of Between-Subjects Effects

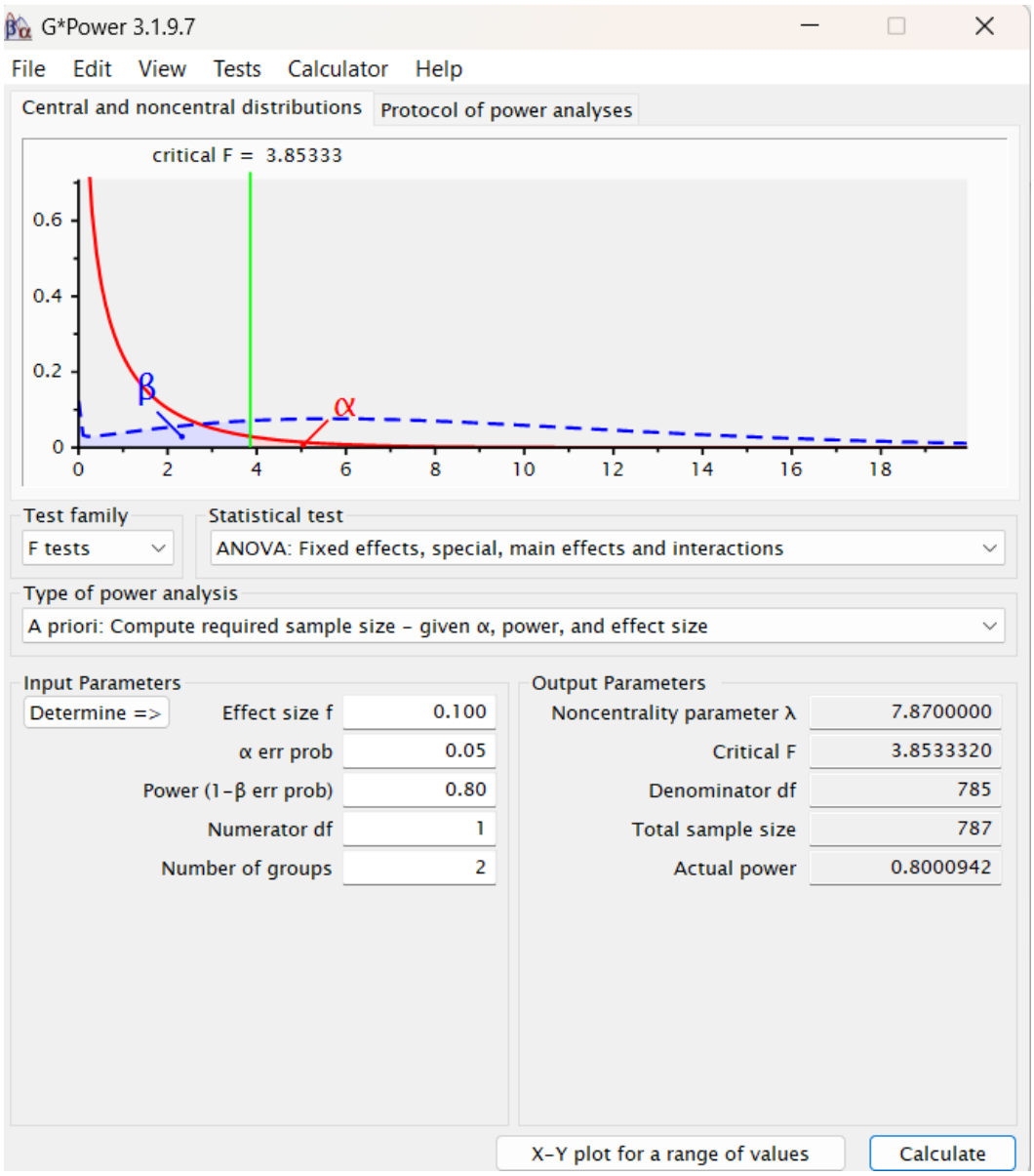
Measure: exp

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	1170.450	1	1170.450	1075.625	<.001	.966
Treatments	7.200	1	7.200	6.617	.014	.148
Error	41.350	38	1.088			

Appendix P – Power Analyses for Peak Data




G*Power 3.1.9.7
File Edit View Tests Calculator Help
Central and noncentral distributions
Protocol of power analyses
critical F = 3.85333
0.6
0.4
0.2
0
0 2 4 6 8 10 12 14 16 18
β
α
Test family
F tests
Statistical test
ANOVA: Fixed effects, special, main effects and interactions
Type of power analysis
A priori: Compute required sample size – given α, power, and effect size
Input Parameters
Determine =>
Effect size f 0.100
α err prob 0.05
Power (1–β err prob) 0.80
Numerator df 1
Number of groups 2
Output Parameters
Noncentrality parameter λ 7.8700000
Critical F 3.8533320
Denominator df 785
Total sample size 787
Actual power 0.8000942
X-Y plot for a range of values
Calculate

Appendix Q - Chi Square/Fisher's Exact Test for Peak Scores

Session2Plevel2 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session2Plevel2	no	Count	17	20	37
		Expected Count	18.5	18.5	37.0
		% within Session2Plevel2	45.9%	54.1%	100.0%
		% within Treatment	85.0%	100.0%	92.5%
		% of Total	42.5%	50.0%	92.5%
	yes	Count	3	0	3
		Expected Count	1.5	1.5	3.0
		% within Session2Plevel2	100.0%	0.0%	100.0%
		% within Treatment	15.0%	0.0%	7.5%
		% of Total	7.5%	0.0%	7.5%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session2Plevel2	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.243[a]	1	.072		
Continuity Correction[b]	1.441	1	.230		
Likelihood Ratio	4.402	1	.036		
Fisher's Exact Test				.231	.115
Linear-by-Linear Association	3.162	1	.075		

Session6Plevel2 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session6Plevel2	no	Count	17	18	35
		Expected Count	17.5	17.5	35.0
		% within Session6Plevel2	48.6%	51.4%	100.0%
		% within Treatment	85.0%	90.0%	87.5%
		% of Total	42.5%	45.0%	87.5%
	yes	Count	3	2	5
		Expected Count	2.5	2.5	5.0
		% within Session6Plevel2	60.0%	40.0%	100.0%
		% within Treatment	15.0%	10.0%	12.5%
		% of Total	7.5%	5.0%	12.5%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session6Plevel2	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.229[a]	1	.633		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.230	1	.632		
Fisher's Exact Test				1.000	.500
Linear-by-Linear Association	.223	1	.637		

Session2Plevel3 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session2Plevel3	.00	Count	9	13	22
		Expected Count	11.0	11.0	22.0
		% within Session2Plevel3	40.9%	59.1%	100.0%
		% within Treatment	45.0%	65.0%	55.0%
		% of Total	22.5%	32.5%	55.0%
	1.00	Count	11	7	18
		Expected Count	9.0	9.0	18.0
		% within Session2Plevel3	61.1%	38.9%	100.0%
		% within Treatment	55.0%	35.0%	45.0%
		% of Total	27.5%	17.5%	45.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session2Plevel3	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.616[a]	1	.204		
Continuity Correction[b]	.909	1	.340		
Likelihood Ratio	1.628	1	.202		
Fisher's Exact Test				.341	.170
Linear-by-Linear Association	1.576	1	.209		

Session6Plevel3 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session6Plevel3	.00	Count	12	18	30
		Expected Count	15.0	15.0	30.0
		% within Session6Plevel3	40.0%	60.0%	100.0%
		% within Treatment	60.0%	90.0%	75.0%
		% of Total	30.0%	45.0%	75.0%
	1.00	Count	8	2	10
		Expected Count	5.0	5.0	10.0
		% within Session6Plevel3	80.0%	20.0%	100.0%
		% within Treatment	40.0%	10.0%	25.0%
		% of Total	20.0%	5.0%	25.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session6Plevel3	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.800[a]	1	.028		
Continuity Correction[b]	3.333	1	.068		
Likelihood Ratio	5.063	1	.024		
Fisher's Exact Test				.065	.032
Linear-by-Linear Association	4.680	1	.031		

Session2Plevel4 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session2Plevel4	.00	Count	16	14	30
		Expected Count	15.0	15.0	30.0
		% within Session2Plevel4	53.3%	46.7%	100.0%
		% within Treatment	80.0%	70.0%	75.0%
		% of Total	40.0%	35.0%	75.0%
	1.00	Count	4	6	10
		Expected Count	5.0	5.0	10.0
		% within Session2Plevel4	40.0%	60.0%	100.0%
		% within Treatment	20.0%	30.0%	25.0%
		% of Total	10.0%	15.0%	25.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session2Plevel4	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.533[a]	1	.465		
Continuity Correction[b]	.133	1	.715		
Likelihood Ratio	.536	1	.464		
Fisher's Exact Test				.716	.358
Linear-by-Linear Association	.520	1	.471		

Session6Plevel4 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session6Plevel4	.00	Count	17	12	29
		Expected Count	14.5	14.5	29.0
		% within Session6Plevel4	58.6%	41.4%	100.0%
		% within Treatment	85.0%	60.0%	72.5%
		% of Total	42.5%	30.0%	72.5%
	1.00	Count	3	8	11
		Expected Count	5.5	5.5	11.0
		% within Session6Plevel4	27.3%	72.7%	100.0%
		% within Treatment	15.0%	40.0%	27.5%
		% of Total	7.5%	20.0%	27.5%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session6Plevel4	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.135[a]	1	.077		
Continuity Correction[b]	2.006	1	.157		
Likelihood Ratio	3.225	1	.073		
Fisher's Exact Test				.155	.078
Linear-by-Linear Association	3.056	1	.080		

SessionnPlevel5 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
SessionnPlevel5	.00	Count	18	14	32
		Expected Count	16.0	16.0	32.0
		% within SessionnPlevel5	56.3%	43.8%	100.0%
		% within Treatment	90.0%	70.0%	80.0%
		% of Total	45.0%	35.0%	80.0%
	1.00	Count	2	6	8
		Expected Count	4.0	4.0	8.0
		% within SessionnPlevel5	25.0%	75.0%	100.0%
		% within Treatment	10.0%	30.0%	20.0%
		% of Total	5.0%	15.0%	20.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within SessionnPlevel5	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.500[a]	1	.114		
Continuity Correction[b]	1.406	1	.236		
Likelihood Ratio	2.594	1	.107		
Fisher's Exact Test				.235	.118
Linear-by-Linear Association	2.437	1	.118		

Session6Plevel5 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session6Plevel5	.00	Count	18	14	32
		Expected Count	16.0	16.0	32.0
		% within Session6Plevel5	56.3%	43.8%	100.0%
		% within Treatment	90.0%	70.0%	80.0%
		% of Total	45.0%	35.0%	80.0%
	1.00	Count	2	6	8
		Expected Count	4.0	4.0	8.0
		% within Session6Plevel5	25.0%	75.0%	100.0%
		% within Treatment	10.0%	30.0%	20.0%
		% of Total	5.0%	15.0%	20.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session6Plevel5	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.500[a]	1	.114		
Continuity Correction[b]	1.406	1	.236		
Likelihood Ratio	2.594	1	.107		
Fisher's Exact Test				.235	.118
Linear-by-Linear Association	2.437	1	.118		

Session2Plevel6 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session2Plevel6	.00	Count	20	19	39
		Expected Count	19.5	19.5	39.0
		% within Session2Plevel6	51.3%	48.7%	100.0%
		% within Treatment	100.0%	95.0%	97.5%
		% of Total	50.0%	47.5%	97.5%
	1.00	Count	0	1	1
		Expected Count	.5	.5	1.0
		% within Session2Plevel6	0.0%	100.0%	100.0%
		% within Treatment	0.0%	5.0%	2.5%
		% of Total	0.0%	2.5%	2.5%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session2Plevel6	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.026[a]	1	.311		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	1.412	1	.235		
Fisher's Exact Test				1.000	.500
Linear-by-Linear Association	1.000	1	.317		

Session6Plevel6 * Treatment Crosstabulation

			Treatment		
			cbt	pcet	Total
Session6Plevel6	.00	Count	16	18	34
		Expected Count	17.0	17.0	34.0
		% within Session6Plevel6	47.1%	52.9%	100.0%
		% within Treatment	80.0%	90.0%	85.0%
		% of Total	40.0%	45.0%	85.0%
	1.00	Count	4	2	6
		Expected Count	3.0	3.0	6.0
		% within Session6Plevel6	66.7%	33.3%	100.0%
		% within Treatment	20.0%	10.0%	15.0%
		% of Total	10.0%	5.0%	15.0%
Total		Count	20	20	40
		Expected Count	20.0	20.0	40.0
		% within Session6Plevel6	50.0%	50.0%	100.0%
		% within Treatment	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.784[a]	1	.376		
Continuity Correction[b]	.196	1	.658		
Likelihood Ratio	.797	1	.372		
Fisher's Exact Test				.661	.331
Linear-by-Linear Association	.765	1	.382		