Key components of effective Psychological Wellbeing Practitioners:

Evidence from routine practice

Helen Green

Thesis submitted to the University of Sheffield

for the degree of

Doctor of Clinical Psychology

July 2011

Declaration

This work has not been submitted to any other institution or for any other qualification

Thesis Structure

The literature review has been prepared in accordance with the guidance for contributors provided by the *Journal of Consulting and Clinical Psychology.* The empirical research report has been prepared in accordance with the guidance for contributors to *Behaviour, Research and Therapy.* Copies of the University journal approval letter and guidelines for authors can be seen in Appendices I, II, and III respectively.

Word Count

Literature Review:	7,998
Research Report:	11,974
Appendices:	6,166
Total word count without references and appendices:	19,972
Total word count with references and appendices:	28,665

Overall Abstract

The purpose of the literature review was to examine the empirical evidence base for therapist effects. Specifically, the review focussed on the methodological issues associated with studying this phenomenon. Seventeen papers were identified that focussed on comparing quantitative treatment outcomes of psychotherapists, nine of which utilised the recommended analytic strategy for examining therapist effects, multi-level modelling. Fifteen of the papers found positive evidence for the existence of therapist effects and the methodological strengths and weaknesses of the papers are discussed.

The purpose of the research report was to examine whether therapist effects existed in a sample of Psychological Wellbeing Practitioners (PWPs) operating at step two of the Improving Access to Psychological Therapies programme. A mixed methods approach was employed, utilising quantitative analysis of electronic client data and PWP measures, and qualitative analysis of interview data from PWPs and their supervisors. Outcome data was analysed using multi-level modelling, which resulted in the finding that almost 9% of the variance in outcome scores was attributable to PWPs. Rankings of PWPs were created from the MLM results and the most effective PWPs were found to have higher rates of resilience than less effective PWPs. Qualitative analysis showed that the more effective group of PWPs described approaching their work in a confident and organised manner, and appeared to be at a more advanced practitioner developmental level than PWPs in the less effective group.

Acknowledgements

This project would not have been possible without the collaboration and involvement of a number of people. I am grateful for the help of everyone involved in recruitment, including the lecturers and administrators at the universities who helped facilitate initial contact with the PWPs; the staff in the governance offices who helped with the formal access procedures and all the managers and administrators in the individual services from which PWPs were recruited.

I would like to thank all the PWPs and supervisors who gave their time to participate in this research, their enthusiasm for the PWP role and involvement in the project helped to make the process enjoyable. My thanks also go to the data managers from all the IAPT services who provided the electronic outcome data and I am extremely appreciative of their help.

I am eternally grateful for the guidance, advice and support of my supervisors, Prof. Michael Barkham and Dr Steve Kellett. Their encouragement and help has kept me going throughout! I would also like to thank Dave Saxon for his help and advice throughout this project.

The continued support from my family, particularly my parents and partner, has been vital in helping me through not only the development of this research, but the whole three years. I would particularly like to thank Paul, as he has been with me throughout the doctoral process and has been the best source of support I could have hoped for.

Contents

Chapter 2: Research Report

Chapter 1: Literature Review

The impact of individual therapists on psychotherapy outcome: A review and synthesis

Literature Review Abstract

Objective: The objective of this review was to examine the published literature on therapist effects. Specific focus was given to the methodologies used by such studies, in order to determine the appropriateness of methods and analytic strategies utilised in empirical studies of therapist effects.

Method: A search of the literature was carried out using electronic databases and visual search strategies. Following the application of inclusion/exclusion criterion, a grand total of 17 studies of therapist effects were included in this review. Studies were reviewed based on their design of Randomised Controlled Trial; Non-Randomised Trial or Routine Outcome Study and were further subdivided based on their use of either single-level or multi-level analysis

Results: Fifteen of the seventeen studies reviewed found evidence of variability in therapist outcomes. The studies varied in their methods of assessing therapist effects. Methods included ranking therapist outcomes, creating and comparing more and less effective therapist groups, and examining the percentage of the variance in outcomes attributable to therapists. The latter method resulted in percentages ranging from 0 to 17% in the four studies that utilised this method.

Conclusions: The majority of studies support the existence of therapist effects. Multi-level modelling provides the most appropriate statistical analysis for examining therapist effects as it accounts for the “nested” nature of the data.

1. Introduction

The dominant approach that largely drives efforts to improve the effectiveness of psychological therapies focuses on the development of treatments (i.e., psychological interventions) for specific clinical presentations. This approach is a response to the need to ensure that effective treatments are available for the range of diagnostic presentations. Delivery of these treatments is supported by clinical guidance from the National Institute for Health and Clinical Excellence (NICE). Overall, these components support activity that focuses on *process* – that is, in the event of diagnosing a specific clinical presentation, the clinical guidelines will indicate an evidence-based intervention that is the treatment of choice. Accordingly, standards of service delivery are achieved by ensuring that specific clinical presentations are matched with the treatment of choice as determined by NICE guidance.

By contrast, current government documents prompt movement towards an *outcomes* approach or framework, whereby the key indicator is outcome for the patient rather than the process of delivering a specific intervention (e.g., *Payment by results*, Department of Health, 2011). This shift in focus necessitates considering factors that might reasonably contribute to or influence outcomes. Among these would be therapists (or practitioners), as it seems unlikely that all practitioners would, or could, be equally effective. This proposition has been reflected in two separate strands of research activity. The first strand focuses on the substantive question of whether some practitioners are more effective than others and, if so, why. The second strand captures the methodological issues and debates associated with how best to establish the presence of practitioner effects.

1.1 Effective therapists

The drive to test the efficacy of specific psychotherapeutic modalities as described above has resulted in a comparative lack of research focussing on individual therapists (Lambert & Okiishi, 1997). However, a large body of research literature attests to the broad equivalence of outcome across modalities and that therapeutic techniques are not strongly linked to client improvement (Lambert & Bergin, 1994; Lambert & Okiishi, 1997; Smith, Glass, & Miller, 1980). Researchers, therefore, have attempted to examine other factors associated with client change. These include client and therapist characteristics, common factors across modalities, and critical external events – all of which have shown some degree of association with client improvement (Lambert & Bergin, 1994).

As these investigations have proved inconclusive, some researchers have argued for a greater emphasis on researching therapists' contribution to outcome in order to further our understanding of crucial change factors (Lambert, 1989). This shift in research strategy places an emphasis on evaluating the performance of individual therapist outcomes rather than on therapist characteristics (e.g., age, gender) and process measures (Luborsky, McLellan, Diguer, Woody, & Seligman, 1997). Lambert (1989) reviewed studies of therapists' contribution to outcome and concluded: "there is empirical support for the notion that the individual therapist can have a substantial effect on process and outcome" (p.480). However, it was noted that the review contained a small number of articles and that methodological limitations may have produced error variance rather than therapist effects (Lambert, 1989). This has led critics to argue that the phenomenon of therapist effects may be due to

methodological issues rather than a true effect of therapist difference (Elkin et al., 2006). This leads to the second strand of work, namely method.

1.2 Methodological issues

Martindale (1978) first reported concerns about the design and analysis of psychotherapy outcome data that did not consider the therapist variable. It was argued that if findings of psychotherapy outcome studies were to be generalised to populations of clients *and* therapists, then therapists must be treated as a random, rather than fixed, factor in research studies. Intuitively it may seem more appropriate to treat therapists as a fixed factor, as these are factors specifically chosen by the research team and not sampled from a population. It would be very difficult to randomly select a sample of therapists for inclusion in a research trial, leading some researchers to argue that treating therapists as fixed effects is more appropriate (Siemer & Joormann, 2003). However, fixed effects assume equivalency of therapists and evidence suggests that there are differences between them (e.g., Lambert, 1989). It has therefore been argued that therapists should be treated as a random factor in such analyses to account for differences, even though this sampling is not strictly random (Crits-Christoph & Mintz, 1991).

The psychotherapy outcomes literature was reviewed by Martindale (1978) and again by Crits-Christoph and Mintz (1991) to examine the methodologies and statistical analyses used by such studies. They found that in the majority of studies, therapists were treated as fixed effects and only small numbers of therapists were included in the samples. It was concluded, therefore that most studies adopted inappropriate statistical analyses to accurately examine

therapist effects and were likely to contain a high proportion of type I errors. (Martindale, 1978 and Crits-Christoph & Mintz, 1991). Based on this evidence, Martindale (1978) suggested that therapists, as well as patients, must be considered when analysing treatment outcome studies.

Traditional models of statistical analysis used in therapist effects studies have also been criticised because they ignore the hierarchical nature of the data (Kim, Wampold & Bolt, 2006). Client and therapist populations are not independent – that is clients are nested within individual therapists who may vary in their levels of effectiveness. If groups of clients are treated by different therapists, it is necessary to consider whether these therapists had an effect on their group of clients' outcome.

The recent development of complex analyses strategies have provided researchers with an alternative to traditional analyses that ignore this nested data, namely multi-level modelling (MLM; also called hierarchical linear modelling). This is a statistical technique that allows data to be analysed on two levels of client and therapist, thus taking into account the nested, multi-level nature of such data (Soldz, 2006). Although there is no formal power analysis for using MLM, a recommendation has been made for inclusion of a minimum of 30 therapists each treating 30 clients for a two level model, in order to ensure reliable results (Soldz, 2006).

1.3 The current review

In light of the two strands of literature relating to (a) the effective practitioner and (b) methodological issues, the purpose of the current review is to review and appraise the published therapist effects research since Lambert's (1989) paper in order to determine the current status of the empirical evidence for the contribution of therapist effects to treatment outcome. This review will focus on two main questions: (1) what research designs and methodologies have been used to study therapist effects? (2) when an appropriate methodology has been used, do therapist effects exist?

2. Method

2.1 Identification of Studies

A literature search was conducted using electronic databases accessed through OvidSP including PsycInfo (1806-July week 1 2011), and Ovid MEDLINE (1948 to July week 1 2011) and through Web of Knowledge including Web of Science (1899-2011); BIOSIS Previews (1969-2011) and MEDLINE (1950-2011). The last search was dated 16th July 2011. The initial search strategy used the term "therapist effects" as a "key word," "in topic" or "in title" to target articles specifically examining this phenomenon. The search yielded a total of 2,130 articles.

A wider search was then conducted to ensure inclusion of all relevant papers. Electronic search terms used were within "key word" or "in topic" and consisted of: (1) "therapist" OR "psychological therapist" OR "counsellor" OR "counselor"

OR “psychologist” OR “psychotherapist” and (2) “treatment outcomes” OR “outcome” OR “effectiveness”. This search strategy yielded 16,483 articles.

Both searches (i.e., therapist effects and the wider search) were combined using the selection “OR” and limited to articles in peer reviewed journals, written in English, published since 1989, and using adult samples. This resulted in 8,234 articles.

2.2 Selection of Studies

2.2.1 Inclusion Criterion

Studies selected for the review had to meet the following inclusion criterion:

a. published in a peer reviewed journal
b. published since 1989
c. written in English
d. adult participants
e. empirical study examining quantitative treatment outcomes
f. treatment provided by a psychological therapist.

2.2.2 Exclusion criterion

Studies were excluded from the review if they met the following exclusion criterion:

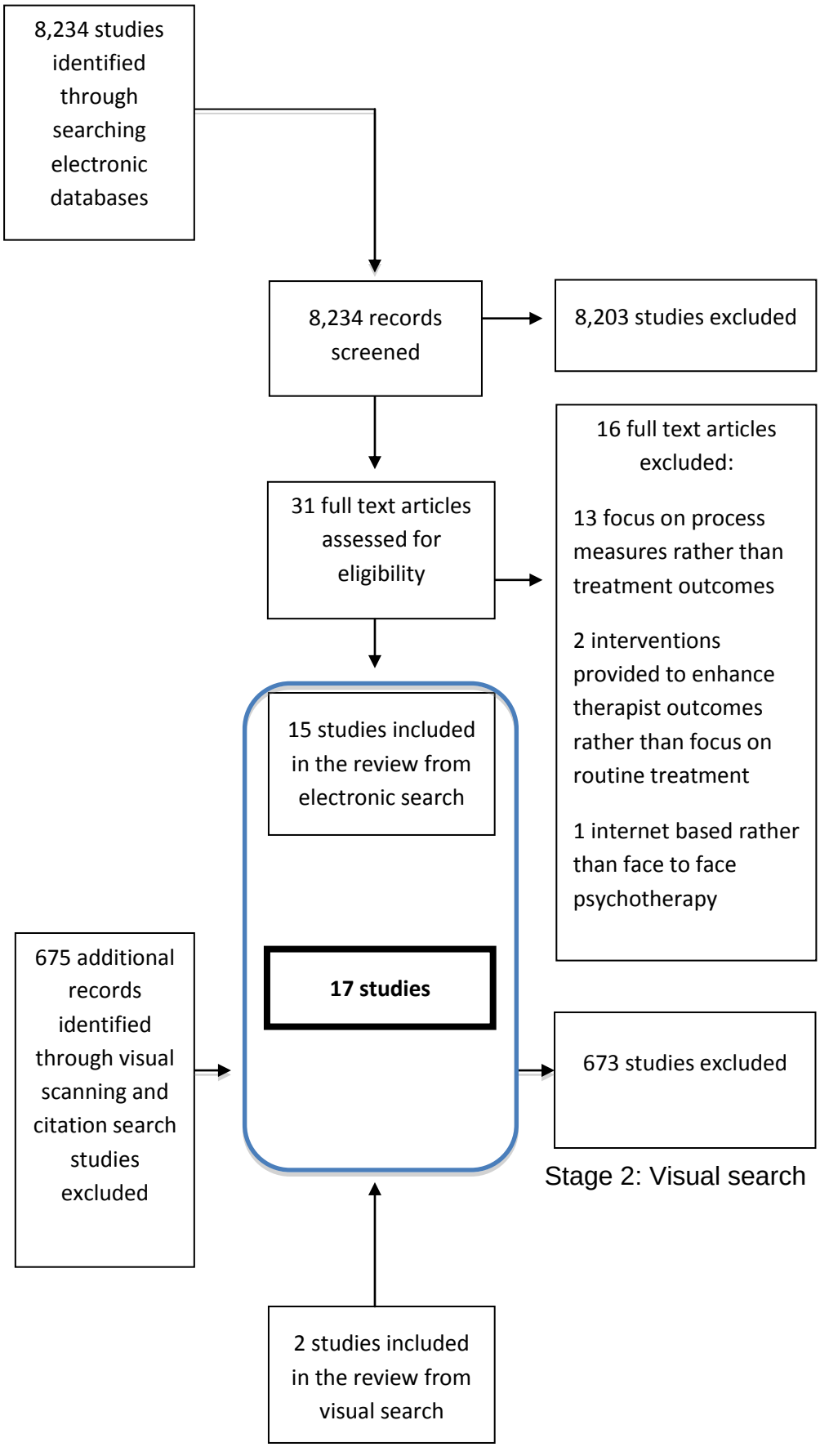
a. not published in a peer reviewed journal
b. published before 1989
c. not written in English

d. no primary focus on quantitative treatment outcomes (i.e., those with a focus on process variables such as working alliance).
e. psychotherapy not the primary treatment
f. child participants.

2.2.3 Process of Selection

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) provides guidance for reviewers on the optimum way of presenting information through the phases of a systematic review (Moher, Liberati, Tetzlaff, Altman & The PRISMA group, 2009). Figure 1 presents a PRISMA diagram of the process of study selection. The search comprised two stages: (a) searches of the electronic data bases, and (b) visual/hand searching. *Stage 1:* After initial identification of studies (k=8,234), abstracts were screened with reference to the inclusion/exclusion criterion. Those not meeting inclusion criterion (k=8,203) were excluded on the basis of: not a study of psychotherapy (i.e., medical interventions or physical therapies, k=3,753), a single intervention effectiveness study with no comparison of therapists (k=2,656); studies comparing more than one intervention with no focus on therapists (k=1,304), and studies focussing on children (k=490). This left a total of 31 papers. Full texts of these 31 papers were reviewed and 16 were excluded for the reasons given in the PRISMA diagram. *Stage 2:* Reference lists of the remaining 15 papers (k=675) were reviewed to ensure inclusion of any relevant articles not found in the search. This search produced an additional 2 papers that met the inclusion criterion and created a grand total of 17 papers included in the review.

Figure 1: PRISMA diagram of studies identified, included or excluded.



2.2.4 Quality Ratings

To assess the quality of the 17 identified papers, the Downs and Black (1998) quality checklist was used. The checklist facilitates reviewers in assessing the methodological and reporting quality of randomised and non-randomised studies. This is achieved by comparison of a Quality Index rating to a mean score of 14 for Randomised Controlled Trials (RCTs) and 11.7 for non-randomised studies found by Downs and Black (1998).

The checklist was adapted for use in the present review. Two items were removed (14 and 24) as they related to blinding participants to the intervention they were receiving, which was inappropriate for use in psychotherapy studies. Item 27 relating to power was scored on a 0/1 basis rather than 0-5 to reflect the recommendations of numbers of therapists needed to achieve reliable results in multi-level modelling (Soldz, 2006). A score of 1 was achieved if the sample included 30 therapists each treating 30 clients. This criteria was also adopted for studies not utilising MLM. With these changes, a new average score was calculated taking the original mean score as a percentage (45.16 for RCTs, 37.74 for non-randomised), which resulted in a score of 11.74 for randomised and 9.80 for non-randomised studies. The results of the quality ratings for the final 17 studies are presented in Table 1.

To determine the reliability of the quality ratings, an independent rater was employed to rate approximately 20 per cent (n=4) articles (Anderson, Ogles, Patterson, Lambert & Vermeersch, 2009; Blatt, Sanislow, Zuroff, & Pilkonis, 1996; Dinger, Strack, Leichsenring, Wilmers & Schauenburg, 2008; Luborsky et al., 1997). These studies were chosen to represent two with Randomised

Controlled Trials designs and two with Routine Outcome Studies (ROS) designs. The independent rater was a doctoral level student, familiar with the Downs and Black (1998) scale. Coaching was given on the adaptations made to the scale for this review. Inter-rater reliability was calculated for using Spearman correlation coefficient, as used by Downs and Black (1998). Agreement was acceptable (r = .68, p < .001). The relatively low inter-rater agreement may have been due to the explicit focus on therapists in the papers reviewed, which has resulted in less emphasis being placed on details of client sample. It was therefore difficult to accurately ascertain some of the details of the client sample required by the Downs and Black scale, leaving more subjectivity in completing related items. Although coaching was given to the independent rater, this could have been enhanced through providing more clarity on coding client sample related items, to ensure less subjectivity and more agreement between raters.

3 Results

3.1 Results overview

The main features of the design and findings of the 17 papers included in the review are summarised in Table 1 with studies ordered chronologically. The studies fell into three categories based on their design: randomised controlled trials (RCTs; n= 5); routine outcome studies (ROSs; n = 11) and non-randomised trial (n=1). The analytic strategy employed was categorised as multi-level (MLM; n=9) or single level analysis (n=8), reflecting higher-order differences in level of analyses. The data in Table 1 shows the increased use of MLM as an analytic strategy since 2006, with more recent studies utilising this type of analysis. The variety of design and methodologies resulted in 15 studies finding variability between therapists and two studies finding no variance.

Table 1

Overview of salient factors involved in the reviewed studies

Author (Year)	Setting	Design	Client Diagnosis	Treatment modality	Analytic strategy	Quality Index	Therapist effect present
Lafferty, Beutler & Crago (1989)	University psychiatric clinic – US	ROS	Anxiety or affective disorders	Mixed	Single	11	Yes
Najavits & Strupp (1994)	Outpatient psychotherapy – US	Non-randomised trial	Mixed	Time limited dynamic	Single	10	Yes
Blatt et al. (1996)	University psychiatric clinic – US	RCT	Depression	IPT & CBT	Single	14	Yes
Luborsky et al. (1997)	Community psychotherapy – US	RCT	Drug addiction & depression	Mixed	Single	11	Yes
Huppert et al. (2001)	Outpatient psychotherapy – US	RCT	Panic Disorder	CBT	Single	14	Yes
Okiishi et al. (2003)	University counselling centre - US	ROS	Mixed	Mixed	MLM	10	Yes
Brown et al. (2005)	Managed Care* - outpatient practice - US	ROS	Mixed	Mixed	Single	10	Yes
Elkin et al. (2006a)	University psychiatric clinic - US	RCT	Depression	IPT & CBT	MLM	13	No
Kim, Wampold, & Bolt (2006)	University psychiatric clinic - US	RCT	Depression	IPT & CBT	MLM	13	Yes

Author (Year)	Setting	Design	Client Diagnosis	Treatment modality	Analytic strategy	Quality Index	Therapist effect present
Sandell et al. (2006)	Subsidized outpatient Psychotherapy - Sweden	ROS	Mixed	Psychodynamic and psychoanalytic	Single	11	Yes
Okiishi et al. (2006)	University counselling centre - US	ROS	Mixed	Mixed	MLM	13	Yes
Lutz et al. (2007)	Managed care* - outpatient practice - US	ROS	Mixed	Mixed	MLM	11	Yes
Dinger et al. (2008)	Inpatient Psychotherapy – Germany	ROS	Severe neurotic and personality disorders	Psychodynamic	MLM	10	Yes
Anderson et al. (2009)	University counselling centre - US	ROS	Mixed	Mixed	MLM	11	Yes
Cella et al. (2011)	Specialized outpatient clinic – UK	ROS	Chronic Fatigue Syndrome	CBT	MLM	11	No
Kraus et al. (2011)	Outpatient psychotherapy – US	ROS	Mixed	Mixed	Single	11	Yes

*Managed Care Organisations (MCOs), are US based insurance organisations whose purpose is to deliver cost-efficient healthcare.

The quality of the studies as rated by the Downs and Black (1998) Quality Index is shown in Table 1. The range of Quality Index ratings is small (10-14) indicating that all the selected studies were of good quality. Studies utilising an RCT design had a mean score above the cut-off of 11.74 of 1.40; the one non-randomised trial scored .20 above the threshold of 9.8 and the ROS studies had a mean of 1.02 above the cut-off of 9.8. This shows that the studies were comparable in terms of their quality. Only one study fell marginally below the mean (Luborsky et al., 1997; score of 11 which is .74 below the cut-off for randomised studies). However as this score was close to the mean, it was included in the review.

The number of therapists and clients included in each sample is reported in Table 2 to examine the fulfilments of Soldz (2006) recommendations relating to sample size. Table 2 provides a breakdown of the mean and minimum numbers of clients within these samples and shows that only one study reached the recommended sample of 30 therapists each with 30 clients (Soldz, 2006). The remaining 16 studies included a wide range of sample sizes, with a total client sample of 60 - 10,812 and therapist samples of 12 - 696.

The studies are discussed separately according to their use of RCT, ROS or non-randomised trial design. Studies are further subdivided based on their analytic strategy of whether they used MLM or another type of analysis.

Table 2:

Numbers of clients and therapists in the samples of reviewed studies

Author (Year)	No. Clients	No. Therapists	N of clients per therapist	
			Mean	Minimum
Lafferty et al. (1989)	60	30	2	2
Najavits & Strupp (1994)	80	16	5	5
Blatt et al. (1996)	119	28	4	1
Luborsky et al. (1997)	198	22	9	2
Huppert et al. (2001)	183	14	13	-
Okiishi et al. (2003)	1,779	56	32	15
Brown et al. (2005)	10,812	281	38	15
Elkin et al. (2006)	119	17	7	4
Kim, et al. (2006)	119	17	7	4
Wampold & Brown (2006)	6,146	581	11	4
Sandell et al. (2006)	327	160	2	-
Okiishi et al. (2006)	7,628	72	106	30
Lutz et al. (2007)	1,198	60	20	10
Dinger et al. (2008)	376	50	8	10
Anderson et al. (2009)	1,141	25	46	15
Cella et al. (2011)	374	12	31	8
Kraus et al. (2011)	6,960	696	10	10

3.2 Studies employing a Randomised Controlled Trial (RCT) design

3.2.1 Overview of RCT studies

RCT research traditionally assumes that providers of a particular treatment or intervention (e.g., an anti-depressant medication) are interchangeable. Accordingly it is the intervention in question that is creating the effect and not the provider (Kim, Wampold, & Bolt, 2006). As a consequence of the literature on therapist effects, this assumption has been questioned in relation to psychotherapy research with the suggestion that the provider of psychotherapy plays an important role in outcome (Elkin, 1999; Lambert, 1989; Luborsky et al., 1986).

RCT designs control the therapist variable by training therapists to a specified standard in the identified clinical mode and thus minimise treatment differences between practitioners (Barkham, Stiles, Connell, Twigg, Leach, Lucock, et al., 2008). The five studies reviewed in this section have attempted to examine whether controlling therapists in this way does minimise variation between therapists.

Five studies of therapist effects were found that utilised randomisation of clients to therapists. Four of these studies were part of larger RCTs with a primary purpose of examining the efficacy of different treatment modalities, three of which re-analysed data from the same study - the National Institute of Mental Health's Treatment of Depression Collaborative Research Program (NIMH TDCRP; Blatt, et al., 1996; Elkin, Falconnier, Martinovich, & Mahoney, 2006a; Kim, Wampold, & Bolt, 2006). The fourth study examined the efficacy of different treatment modalities for panic disorder (Huppert, Bufka, Barlow,

Gorman, Shear, & Woods, 2001). The fifth paper in this section examined the outcomes of therapists from a number of different treatment samples (Luborsky et al., 1997).

Although the designs of these studies were similar, they differed in their analytic strategy. Three studies used a range of analyses to examine therapist variability including analysis of variance (ANOVA), analysis of covariance (ANCOVA), and percentage of mean change scores, (Blatt et al., 1996; Huppert et al., 2001; Luborsky et al., 1997 respectively). The remaining two studies utilised MLM (Elkin et al., 2006a; Kim, Wampold, & Bolt, 2006).

3.2.2 RCT studies utilising single level analysis

Three studies met the criterion of RCTs using single level analysis. Blatt et al. (1996) used ANOVA to examine therapist variance across all therapists in the sample, which resulted in finding no differences between therapists. However this finding is limited due to the treatment of therapists as a fixed effect. Huppert et al. (2001) used a similar strategy in their ANCOVA where therapists again were treated as a fixed effect. However, this study did find a range of therapist effect sizes, ranging from 1% to 18% depending on the outcome measure used (Huppert et al., 2001). Luborsky et al. (1997) reported differences between therapists ranging from a small negative change rate to over 80% improvement. Additionally, in this study, therapists who had large improvements in their caseload in one sample were consistent in showing improvement in client scores in other samples. This finding suggests that effective therapists are consistent in their effectiveness, regardless of the context of the client's presenting problem.

All three studies compared therapists based on rankings of their therapeutic outcomes. Two studies grouped therapists based on these rankings and examined the differences between more and less effective therapists and found differences between the groups in terms of outcome (Blatt et al., 1996; Huppert et al., 2001). The findings suggest that more effective therapists helped their clients to improve to a greater degree than less effective therapists (Blatt et al., 1996). On a measure of panic disorder, the most effective therapists had 66% of their caseload achieving the criteria for reliable and clinically significant change compared to 45% of the less effective therapists (Huppert et al., 2001).

Although Blatt et al. (1996) did not find a therapist effect when the entire sample of therapists was included, when comparisons of the most and least effective therapists were made, differences were found. This means that all studies found variation between therapists. This variation exists even when researchers aim to maximise therapist skill and minimise therapist differences (Luborsky et al., 1997). Crucially, no differences were found on competency and adherence to the model between the groups in one study, indicating that something other than the standardised intervention was creating the variation in therapists (Huppert et al., 2001).

3.2.3 Limitations of RCT studies utilising single level analysis

A major limitation of these studies was the use of small samples of therapists, each of whom treated a relatively small number of clients. The low statistical power limits confidence in the findings and, in particular, their generalisability (Crits-Christoph & Mintz, 1991).

Another factor limiting generalisability is the treatment of therapists as fixed effects (see Blatt et al., 1996 and Huppert et al., 2001). As previously discussed, the use of fixed effects in such studies means that the results cannot be generalised to the therapist population as a whole and therefore only relate to the sample used in the studies.

The study conducted by Blatt et al. (1996) has been critiqued for using a selective sample of therapists to compare the most and least effective therapists' groups, subsequent to a non-significant finding of therapist effects when the entire therapist sample was included (Elkin et al., 2006a). Additionally, it was suggested that their classification strategy ensured that therapist differences would be found (Kim, Wampold, & Bolt, 2006). This criticism could extend to all studies in this section, as the methodology of using change scores to create ranks provides a strategy where differences, however small, may be found and used to distinguish groups of therapists. The use of MLM analysis allows for examination of the variation in client outcomes depending on the treating therapist that does not automatically assume that differences in therapist outcomes exist.

3.2.4 RCT studies utilising MLM

Two papers employing a RCT design used a MLM approach to analysing the data (Elkin et al., 2006a; Kim, Wampold & Bolt, 2006). The data for both papers derived from the NIMH TDCRP, multi-site randomised clinical trial that evaluated psychological, drug, and placebo conditions in the treatment of depression. The trial was designed to examine the efficacy of the different

treatment approaches and, in brief, found that the two psychological treatments examined (Cognitive Behavioural Therapy and Interpersonal Therapy) produced similar outcomes (Elkin, Shea, Watkins, Imber, Sotsky, Collins et al., 1989). The analysis of therapist effects has been conducted subsequently.

MLM analysis was adopted by both groups of researchers (Elkin et al., 2006a; Kim, Wampold & Bolt, 2006). However each reported different findings: no therapist effect (Elkin et al., 2006a) versus 8% of variance in patient outcomes being attributable to therapists (Kim, Wampold, & Bolt, 2006). Two major differences in the models used have been cited in the literature as being potential reasons for the difference in the findings: the number of time points sampled, and the inclusion or exclusion of outliers (Crits-Christoph & Gallop, 2006; Elkin, Falconnier, Martinovich, & Mahoney, 2006b; Soldz, 2006; Wampold & Bolt, 2006). Each of these points is discussed respectively.

Elkin et al. (2006a) used session-by-session data to capture outcomes over time, as they argued that this allowed for the development of growth curves of client outcomes. In contrast Kim, Wampold, and Bolt (2006) used only pre- and post-treatment scores, as they argued that these scores of clinically significant change were more important than a client's journey to that end point. It was also argued that using scores across time points increases the variation between clients that can then reduce variation between therapists. As variation between therapists is the primary focus of the study, pre-post treatment scores were seen as a more appropriate method for adoption (Wampold & Bolt, 2006).

Kim, Wampold, and Bolt (2006), included outliers in their statistical model, stating that these represented a natural part of the variability in the results of different therapists. Elkin et al. (2006a), excluded any therapist outlier scores, as would be done in other forms of statistical analysis. This decision of Elkin et al. (2006a) was critiqued on the basis that it appears counterintuitive to eliminate outliers from the dataset as the major focus of therapist effects is to look for variation in the outcomes of therapists (Wampold & Bolt, 2006).

When reviewing the differences between these two studies, it has been suggested that the different methods used within their statistical models caused the conflicting findings, but the main limitation of both studies was the small sample of therapists used (Soldz, 2006). It was proposed that this sample was too small to reliably analyse the effect of individual therapists (Soldz, 2006).

3.2.5 Conclusions about RCTs

The five RCT studies of therapist effects reviewed have reported conflicting findings, with four studies reporting variation in therapists' outcomes and one reporting that therapists do not contribute to outcome. The differing designs of the studies and analytic strategies used contributed to the likelihood of finding therapist differences in some cases (specifically, Blatt et al., 1996). The methodologies adopted in all the studies were appropriate although it can be argued that the use of MLM was superior to single level models, as this strategy does not assume that variation exists. However, none of these RCT studies utilised a large enough dataset to reliably assess the extent of the effects of individual therapists.

3.3 Studies utilising a non-randomised trial design

One study met the criteria for a non-randomised trial design (Najavits & Strupp, 1994). This study was part of a wider research effort, the Vanderbilt II study, primarily examining the effect of training on treatment outcome where clients were not randomly allocated to therapists (Henry, Strupp, Butler, Schacht, & Binder, 1993). The Vanderbilt II study recruited participants via newspaper announcements and treatment was delivered by experienced therapists undergoing training in Time Limited Dynamic Psychotherapy (Strupp & Binder, 1984). Najavits and Strupp (1994) examined the differences between therapists by using treatment outcome scores and drop-out rates. They were able to classify groups of therapists as “more” or “less” effective based on these outcomes and found that the most effective therapists were superior in terms of their drop-out rates, which was linked to therapists’ ability to create positive change in their clients. Hence the most effective therapists enabled clients to make this change and therefore their clients remained in therapy (Najavits & Strupp, 1994).

A limitation of this study is the method of recruitment of clients as this limits the generalisability of the findings. The authors state that “subjects were selected to be comparable in severity to ordinary outpatient samples” (p.116), however it is unlikely that participants responding to newspaper articles would be completely comparable to a sample of psychotherapy outpatient clients in routine practice.

A further limitation is the small sample of therapists (n=15) who each treated a small number of clients each (n=5). As previously discussed, this limits the

statistical power of the findings and therefore their generalisability (Crits-Christoph & Mintz, 1991).

3.4 Studies utilising routine outcomes

3.4.1 Routine outcomes studies (ROSs) overview

This section reviews the growing body of evidence utilising routinely collected outcome data in real world settings where clients have not been randomised to therapists. Table 1 shows that the settings of the papers varied, but all included samples of therapists who were practising routinely with a heterogenous group of clients that were more representative of real world clinical work than those found in RCTs (Kraus, Castonguay, Boswell, Nordbery, & Hayes, 2011). Although similar in initial design, the studies reported here differed in their analysis of therapist effects, namely by using a single or multi-level approach. The studies will be discussed based on their analytic strategy of choice.

3.4.2 Routine outcome studies utilising single level analysis

Four studies carried out in routine practice utilised a single level analytic strategy (Brown, Lambert, Jones & Minami, 2005; Kraus et al., 2011; Lafferty, Beutler, & Crago, 1989; Sandell et al., 2006). All four studies adopted an approach that contrasted more versus less effective therapists, although the strategies for achieving these contrasts were different.

Two studies, Brown et al. (2005) and Lafferty, Beutler, and Crago (1989), ranked therapists based on residualized change scores, accounting for initial case mix of clients on therapists' caseload and then creating groups of more

effective and less effective therapists. In contrast, the study by Kraus et al. (2011) created categories of therapists based on outcomes of effect sizes and the reliable change index (Jacobson & Truax, 1991). In this study, "effective" and "harmful" therapists were defined as those whose clients on average reliably improved, or reliably deteriorated respectively. Using reliable change overcomes the limitations of categorising therapists noted by Kim, Wampold and Bolt (2006) as the categorisation in this study was based on more stringent criteria than simply those therapists with the highest level of change.

Sandell et al. (2006) conducted a non-parametric latent class regression to determine therapist effects in their sample. This provided a score for clients' average rate of change across their treatment. Group assignment was made based on a cluster analysis of therapists with clients with similar levels of outcomes. Five classes of therapists were identified, with the most effective therapists (Class 1) accounting for 31% (n=52) of the sample and the least effective (Class 5) accounting for 11% (n=18).

The ability to be able to create distinct categories based on therapist outcomes suggests that there is variability between therapists. However, additional comparisons were made between the groups to examine differences in their outcomes in some studies. Brown et al. (2005) reported that therapists in the "highly effective" group yielded three times as much change in their clients' scores on average than therapists in the "other" category. Kraus et al. (2011) found that there were wide variations in therapist effectiveness across a range of diagnostic domains. For example, the percentage of effective therapists ranged from 29% when treating sexual dysfunction to 67% when treating

depression. Additionally, there was a range of therapists classified as harmful, ranging from 3% treating depression to 16% in treating substance abuse.

Two studies examined therapist variables associated with more and less effective therapist groups (Lafferty, Beutler, & Crago, 1989; Sandell et al., 2006). Sandell et al. (2006) found that the most effective therapists were characterised by high scores on kindness and neutrality. Lafferty, Beutler and Crago, (1989) found that more effective therapists rated themselves higher on measures of empathic understanding and reported greater valuing of intellectual goals than less effective therapists.

3.4.3 Limitations of ROS studies utilising single level analysis

As Table 2 shows, the limitations noted in previous sections relating to sample size equally applies to routine outcome studies. On initial inspection, the study sample sizes appear quite large. However, when these are broken down into the minimum number of clients nested within each therapist, the numbers are much smaller.

The major limitation of these studies is the analytic strategy used by researchers. Although the studies demonstrate that therapist outcomes vary, the comparisons of therapists in groups does not allow for the variance across the whole sample to be analysed. Nevertheless, taken together these studies demonstrate that therapist outcomes in routine practice vary.

3.4.4 Routine studies utilising MLM

Two studies used a MLM analytic strategy and examined the rate of change for clients (Okiishi, Lambert, Nielsen, & Ogles, 2003; Okiishi, Lambert, Eggett, Nielsen & Dayton, 2006). Okiishi et al. (2003) found that differences in therapists' rate of change were substantial, with the most effective therapist showing improvement rates in their clients at ten times the average rate for the whole sample. Not only did the top rated therapists show the greatest rate of change in their clients, but they also achieved this in the shortest amount of time.

Extending this study to a larger sample, Okiishi et al. (2006) also ranked therapists on two different indices. The first was on the average rate of change analysed by MLM, and the second was on the overall change scores of clients on each therapist's caseload. Rankings on the two indices were averaged to create a final, composite rank. The top and bottom 10% of therapists were compared on the amount of clients who achieved reliable change (Jacobson & Truax, 1991). It was reported that therapists in the top grouping had an average of 22.4% of their clients who met the reliable change criteria compared to 10.6% in the bottom grouping of therapists (Okiishi et al, 2006).

These two studies highlight that there are differences in client outcome that can be attributed to therapists when MLM analysis is used. They also demonstrate that some therapists can achieve significantly better results faster than other therapists.

One study used MLM analysis and examined the role of Facilitative Interpersonal Skills (FIS) as a predictor of therapist variance in outcomes (Anderson et al., 2009). Anderson et al. (2009) used the same outcomes database as Okiishi et al. (2006), and included a performance task measure of FIS as rated by an independent observer and therapist self-reported social skills. The results showed that therapists with higher FIS scores had clients with greater rates of change and that FIS was a significant predictor of client outcome. This suggests that facilitative interpersonal skills, such as emotional expression and persuasiveness, are a factor that has a positive impact on client outcomes (Anderson et al., 2009).

Four studies used an MLM analytic strategy to report the percentage of variance attributable to therapists (Cella, Stahl, Reme, & Chalder, 2011; Dinger et al., 2008; Lutz Scott, Martinovich, Lyons & Styles, 2007 and Wampold & Brown, 2005). All four studies treated their therapist sample as random factors, allowing generalisation of the findings to the population of therapists from which they were drawn. Three studies used a two level model (Cella, et al., 2001; Dinger et al., 2008 and Wampold & Brown, 2005) and one used a three level model (Lutz et al., 2007).

The percentages of outcome variance attributed to therapists found in these studies varied as follows: 0% (Cella et al., 2011), 3% (Dinger et al., 2006), 5% (Wampold & Brown), and 8% of variance in clients' symptoms but 17% of variance in estimated rates of client improvement (Lutz, et al., 2007). The difference in the variations reported may be due to different samples of clients, therapists, and the methodology employed. The relatively small variation of 3%

was hypothesised as being due to the examination of therapist effects in an inpatient setting (Dinger et al., 2008). This study used outcome data from individual therapy within a hospital setting where other forms of treatment were also ongoing, including therapeutic groups. The magnitude of individual psychotherapy outcomes may be minimised therefore because other therapy components contributed to therapeutic success, thus leading to individual psychotherapy variations being smaller (Dinger et al., 2008).

Cella et al. (2011), reporting 0%, utilised a sample of experienced therapists delivering manualised CBT to clients with a diagnosis of Chronic Fatigue Syndrome (CFS). One major difference between this study and the other studies in this section is the adoption of a homogenous and selective client sample treated by therapists delivering protocol driven treatment carried out in routine practice. Hence, key design features drew on trials methodology but applied in routine settings. The selective client sample meant that scores of clients who did not complete therapy were not analysed. The distinction has been made between this type of approach, classified as an *effectiveness* study, and those conducted in the other papers in this section where treatment is delivered in routine settings with a more heterogeneous population by therapists not bound to a treatment protocol, classified as *practice-based* studies (Barkham, et al., 2008). Therapists in practice-based studies treat clients with a wide range of difficulties, using a variety of approaches, which increases the variability in outcomes. This increased variation in client presentation and treatment may account for the wider variation found in treatment outcomes. Accordingly, the restricted, homogenous group of clients sampled is likely to

have narrowed this variation and therefore reduced the naturally occurring variability between therapists found by Cella et al., (2011).

3.4.5 Limitations of ROS studies

Although a major strength of the design of naturalistic studies is the fact that they capture “real world” outcomes in psychotherapy, this strength can also be a limitation. In routine settings, clients are not randomised to therapists, meaning that there may be a disproportionate number of “difficult to treat” clients assigned to some therapists and not to others, thereby resulting in an inflation or deflation of their effectiveness (Okiishi et al., 2006; Wampold & Brown, 2005).

Most studies attempted to overcome this limitation by examining pre-treatment levels of pathology in clients across therapists’ caseloads in order to check for equivalency. However, the majority of these analyses were based on client self-reported outcome measures of specific or global symptom difficulties of distress. This takes into account only the clients’ self-reported *symptoms* as captured by the measures and infrequently examines variables such as interpersonal difficulties, motivation, and external therapy events. It could be that some therapists had caseloads with clients who were more difficult to treat in ways that were not captured by these outcome measures (Brown et al., 2005).

As with RCTs, the small sample sizes limit the reliability of the findings. Inspection of Table 2 shows that only one study reached the recommended

criteria of 30 therapists each treating 30 clients, and this study did find a therapist effect (Okiishi et al., 2006). However, it should be noted that this criterion is a recommendation and is not, to date, based on a formal power analysis. As such, smaller sample sizes may be equally placed to provide reliable results, although future studies should aim to replicate these results with larger samples in order address the question as to whether smaller samples using MLM yield similar results.

The samples used in these studies were mainly from large datasets collected routinely in managed care and outpatient psychotherapy services. One limitation of this type of data is that it relies on the inclusion of a convenience sample of clinics that adopt the relevant outcome data collection system (Kraus et al., 2011). Many therapists may opt out of such a system and the therapist sample included may therefore not be representative. Additional limitations to the use of datasets from managed care companies include the issue of restricting their samples, by definition, to those clients with particular types of insurance (Brown et al., 2005). It is unknown as to whether the findings extend to clients and therapists in other, non-insurance based populations (Wampold & Brown, 2005).

3.4.6 Conclusion of routine practice studies:

In studies of therapist effects in routine outcomes, findings suggest that therapists will have variable outcomes. When an *effectiveness* study design is employed with a homogenous group of clients with therapists who are regularly supervised using a manualised treatment in a specific model, therapist variation

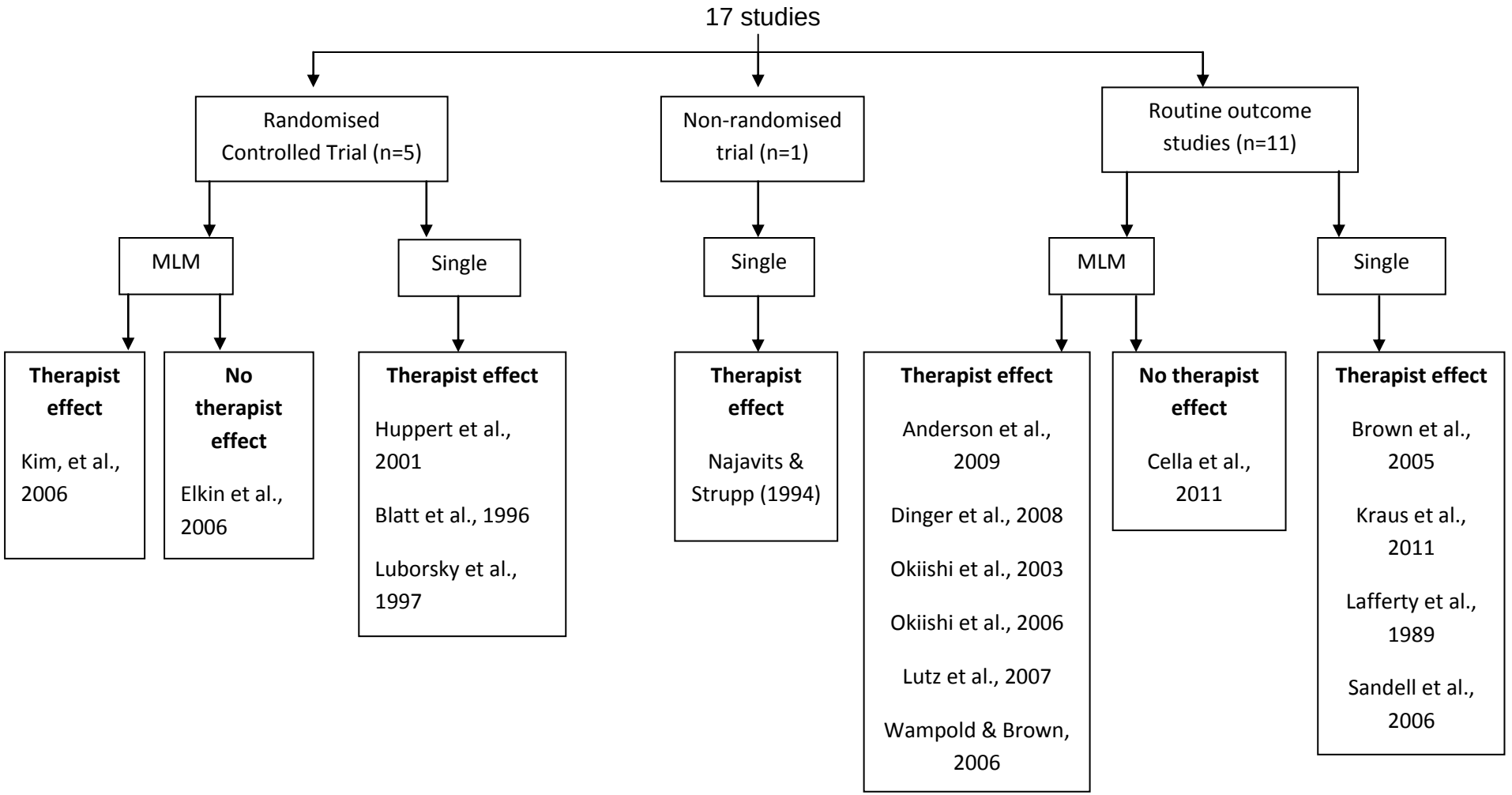
is reported to be as low as 0%. However, in *practice-based studies* sampling a real-world heterogeneous group of clients, therapist effects are found.

4. Synthesis and discussion

Figure 2 presents a synthesis of the 17 studies included in the review in terms of their design, analysis, and findings. Of the 17 studies, 15 (88%) reported variation in the outcomes of therapists. The two studies that did not find a therapist effect utilised client samples that were homogenous and therapists who were highly trained and supervised in a specific treatment modality (Cella et al., 2011; Elkin et al., 2006). Both used MLM as their analytic strategy. The reasons for the lack of finding of a therapist effect may be a function of the type of MLM used and the homogenous client sample.

The amount of variation found between therapists in studies finding a therapist effect appears to be related to the size and type of client sample used, the setting therapy is delivered in, and the analytic strategy used. Many studies used small sample sizes with few clients per therapist, which limits the generalisability of the results and increases the likelihood of Type I errors (Crits-Christoph & Gallop, 2006). Studies comparing more and less effective therapist groups have been critiqued for utilising a design that ensures differences between therapists, however small, will be found. For this reason, the analytic strategy of MLM was considered to be superior for assessing therapist effects.

Figure 2: Synthesis of results



A total of 9 papers reported using MLM analysis, which has been described as the analysis of choice for studying therapist effects where clients are nested within therapists (Lutz et al., 2007). Of these, 7 found a therapist effect. This suggests that the majority of studies of therapist effects with the most appropriate methodology support the existence of this phenomenon. However, these results pertained despite the small sample sizes, with most studies not meeting the recommended 30 by 30 requirement recommended by Soldz (2006). The results of the one study utilising this recommended sample size found similar results to other ROS using MLM analysis (Okiishi et al., 2006). This suggests that smaller samples can detect therapist effects. Soldz's (2006) recommendation was not based on a formal power analysis, and therefore it could be argued that the notion of requiring a sample of 30 by 30 needs to be tested and evidenced before it can be conclusively used as a benchmark equating to a power analysis.

The findings of this review corroborate Lambert's (1989) review in that the empirical evidence base of studies examining the individual therapist's contribution to outcome indicate that therapist effects exist. Methodological considerations and limitations remain, but the development and increased use of multi-level modelling since 2006 is helping to overcome these limitations.

5. Clinical Implications

The majority of the evidence indicates that therapist effects exist. This means that individual therapists will vary in their outcomes despite the therapeutic modality in use. The clinical implication of such a finding is that some therapists within services are likely to be more effective than others, and it should not be

assumed that all therapists will be effective with all clients. Therapists should routinely monitor their outcomes to review their own effectiveness. If outcomes are found to be unsuccessful, therapists should seek feedback on their clinical performance as this has been found to improve outcomes (Lambert, Whipple, Vermeersch, Smart, Hawkins, et al., 2002)

6. Directions for future studies

In order to carry forward research investigating therapist effects, future studies should utilise the most appropriate statistical tests. MLM is considered the most appropriate analysis for examining nested data (Soldz, 2006). In order to generalise the results to therapists outside of the sampled population, therapists should be treated as random factors in the analysis (Crits-Christoph & Mintz, 1991). Researchers should aim to include sufficiently large samples in order to test the recommendation of requiring 30 therapists each treating 30 clients.

Obtaining larger sample sizes is often problematic in research. It may be particularly difficult for RCTs aiming to examine therapist effects to recruit the required number of therapists. However, the development of more innovative psychological interventions and the need for these to be evaluated in trials – and funded through the NIHR and supported by the Mental Health Collaborative Network – may result in such studies yielding much larger Ns of therapists than previously achieved. However, to date, large data sets may be more achievable in practice-based studies. Managed care organisations in the USA have provided an opportunity to examine large datasets but there are limitations with these as previously discussed. Also in managed care, there is wide variation in the therapists and the types of treatment delivered which potentially

increases the variance. It would be interesting to conduct studies of therapist effects where therapist interventions being delivered were more homogenous.

One initiative in the UK that may provide such an opportunity is the Improving Access to Psychological Therapies (IAPT) programme. Psychological Wellbeing Practitioners (PWPs) operating at step two of the stepped care model and High Intensity Workers operating at step three all receive year-long training in CBT or guided self-help based on CBT principles. The variation in outcomes for these therapists should, therefore, be reduced by receiving similar training methods and delivering treatments within a specific modality. These practitioners routinely collect client outcomes as part of the IAPT model, making large client outcome datasets readily available (CSIP Choice & Access Team, 2008).

Whilst studies investigating the existence of therapist effects are useful and needed, future studies would benefit from including measures hypothesised to be associated with effective therapists. This would facilitate development of knowledge about what enables an effective therapist (Lambert & Okiishi, 1997)

In conclusion, studies utilising MLM provide the best indication of therapist effects and have begun to establish an evidence-base regarding this phenomenon. However future research needs to build on the use of routine samples in order to provide estimates of therapist effects that occur in “real-world” situations. RCT studies should also incorporate the examination of therapist effects, potentially with the use of cluster trials, as required by the extensions to the 2010 Consolidated Standards of Reporting Trials statement currently in development (CONSORT, 2010).

References

NB: Studies included in the review are preceded by † (electronic search) or ‡ (visual search)

† Anderson, T., Ogles, B.M., Patterson, C.L., Lambert, M.J., & Vermeersch, D.A. (2009). Therapist effects: Facilitative interpersonal skills as a predictor of therapist success. *Journal of Clinical Psychology, 65*, 755-768. doi:10.1002/jclp.20583

Barkham, M., Stiles, W.B., Connell, J., Twigg, E., Leach, C., Lucock, M., & Angus, L. (2008). Effects of psychological therapies in randomized trials and practice-based studies. *British Journal of Clinical Psychology, 47*, 397-415. doi:10.1348/014466508X311713

† Blatt, S.J., Sanislow, C.A., Zuroff, D.C., & Pilkonis, P.A. (1996). Characteristics of effective therapists: Further analyses of data from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology*, *64*, 1276-1284. doi:10.1037/0022-006X.64.6.1276

‡ Brown, G.S., Lambert, M.J., Jones, E.R. & Minami, T. (2005). Identifying highly effective psychotherapists in a managed care environment. *American Journal of Managed Care, 11*, 513-520.

Care Services and Improvement Partnership Choice and Access Team (2008) *Improving Access to Psychological Therapies (IAPT) Commissioning Toolkit.* London: Department of Health.

† Cella, M., Stahl, D., Reme, S.E., & Chalder, T. (2011). Therapist effects in routine psychotherapy practice: An account from chronic fatigue syndrome. *Psychotherapy Research*, *21*, 168-178. doi:10.1080/10503307.2010.535571

Crits-Christoph, P., & Gallop, R. (2006). Therapist effects in the National Institute of Mental Health Treatment of Depression Collaborative Research Program and

other psychotherapy studies. *Psychotherapy Research, 16*, 178-181. doi:10.1080/10503300500265025

Crits-Christoph, P., & Mintz, J. (1991). Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *Journal of Consulting and Clinical Psychology, 59*, 20-26. doi:10.1037/0022-006X.59.1.20

Department of Health. (2008). A report on the development of a Mental Health Currency Model (including suggestions on the development and testing of PbR tarrifs) London: Self, R., Painter, J., & Davis, R.

† Dinger, U., Strack, M., Leichsenring, F., Wilmers, F., & Schauenburg, H. (2006). Therapist effects on outcome and alliance in inpatient psychotherapy. *Journal of Clinical Psychology, 64,* 344-354.

Downs, S.H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality of both randomised and non-randomised studies of health care interventions. *Journal of Epidemiological Community Health, 52*, 377-384. doi:10.1136/jech.52.6.377

Elkin, I. (1999). A major dilemma in psychotherapy outcome research: Disentangling therapists from therapies. *Clinical Psychology: Science and Practice, 6,* 10-32.

† Elkin, I., Falconnier, L., Martinovich, Z., & Mahoney, C. (2006a). Therapist effects in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Psychotherapy Research, 16,* 144-160. doi:10.1080/10503300500268540

Elkin, I., Falconnier, L., Martinovich, Z., & Mahoney, C. (2006b). Rejoinder to commentaries by Stephen Soldz and Paul Crits-Christoph on therapist effects. *Psychotherapy Research, 16,*182-183. doi:10.1080/10503300500282186

Elkin, I., Shea, M.T., Watkins, J.T., Imber, S.D., Stotsky, S.M., Collins, J.F.& Parloff, M.B. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archive of General Psychiatry, 46*, 971-982.

Henry, W.P., Strupp, H.H., Butler, S.F., Schacht, T.E., & Binder, J.L. (1993). Effects of training in time-limited dynamic psychotherapy: Changes in therapist behavior. *Journal of Consulting and Clinical Psychology, 61*, 434-440. doi:10.1037/0022-006X.61.3.434

† Huppert, J.D., Bufka, L.F., Barlow, D.H., Gorman, J.M., Shear, K.M, & Woods, S.W. (2001). Therapists, therapist variables, and Cognitive-Behavioral Therapy Outcome in a multicenter trial for panic disorder. *Journal of Consulting and Clinical Psychology, 69*, 747-755. doi:10.1037/0022-006X.69.5.747

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

† Kim, D., Wampold, B.E., & Bolt, D.M. (2006). Therapist effects in psychotherapy: A random-effects modelling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research, 16,* 161-172. doi:10.1080/10503300500264911

† 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*, 267-276. doi:10.1080/10503307.2011.563249

† Lafferty, P., Beutler, L.E. & Crago, M. (1989) Differences between more and less effective psychotherapists: A study of select therapist variables. *Journal of Consulting and Clinical Psychology, 57*, 76-80. doi:10.1037/0022-006X.57.1.76

Lambert, M.J. (1989). The individual therapist's contribution to psychotherapy process and outcome. *Clinical Psychology Review, 9,* 469-485. doi:10.1016/0272-7358(89)90004-4

Lambert, M.J., & Bergin, A.E. (1994). The effectiveness of psychotherapy. In A. Bergin & S. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed) (pp. 141-150). New York: John Wiley & Sons.

Lambert, M.J.. & Okiishi, J.C. (1997). The effects of individual psychotherapist and implications for future research. *Clinical Psychology: Science and Practice, 4*, 66-75. doi:10.1111/j.1468-2850.1997.tb00100.x

Lambert, M.J., Whipple, J.L., Vermeersch, D.A., Smart, D.W., Hawkins, E.J., Nielsen, S.L., & Goates, M. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology and Psychotherapy, 9*, 91-103. doi:10.1002/cpp.324

† Luborsky, L., McLellan, A.T., Diguer, L., Woody, G., & Seligman, D.A. (1997). The psychotherapist matters: Comparison of outcomes across twenty-two therapists and seven patient samples. *Clinical Psychology: Science and Practice, 4,* 54-65. doi:10.1111/j.1468-2850.1997.tb00099.x

Luborsky, L.L., McLellan, A.T., Woody, G.E., O'Brien, C.P. & Auerbach, A. (1986). Therapist success and its determinants. *Archives of General Psychiatry, 42*, 602-611.

† Lutz, W., Scott, L., Martinovich, Z., Lyons, J.S., & Stiles, W. B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology, 54*, 32-39. doi:10.1037/0022-0167.54.1.32

Martindale, C. (1978). The therapist-as-fixed-effect fallacy in psychotherapy research. *Journal of Consulting and Clinical Psychology, 46*, 1526-1530. doi:10.1037/0022-006X.46.6.1526

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses. The PRISMA statement. *Public Library of Science Medicine, 6,* e1000097. doi:10.1371/journalpmed10000097

‡ Najavits, L.M., & Strupp, H. H. (1994). Differences in the effectiveness of psychodynamic therapists: A process-outcome study. *Psychotherapy, 31*, 114-123. doi:10.1037/0033-3204.31.1.114

† Okiishi, J.C., Lambert, M.J., Eggett, D., Nielsen, L., & Dayton, D.D. (2006). An analysis of therapist treatment effects: Towards providing feedback to individual therapists on their clients' psychotherapy outcome. *Journal of Clinical Psychology, 62*, 1157-1172. doi:10.1002/jclp.20272

† Okiishi, J.C., Lambert, M.J., Nielsen, S.L., & Ogles, B.M. (2003). Waiting for supershrink: Empirical analysis of therapist effects. *Clinical Psychology and Psychotherapy, 10,* 361-373. doi:10.1002/cpp.383

† Sandell, R., Lazar, A., Grant, J., Carlsson, J., Schubert, J., & Broberg, J. (2006). Therapist attitudes and patient outcomes. III. A latent class analysis of therapists. *Psychology and Psychotherapy: Theory, Research and Practice, 79,* 629-647. doi:10.1348/147608306X105694

Siemer, M., & Joormann, J. (2003). Power and measures of effect size in analysis of variance with fixed versus random nested factors. *Psychological Methods*, *8*, 497–517. doi:10.1037/1082-989X.8.4.497

Smith, M., Glass, G. & Miller, T. (1980). *The benefits of psychotherapy.* Baltimore: Johns Hopkins.

Soldz, S. (2006). Models and meanings: Therapist effects and the stories we tell. *Psychotherapy Research, 16,* 173-177. doi:10.1080/10503300500264937

Strupp, H.H., & Binder, J.L. (1984). *Psychotherapy in a new key.* New York: Basic Books.

Wampold, B.E., & Bolt, D.M. (2006). Therapist effects: Clever ways to make them (and everything else) disappear. *Psychotherapy Research, 16*, 184-187. doi:10.1080/10503300500265181

† Wampold, B.E., & Brown, G.S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73,* 914-923. doi:10.1037/0022-006X.73.5.914

Chapter 2: Research Report

Key components of effective Psychological Wellbeing Practitioners: Evidence from routine practice.

Research Report Abstract

The purpose of this research was twofold; (1) to examine whether therapist effects were present in a sample of Psychological Wellbeing Practitioners (PWPs) working within step two in Improving Access to Psychological Therapies services and (2) to examine what factors were associated with effectiveness of PWPs.

A therapist effect was found, using multi-level modelling, where PWPs accounted for almost 9 per cent of the variance in outcome. PWP ranks were created in order to compare the most and least effective PWPs. More effective PWPs had higher levels of self-rated resilience, and supervisors rated less effective PWPs as using more “experiential” intuition (i.e., affect driven processing style). Qualitative analysis of PWP and supervisor interviews revealed that more effective PWPs approached their work with confidence and in an organised manner. The qualitative analysis results suggest that more effective PWPs may have been at an advanced practitioner developmental stage than the less effective PWPs.

The findings indicate that therapist effects exist in a sample of PWPs, delivering a set of standardised interventions, and contributes to the literature suggesting that therapist effects is a real phenomenon. Resilience appears to play a role in effective PWPs and future studies should aim to explore this further.

1. Introduction

1.1 Improving Access to Psychological Therapies

In recent years the landscape of psychological therapies in the UK has radically changed with the introduction of the UK government's Improving Access to Psychological Therapies (IAPT) programme. This was introduced in response to the Layard Report, which described how a lack of psychological therapists trained in evidence-based psychological therapies was preventing delivery of effective interventions for people experiencing depression and anxiety (Layard, 2006). This government initiative led to the development of the IAPT programme that was piloted in two demonstration sites from 2006 (Clark et al., 2009). It was subsequently rolled out nationally from 2008 whereby greater access to evidence based psychological therapies was provided for people suffering from anxiety and depression (CSIP Choice & Access Team, 2008).

1.1.1 Psychological Wellbeing Practitioners (PWPs)

A key characteristic of the IAPT initiative has been the creation of the psychological wellbeing practitioner (PWP) operating at step two of the IAPT stepped care model (CSIP Choice & Access Team, 2008). The PWP role differs from a traditional psychological therapist role in that PWPs restrict their work to clients presenting with mild levels of psychological distress. PWPs tend to have a "low contact high volume" approach to intervention, ensuring high caseloads with a recommendation of between 175 and 250 clients per year (IAPT, 2008). This caseload is much higher in comparison to, for example, traditional psychotherapy caseloads. PWPs use primarily cognitive-behaviourally based, low intensity psychological interventions, such as assisted self-help and behavioural activation (Richards & Whyte, 2009). This type of approach

requires PWP sessions to be short, usually lasting 30 minutes and may be over the telephone rather than face to face (Richardson & Richards, 2009). With such high caseloads, supervision plays a vital role in helping PWPs manage their cases and to help facilitate their continuing development (Turpin & Wheeler, 2011).

PWPs undertake a 12-month training course leading to a Post Graduate Certificate (PG Cert) in Low Intensity Working, at a Higher Education Institute (HEI) local to their employing service (Richardson & Richards, 2009). This course combines training days within the HEI and concurrent clinical work in the employing service for 4 days per week. In order to pass the training, PWPs must demonstrate competence in delivering seven standardised treatment protocols of behavioural activation, exposure therapy, cognitive restructuring, medication support, problem solving, panic management and sleep hygiene (Richardson & Richards, 2009). These seven, manualised treatments form the basis of their day-to-day work with clients and constitute the "PWP clinical method".

Due to its relatively recent development, no research currently exists into the PWP role. An evaluation of the first year national rollout of IAPT to 32 sites illustrated that 37.5 per cent of the total clients receiving low intensity interventions reached recovery (Glover, Webb, & Evison, 2010). However, there was a considerable range in recovery rates across services from 14 to 54 per cent (Glover et al., 2010).Variation across services is likely to be matched by variation within services. However, there has been no research examining what makes the PWP role effective, or whether there is indeed variation between

PWPs (i.e., whether there is a therapist effect). In order to explore these questions, consideration of methodologies used to examine therapist effects and effective practitioners more generally is presented.

1.2 Key components of traditional therapists

Identifying what factors contribute to effective therapists has proved difficult to capture accurately due to methodological difficulties (Hubble, Duncan, & Miller, 1999). Two strands of research have therefore emerged; (1) a focus on the outcomes achieved by individual therapists in order to examine possible variation between therapists, thereby determining the presence or not of therapist effects, or (2) a focus on therapists themselves and identifying factors associated with effectiveness.

1.2.1 Therapist effects

The research paradigm focusing on treatments as delivered in routine practice has found that there is variance in the effectiveness among individual therapists within specific approaches, (e.g., Luborsky et al., 1985; Okiishi, Lambert, Nielsen & Ogles, 2003; Wampold & Brown, 2005). This variation appears to be consistent across studies in practice-based settings, as reported in the preceding literature review (e.g., Brown, Lambert, Jones, & Minami., 2005; Lutz, Scott, Martinovich, Lyons, & Stiles, 2007; Okiishi, et al., 2003). However the evidence is less clear when therapist effects are examined in randomised controlled trials (RCTs). In these studies, interventions are protocol driven and therapists are highly supervised to a specific therapeutic modality, with an aim of minimising variation in the delivery of treatment across therapists. (Barkham et al., 2008). Despite attempts to minimise such differences, some studies have

found a therapist effect (e.g., Huppert, Bufka, Barlow, Gorman, & Shear, 2001). Moreover, conflicting results have arisen from analyses of the same data set, namely the National Institute of Mental Health's Treatment of Depression Collaborative Research Program (NIMH TDCRP; Elkin et al., 1989). Two studies analysed the same data from the NIMH TDCRP with Kim, Wampold, and Bolt (2006) reporting evidence of a therapist effect while Elkin et al. (2006a) reported no effect. A major reason for the contrasting findings has been attributed to differences in the multi-level models used in the analysis (Crits-Christoph & Gallop, 2006; Soldz, 2006).

Multi-level modelling (MLM) has been suggested as the most appropriate method for analysing therapist effects (Okiishi et al., 2003; Raudenbush & Bryk, 2002). MLM takes into account the nested nature of the data, acknowledging that clients are treated by a particular therapist and so all clients who receive an intervention from the same therapist are "nested" within that practitioner. In order for the effect of therapists to be analysed, the analysis is carried out on two levels with clients at level one and therapists at level two. The controversy surrounding the different outcomes from the same dataset, as outlined above, stems from differences in the models used but also, it was argued, that the therapist sample used in the study was not large enough to detect reliable results (Soldz, 2006). In response, it was recommended that when using MLM researchers should aim to include a minimum of 30 therapists with each therapist treating at least 30 clients (Soldz, 2006). However, this 30x30 rule is only a research recommendation and, at present, there is no empirical evidence to support it.

Although methodological issues and controversies remain in the examination of therapist effects, it is widely accepted that therapist effects exist (Wampold, 2001). However research has as yet been unable to conclusively identify what makes some therapists more effective than others.

1.2.2 The effective practitioner

Factors that are not associated with effective practitioners have been widely reported in the literature. These include age and gender (Okiishi et al., 2003), ethnicity (Roth & Fonagy, 2005), years of experience (Faust & Zlotnick, 1995), and training discipline (Najavitis & Strupp, 1994).

Studies examining the role of specific factors associated with effective therapists have been criticised on methodological grounds, specifically the use of case notes, self-report measures, and single measures of outcome (Najavitas & Strupp, 1994). Notwithstanding these limitations, factors associated with effective therapists can broadly be separated into two categories of in-therapy and out-of-therapy factors.

In-therapy factors include having superior ability to establish a therapeutic alliance (Luborsky, et al.,1985), showing more warmth, affirmation, understanding and helping and protecting (Najavitis & Strupp, 1994), showing less active hostility (Najavitis & Strupp, 1994), and demonstrating good relationship skills (Jennings & Skovholt, 1999). Many of these factors appear to be pan-theoretical and are often referred to as common factors (Weinberger, 1993). A number of factors have been muted as being common factors including therapist intuition (Welling, 2005). In the past intuition has been

associated with mysticism and spirituality, leading it to be criticised for being unscientific (English, 1993). More recently attempts have been made to understand intuition as being a rapid combination of cognitive and affective information leading to a decision (Pretz & Totz, 2007). Research into intuition in psychotherapy is limited but is beginning to become more common, although how this relates to the effective practitioner remains unknown (Rea, 2001; Welling, 2005).

Out-of-therapy factors include good emotional adjustment (Luborsky, et al.,1985), being highly self-critical of therapeutic performance in sessions (Najavitis & Strupp, 1994), self-reflective abilities (Jennings & Skovholt, 1999), and an emphasis on hard work and using feedback to improve performance (Miller, Hubble, & Duncan, 2008). These out-of-therapy factors suggest that therapists who seek out feedback and who are emotionally mature and adjusted enough to be able to use this feedback are more effective. Therapists who wish to increase their effectiveness, therefore, may need to be able to emotionally handle receiving feedback and use this constructively to enhance their practice. Two factors that may capture this quality in therapists are ego strength and resilience.

Ego strength can be defined as an ability to maintain a sense of self in the face of challenges, and an ability to manage conflicts without becoming overwhelmed (Markstrom, Sabino, Turner, & Berman, 1997). This construct has been positively associated with measures of internal locus of control and self-esteem, which has led to the suggestion that ego strength is “indicative of psychosocial maturity and adjustment” (Markstrom & Marshall, 2007, p. 67).

Resilience has been defined as abilities and characteristics that provide individuals with the skills to cope with, and bounce back from, adverse situations (Rutter, 1993). Tusai and Dyer (2004) provided a review of the construct of resilience and reported that characteristics associated with high resilience were both intrapersonal (including optimism and intelligence) and environmental (including perceived social support). The role of these constructs in relation to therapists is unknown.

1.3 Identifying key components of effective PWPs

Research evidence suggests that multiple factors are associated with effective therapists, although methodological limitations apply. As the PWP role has only been in operation since 2008 and because PWPs provide clinical interventions in a different way to traditional therapists, little is known about what factors contribute to effective PWPs. The present research, therefore aimed to explore the PWP role and address the following research questions:

1. Do therapist effects exist in a population of PWPs – that is, do PWPs vary in their contribution to client outcomes?
2. What factors are associated with effective PWPs?

The experimental hypotheses for the current project are as follows:

1. There will be variation in the outcomes of PWPs, with some PWPs having superior outcomes to their peers
2. PWPs in the upper quartile, based on practitioner ranks, will have higher levels of self-reported intuition, ego strength and resilience than PWPs in the lowest quartile.

3. Supervisors of PWPs in the upper quartile will rate their supervisees as having higher levels of self-reported intuition than supervisors of PWPs in the lowest quartile.
4. PWPs in the upper quartile will report differences in the characteristics that describe their working practice compared to PWPs in the lowest quartile.
5. Supervisors of PWPs in the upper quartile will report differences in how the PWP engages with their work compared to supervisors of PWPs in the lowest quartile.

2. Method

2.1 Design Overview

The design of this study was driven by the recommendations of a 30 therapist by 30 clients sample size needed for reliably examining therapist effects (Soldz, 2006). Recruitment was therefore designed to aim to achieve a sample of 30 PWPs.

The study adopted a cross-sectional design comprising a volunteer sample of PWPs who completed their PG Cert Low Intensity in 2010 and subsequently worked within six IAPT services located across the North of England, UK. Four sources of information were collated and analysed: (a) electronic download data of client outcomes routinely collected within their IAPT service, (b) questionnaires of self-rated intuition, ego strength and resilience, (c) interview data with PWPs focusing on work engagement, and (d) supervisor-rated questionnaire (intuition) and interview data to gain a supervisor perspective on PWP effectiveness. Accordingly, the design utilised a triangulated view of

effectiveness (practitioner, client, and supervisor) as well as a mixed methods approach to the analyses of the data.

2.2 Participants

Participants were drawn from three perspectives: PWPs, supervisors, and clients. Each perspective is detailed below.

2.2.1 Psychological Wellbeing Practitioners (PWPs)

All PWPs had completed their Low Intensity PG Cert training in 2010 at the Universities of Sheffield, Nottingham or York and were employed by NHS trusts (n=13), private (n=1), or voluntary (n=1) organisations. These 15 possible sites offering IAPT services within the regions of Yorkshire, Derbyshire and Nottinghamshire, were approached and invited to participate, of which 9 (60%) agreed. Across these 9 services, all eligible PWPs were approached by email or by presentations about the research from the lead researcher, and invited to participate (n=47). Of these 47 PWPs, 31 (66%) agreed and were then sent information sheets (Appendix V), consent forms (Appendix VI) and questionnaire packs. Subsequently, 3 of the 9 services were unable to provide client outcome data for 8 participating PWPs in their service due to technical difficulties with data management and retrieving data from storage. In addition, two PWPs dropped out of the project in the remaining participating services.

The final sample comprised 21 PWPs (5 male, 16 female) across 6 services providing full client outcome datasets. They had a mean age of 29.9 years (SD = 7.6, range 23 – 52 years). These PWPs treated a mean of 53.5 clients in the study period, from when they started in their service to the end of February

2011, ranging from 8 to 197 clients. The mean age of clients on each PWP's caseload ranged from 36.09 to 45.68 years and mean number of sessions each PWP saw clients for ranged from 2.70 to 7.06 sessions.

2.2.2 Supervisors

All supervisors of the 21 participating PWPs were approached and invited to take part in the research. After two reminders, a total of 17 (81%) supervisors agreed to participate and were sent information sheets (Appendix VII) and consent forms (Appendix VIII). Supervisors were required to complete a questionnaire and participate in an interview about how their PWP supervisee approached their work. Demographic information was not available for supervisors.

2.2.3 Clients

Routinely collected, anonymised client data was obtained from electronic downloads from the participating services. Requests for data were made to data managers in IAPT services, who ensured anonymity of client details. Client outcomes were included in the project data set if; (1) clients had attended at least two sessions with a participating PWP, which included an assessment; (2) clients had completed a standard battery of outcome measures at the first and last session and (3) clients had attended individual sessions with PWPs as group work data was not included. Accordingly, data included both completed cases and "in treatment" cases in order to increase the number of clients seen by these PWPs and to get the most up-to-date reflection of their skills and effectiveness since completing their training. However, two services were only able to provide completed cases.

Complete datasets were obtained for 1,122 clients. Clients had a mean age of 41.1 years (SD = 14.2 years; range = 16 - 92 years) with females comprising 64.7% of the sample. In terms of ethnicity, 65.8% identified themselves as Caucasian, 2.8% as Asian, 0.7% as Black Caribbean or African, and 1.1% as mixed race. Ethnicity information was not available for 29.3% of the sample. Clients received an average of 4.57 sessions/contacts (SD = 2.88; range 2 – 21).

2.3 Measures

2.3.1 PWP measures

PWPs completed a battery of measures focusing on intuition, ego strength, and resilience. Measures were mailed to PWPs and were packaged in order as listed below:

Psychological Wellbeing Practitioner Demographic Information Sheet (Appendix IX): This sheet requested demographic information relating to the PWP's gender, age, previous experience, and training in mental health.

Ego strength (Appendix X): The Psychosocial Inventory of Ego Strengths (PIES; Markstrom, et al., 1997) was used to measure ego strength. The PIES comprises 64 items, rated on a 5-point likert scale (rated 1-5) that can be summed to give a total Ego Strength score. An example item is: *"I have strengths that enable me to be effective in certain situations".* The PIES has been shown to have good internal consistency (α = 0.94; Markstrom et al., 1997) and good construct validity (Markstrom & Marshall, 2007).

Intuition (Appendix XI): The Rational-Experiential Inventory (REI; Pacini & Epstein, 1999) was used to measure intuition. The REI assesses an

individual's preference for either rational or experiential cognition (20-item scales each). Items are rated on a 5-point scale (rated 1-5). Rationality scale items include: *"I have a logical mind"* and experiential items include: *"I believe in trusting my hunches".* The two REI scales have good internal consistency (Rationality scale: α = 0.90; Experientiality scale: α = 0.87; Pacini & Epstein, 1999) and test-retest reliability (Rationality scale: *r* = 0.76; Experientiality scale: *r* = 0.83; Handley, Newstead, & Wright, 2000).

Resilience (Appendix XII): The Connor-Davidson Resilience Scale was used to measure resilience (CD-RISC; Connor & Davidson, 2003). This is a 25-item measure, with each item rated on a 5 point likert scale (rated 0-4). Items are summed to give a total resilience score and include *"under pressure, I stay focused and think clearly".* The CD-RSIC has good internal consistency (α = 0.89) and test-retest reliability (intraclass correlation coefficient = 0.87).

Interview Schedule: The PWP interview schedule was developed based on the Jennings and Skovholt (1999) qualitative study of master therapists. Four of the original interview questions were included in the original schedule such as "what is particularly therapeutic about you?". These were adapted and revised based on feedback from three pilot interviews and based on the specifics of the PWP role. The original and revised schedules are presented in Appendices XIII and XIV.

2.3.2 Supervisor Measure:

Intuition: Supervisors completed the Rational-Experiential Inventory (REI; Pacini & Epstein, 1999). As the focus was on their named PWP supervisee rather than them as supervisors, the questions were re-worded and framed in the 3rd person (i.e., "the supervisee has a logical mind"; Appendix XV).

However, as the REI is designed as a self-report rather than an other-rated measure, reliability and validity data cannot be directly transferred or assumed.

Interview Schedule: The interview schedule for supervisors was developed based on the Jennings and Skovholt (1999) qualitative study of master therapists, but this time adapted for the specifics a supervisory role. Four of the original interview questions were included in the original supervisor schedule such as “what distinguishes a good therapist from a great therapist?” (therapist changed to PWP). The questions were adapted and revised based on feedback from one pilot interview. The original and revised schedules can be seen in Appendices XVI and XVII.

Supervisors were not asked to complete measures of ego strength and resilience about their PWP due to difficulties in rewording questionnaires and to minimise the burden on supervisors’ time.

2.3.3 Client Measures:

Patient Health Questionnaire-9 (PHQ-9; Appendix XVIII): The PHQ-9 is a brief measure of depression (Kroenke, Spitzer, & Williams, 2001). It comprises 9 items with all items relating to DSM-IV classifications of symptoms of depression. A score of 10 is recommended as a cut-off for clinical samples. Items include: *“feeling down, depressed, or hopeless”.* The PHQ-9 is reported to have high sensitivity (92%) and specificity (80%) when using a cut-off of a score of 10 (Gilbody, Richards, Brearly, & Hewitt, 2007). It is also reported to have good construct validity and internal reliability (α = 0.89; Kroenke, Spitzer, & Williams, 2001).

Generalised Anxiety Disorder (GAD-7; *Appendix XIX*): This is a brief measure of anxiety (Spitzer, Kroenke, Williams, & Lowe, 2006). The GAD-7 comprises 7 items with total scores ranging from 0-21. Items include: *"trouble relaxing".* Using a cut-off of 10 points, the GAD-7 is reported to have good sensitivity (98%) and specificity (82%) (Gilbody et al., 2007). The GAD-7 has good construct validity, internal consistency (α = 0.92) and test-retest reliability (*r* = 0.83; Spitzer et al., 2006).

2.4 Procedures

2.4.1 Recruitment

NHS Ethical approval for the project was received from the South Yorkshire Research Ethics Committee (REC Reference: 10/H1310/56; Appendix XX). Governance approval was received from all participating NHS trusts and organisations. Following these approvals, potential participants were approached via email or from presentations by the lead researcher. Those PWPs agreeing to participate were mailed questionnaire packs, with freepost returns. Once written consent was received from the PWP, supervisors were contacted and outcome data was sought from the employing service.

2.4.2 Quantitative Data and Blinding Procedures

In order to ensure no contamination or bias in the interviews as a result of information gained from the PWP and supervisor measures, the lead researcher did not have sight of these questionnaires until after completion of all interviews. Returned questionnaires were therefore stored by a member of the research team (S.K.) until all interviews had been carried out.

Client routine outcome data was collated electronically from pre-existing datasets held by each employing IAPT service. The results of session-by-session outcome scales, including PHQ-9 and GAD-7, were inputted as mandated by the Department of Health into electronic patient information systems and held by the service.

In order to minimise any potential bias arising from knowledge of the effectiveness of individual PWPs, blinding procedures were used when requesting the outcome data from services. To ensure that the lead researcher, who conducted the interviews, was blind to outcome data whilst the interviews were being conducted, all outcome data was sent to a third party (D.S.) who was not involved in the interviews. Accordingly, the researcher did not access client-completed, PWP-completed, or supervisor-completed measures prior to interviewing PWPs or supervisors.

An additional layer of blinding was added by ensuring that PWPs were anonymised in their datasets by data managers, and anonymity was checked by the third party. This enabled the lead researcher conducting the multi-level modelling analysis to do so without identifying any PWPs. This minimised any bias in the analysis and ranking of PWP effectiveness.

2.4.3 Interview Procedure

The interview schedule was piloted in January 2011 with three PWPs who were trained in earlier IAPT cohorts and therefore were not eligible to participate in the current study. One of these participants was also a trained supervisor and therefore piloted the supervisor interview schedule. Amendments to the

schedules were made based on feedback from these interviews, which ensured the questions were more focussed and less general, with specific elements of how the question related to the PWP's clinical practice. For example, an original question was "how important to you is supervision?" and was changed to "how do you use and engage with clinical and case management supervision to improve your skills as a PWP?" thus reflecting a more focussed and specific question.

Once a revised version of the interview schedule had been developed, the 21 participating PWPs and the 17 supervisors engaged in the interviews over the telephone (n=28) or face-to-face (n=10) depending on geographical proximity to the lead researcher. The order of the interviews (i.e., PWP or supervisor interviewed first) was counterbalanced to minimise any order effects. Interviews were conducted between 18th February and 1st June 2011.

2.5 Data Analyses

2.5.1. Electronic client outcome data

To determine levels of therapist effectiveness, electronic anonymised client outcomes for each PWP were analysed using multi-level modelling. Data was included in the analysis from the date a PWP started in their service to the end of February 2011. This applied to all but one service which was only able to provide data until the end of September 2010.

Analysis was conducted using multi-level modelling software (MLwiN v2.3; Rabash, Charlton, Browne, Healy, & Cameron, 2009). In order to allow generalisation to other populations of PWPs, therapists were treated as random

variables (Kim, Wampold, & Bolt, 2006). Two separate models were developed, one based on the PHQ-9, the other on the GAD-7, to examine differences in therapist performance these measures. Pre-treatment scores and interactions between the measures were accounted for by inclusion of both PHQ-9 and GAD-7 pre-treatment scores in both models. Each model was developed using two levels: clients at level 1 and PWPs at level 2. Due to the small number of services included in the study, it was not possible to test for a third level of service effects.

Multi-level models were developed in stages using Iterative Generalised Least Squares (IGLS) procedures, beginning with a single level regression, progressing to a random intercepts model and finally to a random slope model. At each development stage, improvements in the model were considered using chi squared distribution to test the significance of the difference between the -2*loglikelihoods. The MLM analysis was used in three ways: (1) to examine the amount of variance in the outcomes attributable to PWPs, controlling for pre-treatment scores, (2) to examine the shape of residual plots of PWP variation, and (3) to use this shape to determine quartiles of PWPs based on the rank of their residual plots.

2.5.2 Methods of analysing change on client measures

Client outcomes were also analysed using recovery rates, which were determined based on criteria set out by Clark et al. (2009) and adopted in IAPT services. Recovery is determined by the proportion of clients who meet criterion for “caseness” at pre-treatment; that is, a score of 10 or more on the PHQ-9 *and/or* 8 or more on the GAD-7. Those clients who meet the threshold for

caseness, then need to have a post-treatment score on the PHQ-9 of 9 or less *and* a GAD-7 score of 7 or less to be considered “recovered”.

This definition of recovery does not take into account the concept of reliable change, in which the pre-post change is required to exceed that which might be expected due to measurement error. The proportion of clients who reliably improved or reliably deteriorated was calculated using the Reliable Change Index (Jacobson & Truax, 1991). This was calculated separately for the PHQ-9 and the GAD-7. The formula used in the calculation was;

$$SD \quad \overline{2} \quad \overline{(1-r)}$$

The standard deviation was taken from the population analysed in the original evaluation of the IAPT demonstration site at Doncaster (Clark et al., 2009). Internal reliability estimate used for the PHQ-9 was α = .89 (Kroenke et al., 2001) and for GAD-7 was α = .92 (Spitzer et al., 2006). Clients were considered to have made reliable change if their scores had moved at post-treatment by at least 6 points on the PHQ-9 or by 4 points on the GAD-7. Reliable improvements were achieved if scores decreased by these margins and the criterion for reliable deterioration was met if scores increased by these amounts.

2.5.3 Questionnaire Data

PWPs were allocated into upper and lower effectiveness quartiles (n=5) based on the ranks developed from the MLM. This allowed for non-parametric statistical analysis of differences between the groups on their questionnaire scores. Statistical analysis was carried out using SPSS version19. Mann Whitney U tests were used to test for differences between the groups on self-

rated and supervisor rated questionnaires. A non-parametric test was selected due to the small number of participants in each group.

T-tests were used to examine differences between the quartiles using client outcomes, such as post-treatment scores and change scores. A parametric test was chosen for this analysis due to the large numbers of client data available to analyse. Accordingly the assumptions underpinning a parametric test were not violated.

Uncontrolled effect sizes were calculated for the upper and lower quartiles using Cohen's d. Using the standard criteria, uncontrolled effect sizes of .2 were considered small; .5 medium and .8 large. Although the original definitions were based on controlled effects sizes, with comparison of groups, the meaning of the size of *uncontrolled* effect sizes, (i.e., not compared to a group) remains the same.

2.5.4 Qualitative Analysis

Analysis of the PWP and supervisor interviews was conducted using Template Analysis (TA; King, 1998). In TA the researcher can define *a priori* codes that they expect to find in the data - the template - but modify these throughout the analysis as more codes emerge (King, 2004). This approach is widely used in health research (e.g., King, Thomas, & Bell, 2003). An advantage of this approach lies in it being a flexible approach that can be easily modified for different, specific areas of study (King, 2004).

A key feature of TA is hierarchical coding, with higher order codes overarching a cluster of lower order codes in a similar theme (King, 2004). In this project, *a priori*, high order themes were based on the interview schedules. Interviews were then analysed in two stages: (1) initial exploration of the data of all participants (2) examination of common lower order themes between upper and lower quartiles.

High order themes were used as a guide to examine emerging lower order themes of all 21 PWPs and 17 supervisors in the first stage of analysis. This is consistent with the "listing codes" procedure outlined by King (2004). The notion of "selectivity" was used in this process, identifying themes of central relevance to the research question (King, 2004). However, to minimise bias and facilitate openness, an independent researcher (a doctoral level student familiar with TA) also analysed 15 per cent of the interviews (n=6), as a form of quality control. They independently coded lower order themes from the template of high order themes, which were then compared with the lead researcher's codings.

The second stage of analysis allowed examination of differences between the groups of more and less effective PWPs. At this stage, some higher order themes were deleted or redefined according to the TA procedure (King, 2004). Themes were deleted if less than two PWPs or supervisors had described a similar lower order theme of relevance to the high order theme. Lower order themes that were identified by two or more PWPs or supervisors in their respective group (i.e., upper or lower quartile) were included as a final lower order theme. Quality control procedures were then implemented for this part of

the analysis with another independent rater who was also a doctoral level student familiar with qualitative analysis. They examined the high and lower order themes for the quartile groups to determine whether agreement was met regarding their appropriateness and fit.

3. Results

The results are presented in four specific phases as follows: (1) descriptive statistics provide a context of the data; (2) MLM analysis provides the test of therapist effects and yield groupings of more and less effective therapists; (3) inferential statistics employed to compare key variables between upper and lower quartiles of PWPs; (4) and qualitative results detail the overall comparisons of the more and less effective therapists based on their interview data.

3.1 Descriptive Data: Client, service, and individual practitioner levels

Descriptive results are presented to contextualise the PWP sample studied. Three levels of descriptive data are included; client, service, and PWP level.

3.1.1 Client-level outcome data

Table 1 presents the mean outcomes for the PHQ-9 and GAD-7 for all clients (n=1,122). The mean change scores (i.e., from pre- to post-therapy) on the PHQ-9 and GAD-7 were 3.34 and 3.05 respectively. The corresponding uncontrolled effect sizes, using the respective pre-treatment SD as the denominator, were 0.52 and 0.55. Recovery rates for the overall sample were calculated at 35.4 per cent.

Table 1:

Client outcomes scores at pre and post treatment, change scores and effect sizes

Outcome measure	Pre-treatment score mean (SD)	Post-treatment score mean (SD)	Change score mean (SD)	Uncontrolled effect size
PHQ-9	13.17 (6.43)	9.83 (7.15)	3.34 (6.43)	0.52
GAD-7	12.04 (5.57)	8.99 (6.32)	3.05 (5.82)	0.55

3.1.2 Service-level outcome data

The service level descriptive data is reported in Table 2. One service was excluded from this analysis in order to protect the anonymity of the single PWP employed by this service. Table 2 reports the means and SDs for pre-treatment, post-treatment, change scores, and the resultant uncontrolled effect sizes for each service. The data demonstrates a range in terms of each of these indices. One service was found to have lower pre-treatment scores on the PHQ-9 and GAD-7 than the rest of the services (9.51 and 9.59 respectively for service 4). Service 2 was also found to have low effect sizes compared to other services, with .44 and .46 on the PHQ-9 and GAD-7 respectively.

Table 2: *Change Scores, Effect Sizes and Recovery Rates by Service*

	PHQ-9				GAD-7				PHQ-9 & GAD-7 combined	
Service	*M (SD)*		Change score *M (SD)*	Uncontrolled effect size *M*	*M (SD)*		Change score *M (SD)*	Uncontrolled effect size *M*	Caseness % at intake*	Recovery %**
	Pre-treatment	Post-treatment			Pre-treatment	Post-treatment				
1	12.82 (6.46)	7.01 (6.22)	5.81 (6.52)	0.90	12.15 (5.34)	6.75 (5.44)	5.39 (5.92)	1.01	83.16	51.90
2	13.66 (6.23)	10.89 (7.24)	2.77 (5.42)	0.44	12.62 (5.60)	10.02 (6.58)	2.60 (5.18)	0.46	86.12	29.67
3	14.60 (5.75)	10.10 (6.84)	4.50 (6.17)	0.78	13.08 (4.90)	9.00 (5.88)	4.08 (5.59)	0.83	90.27	35.29
4	9.81 (6.83)	5.81 (6.89)	4.00 (4.80)	0.59	9.59 (5.64)	5.44 (5.92)	4.12 (4.48)	0.73	61.54	53.12
5	13.60 (6.25)	8.75 (6.68)	4.84 (5.97)	0.77	12.14 (5.44)	7.87 (5.79)	4.28 (5.51)	0.78	83.00	43.33

* To meet “caseness” at pre-treatment, clients must have a score of 10 or more on the PHQ-9 ***and/or*** 8 or more on the GAD-7 – that is a ‘case’ score above the threshold on ***either*** measure.
** Clients are considered “recovered” if they met the threshold for “caseness” at pre-treatment, and then have a post-treatment score on the PHQ-9 of 9 or less ***and*** a GAD-7 score of 7 or less – that is scores on ***both*** measures must be met for recovery.

3.1.3 PWP level descriptive data

Tables 3 and 4 present breakdowns of client data by individual PWPs. Table 3 focuses on descriptive information of the number of clients per PWP and sessions delivered. Table 4 documents measures of change including pre-post change scores, uncontrolled effect sizes, recovery rates and reliable change. These tables illustrate the range of scores across the PWP sample.

Table 3 also shows that the number of clients on each PWP's caseload ranged between 8 and 197. The outcome of Spearman's correlation showed that there was a non-significant relationship between PWP caseload and PHQ-9 change score (r = -.194, p = .399) and between caseload and PHQ-9 post-treatment score (r = .*318,* p = .160).

Table 3: *Individual PWP caseload variables*

PWP Id. No.	Clients per PWP N*	Number of sessions *M*	PHQ-9 *M (SD)*		GAD-7 *M (SD)*	
			Pre-score	Post-score	Pre-score	Post-score
1	17	4.65	15.94 (5.68)	9.53 (5.86)	14.59 (4.87)	8.88 (6.09)
2	39	3.77	13.77 (6.05)	10.90 (7.61)	13.33 (5.11)	10.38 (6.34)
3	15	3.73	18.00 (4.21)	12.67 (6.01)	14.27 (4.65)	9.20 (4.59)
4	19	4.37	12.42 (6.14)	8.42 (7.06)	10.79 (5.22)	7.79 (6.54)
5	23	3.57	14.61 (4.96)	8.87 (6.25)	12.65 (4.00)	7.61 (4.93)
6	64	4.95	11.27 (6.11)	6.75 (6.07)	11.16 (5.09)	6.63 (5.48)
7	46	5.15	13.28 (5.88)	8.37 (6.66)	12.80 (5.21)	7.65 (5.91)
8	197	5.58	12.23 (6.85)	13.15 (6.91)	10.91 (5.95)	11.87 (6.16)
9	15	6.60	10.93 (8.04)	7.13 (8.55)	10.73 (5.71)	6.47 (7.24)
10	12	5.75	9.58 (5.63)	4.17 (5.77)	8.83 (5.36)	3.58 (5.02)
11	17	4.47	10.35 (6.33)	5.76 (5.87)	10.53 (5.93)	5.82 (5.65)
12	8	3.75	6.88 (7.43)	5.88 (7.77)	6.38 (4.84)	5.50 (5.45)
13	80	4.28	13.79 (6.88)	6.43 (5.79)	12.56 (5.56)	6.34 (5.14)
14	54	4.74	10.78 (5.80)	6.31 (5.85)	10.19 (5.48)	6.02 (5.19)
15	194	2.70	14.60 (6.13)	12.28 (7.11)	13.58 (5.16)	11.39 (6.22)
16	69	3.65	13.26 (6.16)	10.57 (7.15)	11.83 (6.24)	9.28 (7.18)
17	60	6.50	13.53 (5.10)	8.17 (5.65)	12.60 (4.63)	7.17 (5.03)
18	37	4.78	12.57 (7.03)	7.84 (7.30)	10.11 (4.90)	6.68 (6.16)
19	35	7.06	15.91 (6.29)	10.40 (7.50)	14.26 (5.74)	9.97 (5.57)
20	49	4.88	13.80 (6.41)	10.69 (7.33)	13.31 (5.76)	9.88 (6.67)
21	72	4.58	12.92 (6.44)	7.60 (5.96)	10.99 (5.47)	6.67 (5.13)

*The timeframe for each practitioner to have individual client contact varied as a function of their PWP training course start dates (ranging from September 2009 to April 2010). Additional factors influencing caseloads included individual service level policies on when client contact could begin, differences in service referral rates and PWP employment status/leave. For further details, please see the discussion section.

Table 4: *Mean (SD) Client Outcome Score, Effect Sizes and Recovery Rates by PWP*

	PHQ-9				GAD-7				PHQ-9 & GAD-7	
PWP Id. No.	Change score *M (SD)*	Uncontrolled Effect size	Reliable improvement (%)	Reliable deterioration (%)	Change score *M* (*SD*)	Uncontrolled effect size	Reliable improvement (%)	Reliable deterioration (%)	Caseness rate (%)	Recovery rate (%)
1	6.41 (6.06)	1.13	52.94	5.88	5.71 (5.79)	1.17	58.82	0.00	100.00	47.06
2	2.87 (6.56)	0.47	23.08	10.26	2.95 (5.80)	0.58	41.03	10.26	89.74	28.75
3	5.33 (6.03)	1.27	46.67	0.00	5.07 (4.43)	1.09	66.66	6.67	100.00	33.33
4	4.00 (6.47)	0.65	47.37	5.26	3.00 (5.30)	0.57	57.89	15.79	68.42	30.77
5	5.74 (5.10)	1.16	39.13	0.00	5.04 (5.84)	1.26	47.83	0.00	95.65	40.90
6	4.52 (6.51)	0.74	42.19	9.38	4.53 (5.70)	0.89	59.38	6.25	81.25	50.00
7	4.91 (5.50)	0.84	39.13	4.35	5.15 (4.99)	0.99	65.22	0.00	82.61	44.74
8	-.91 (6.82)	-0.13	11.68	24.37	-.96 (5.49)	-0.16	14.72	20.81	73.64	13.10
9	3.80 (5.87)	0.47	20.00	0.00	4.27 (4.88)	0.75	53.33	6.67	66.67	40.00
10	5.42 (5.11)	0.96	33.33	0.00	5.25 (4.31)	0.98	66.67	0.00	66.67	75.00

	PHQ-9				GAD-7				PHQ-9 & GAD-7	
PWP ID No.	Change score *M (SD)*	Uncontrolled Effect size	Reliable improvement (%)	Reliable deterioration (%)	Change score *M* (*SD*)	Uncontrolled effect size	Reliable improvement (%)	Reliable deterioration (%)	Caseness rate (%)	Recovery rate (%)
11	4.59 (4.26	0.73	41.18	0.00	4.71 (4.82)	0.79	58.82	0.00	70.59	58.33
12	1.00 (1.41)	0.13	0.00	0.00	0.88 (1.13)	0.18	0.00	0.00	25.00	0.00
13	7.36 (6.82)	1.07	53.75	1.25	6.23 (6.51)	1.12	62.50	5.00	85.00	57.35
14	4.46 (6.53)	0.77	40.74	7.41	4.17 (6.19)	0.76	51.85	5.56	70.37	68.42
15	2.32 (4.74)	0.38	21.13	3.61	2.19 (4.53)	0.42	27.84	3.09	91.24	19.20
16	2.70 (6.02)	0.98	27.54	4.35	2.55 (5.84)	0.41	33.33	5.70	84.58	36.21
17	5.37 (5.27)	1.05	48.53	28.33	5.43 (5.31)	1.17	63.33	3.33	90.00	50.00
18	4.73 (6.55)	0.67	40.54	5.41	3.43 (5.19)	0.7	45.95	5.41	67.57	48.00
19	5.51 (7.46)	0.88	48.57	2.86	4.29 (5.67)	0.75	57.14	8.57	91.43	34.38
20	3.10(5.31)	0.48	24.49	4.08	3.43 (6.18)	0.60	38.78	4.08	87.76	37.21
21	5.32 (5.75)	0.83	43.06	4.17	4.32 (5.23)	0.77	52.78	2.78	77.78	44.64

3.2 Multi-level modelling

This section details the procedures and analysis used in order to determine whether a therapist effect was present in the sample. Separate multi-level models were derived for the PHQ-9 and GAD-7 in order to compare the results.

The initial single level regression analysis allowed examination of the relationship between the pre-treatment scores and post-treatment scores. The estimated regression lines were:

$$Postscore = \beta_0 + \beta_1(\text{prescore}) + e^i$$

Formula 1: $PHQ9\ Postscore = 9.832 + (0.618 \times PHQ9prescore) + e^i$

Formula 2: $GAD7\ Postscore = 8.990 + (0.596 \times GAD7prescore) + e^i$

This indicated a positive relationship between the pre- and post-treatment scores on both measures, with clients obtaining higher scores at pre-treatment also having higher scores at post-treatment. Pre-treatment scores of the alternative measure were then included plus the interaction between the measures. For example, in the model for the PHQ-9, the GAD-7 pre-treatment scores were included as variables plus the interaction between the PHQ-9 and GAD-7 pre-treatment to determine whether this had an effect on the PHQ-9 post-treatment score as the model progressed to a multi-level stage.

A multi-level model was developed, with the inclusion of PWPs at level two and allowing individual PWPs regression lines and intercepts to vary, but keeping a common slope. Using the likelihoods ratio test to estimate the between PWP variation in the intercepts (i.e., comparing this to the results of the single level regression) showed that the difference was significant for both measures: PHQ-

9 $\chi^2(1) = 139.923$, $p < .001$; GAD-7 $\chi^2(1) = 140.434$, $p < .001$. These results indicate that the random intercept model was a better fit for the results than the single level regression. This result suggests that there is significant variability between PWPs even after adjusting for clients' pre-treatment scores.

The next stage of the model, the random slopes model, built on the random intercept model by allowing the slopes to vary between PWPs. The results for the final PHQ-9 model (model 1) and for the GAD-7 (model 2) are presented below.

Model 1, relating to the PHQ-9 scores, is shown below, with full details presented in Appendix XXI. Model 1 shows that the intercepts of the individual PWP lines are varied, with a mean of 8.664 (*SE* 0.437) and a variance of 2.779 (*SE* 1.131). The coefficient of the PHQ-9 average slope is estimated at 0.549 (*SE* 0.046) and individual PWP slopes vary about this mean with an estimated variance of 0.013 (*SE* 0.009). There is a positive covariance between intercepts and slopes estimated as +0.074 (*SE* 0.074) indicating that larger intercepts tend to have steeper slopes. That is, the regression lines for the PWPs 'fan out'. However, the large standard error indicates a large degree of uncertainty. The loglikelihood test for model 1 was significant ($\chi^2(2) = 13.725$, $p < .05$), thereby indicating an improvement from the random intercept model.

Model 1:

$$\text{PHQlast}_{ij} = \beta_{0j} + \beta_{1j}(\text{PHQfirst-gm})_{ij} + 0.117(0.041)(\text{GADfirst-gm})_{ij} + 0.012(0.005)(\text{PHQfirst-gm}).(\text{GADfirst-gm})_{ij} + e_{ij}$$

$$\beta_{0j} = 8.664(0.437) + u_{0j}$$

$$\beta_{1j} = 0.549(0.046) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 2.779(1.131) & \\ 0.074(0.074) & 0.013(0.009) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0, \sigma_e^2) \quad \sigma_e^2 = 29.122(1.250)$$

$$-2*loglikelihood = 7010.868(1122 \text{ of } 1122 \text{ cases in use})$$

Model 2, relating to the GAD-7 scores, is shown below and full details can be found in Appendix XXII. Model 2 shows that the intercepts of the individual PWP lines were varied, with a mean 7.805 (*SE* 0.393) and variance of 2.262 (*SE* 0.917). The coefficient of the PHQ-9 average slope is estimated at 0.478 (*SE* 0.048) and individual PWP slopes varied about this mean with an estimated variance of 0.015 (*SE* 0.010). There is a positive covariance between intercepts and slopes estimated as +0.112 (*SE* 0.073). The 'fanning out' of the PWP regression lines appears more certain for the GAD-7 than for the PHQ-9. The loglikelihood test for model 2 was significant ($\chi^2(2) = 16.337$, $p < .001$), indicating a significant improvement on the previous models.

Model 2:

$$\text{GADlast}_{ij} = \beta_{0j} + \beta_{1j}(\text{GADfirst-gm})_{ij} + 0.131(0.032)(\text{PHQfirst-gm})_{ij} + 0.014(0.004)(\text{GADfirst-gm}).(\text{PHQfirst-gm})_{ij} + e_{ij}$$

$$\beta_{0j} = 7.805(0.393) + u_{0j}$$

$$\beta_{1j} = 0.478(0.048) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 2.262(0.917) & \\ 0.112(0.073) & 0.015(0.010) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0, \sigma_e^2) \quad \sigma_e^2 = 23.436(1.006)$$

$$-2*loglikelihood = 6765.856(1122 \text{ of } 1122 \text{ cases in use})$$

3.2.1 Intraclass Correlation

The extent of similarity between individuals in the same group is indexed by the intra-class correlation (ICC). The ICC measures the proportion of the total residual variation due to differences between groups. In the present study, the ICC measures the extent to which the variance in PHQ-9 and GAD-7 post-treatment scores, when accounting for pre-treatment scores, is attributable to PWPs. It is represented by the formula:

$$\frac{\sigma^2 therapist}{\sigma^2 therapist + \sigma^2 error} \times 100$$

Formula 3: $PHQ9$: $\frac{2.779}{2.779 + 29.122} \times 100 = 8.71$

Formula 4: $GAD7$: $\frac{2.262}{2.262 + 23.436} \times 100 = 8.80$

Formula 3 yields an ICC of 8.71 for the PHQ-9, indicating that approximately 8.7% of the total variance of post-treatment scores is attributed to differences between PWPs. Similarly, formula 4 for the GAD-7 yields an ICC of 8.80, which indicates that approximately 8.8% of the variance in post-treatment scores on this measure is attributed to differences between PWPs in terms of their clinical effectiveness.

3.3 Defining PWP ranks from multi-level modelling

In order to make comparisons between the overall outcomes of PWPs, the residuals of individual PWPs were used. The residuals represent how each PWP departs from the overall outcome mean for all PWPs, which in turn allows examination of the shape of the overall distribution of the residuals. PWPs were then ranked based on the mean value of the residual scores.

Figure 1: Residual plots from model 1

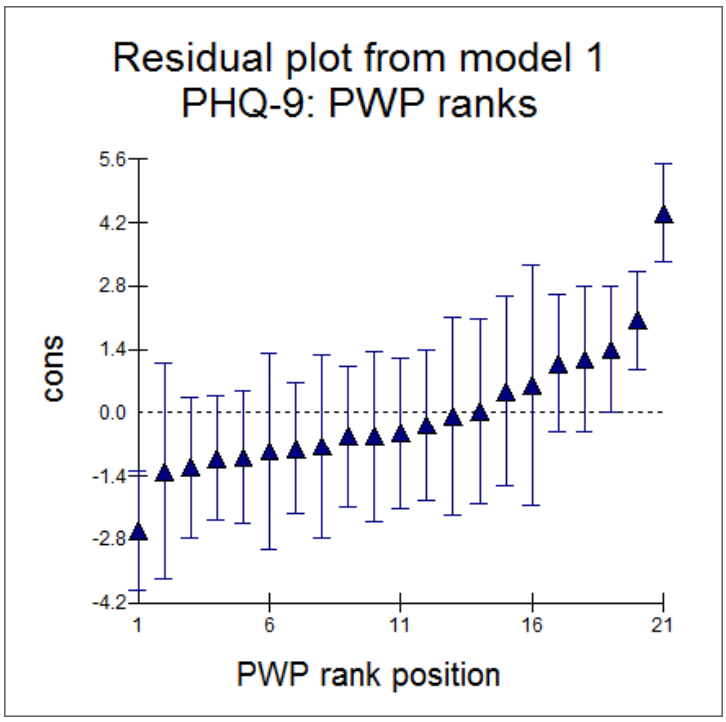
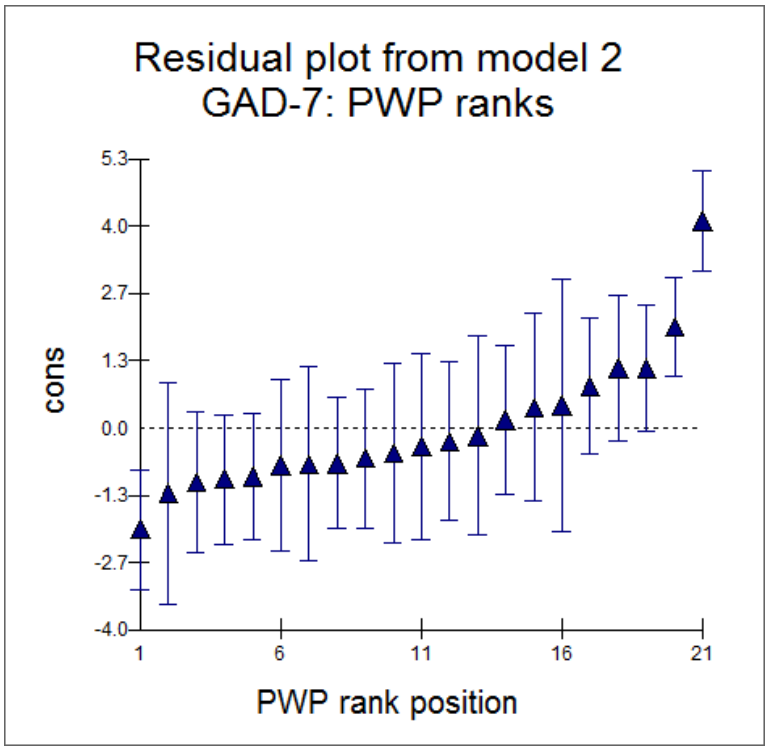


Figure 2: Residual plots from model 2



Figures 1 and 2 present the residuals of the 21 PWPs plotted for the PHQ-9 and GAD-7 respectively. The ranks are shown across the x axis with the PWP ranked the most effective (rank 1) at the bottom left, ranging across to the PWP ranked least effective (rank 21) at the top right. Although the confidence intervals are large, the data indicate that the confidence intervals for the mean values for three PWPs did not cross zero: PWPs with ranked numbers 1, 20 and 21. These non-overlapping confidence intervals suggest that these PWPs yielded outcomes that were significantly different from the average outcomes at the 5% level. The remaining PWPs mean values all had confidence intervals that overlapped zero, indicating that they fell within the mean range.

Table 5 illustrates where each PWP was ranked based on PHQ-9 and GAD-7 outcomes. The rankings were broadly similar but not exactly identical for the

two measures, with a significant concordance rate measured by Kendall's Tau of r = .86, p<.001. A composite of these ranks was created using the mean of the two rank positions and is shown in Table 5. Results of rankings based on recovery rates are also reported and concordance rates with the composite PHQ-9/GAD-7 ranks were also significant (r = .71, p<.001). Unlike MLM rankings, recovery rates do not take into account the whole sample of clients that PWPs worked with (i.e., to meet recovery clients had to meet caseness at pre-treatment). Accordingly, MLM rankings of individual PWPs were used in preference to recovery ranking in subsequent analyses.

The shaded sections of Table 5 represent the upper (n=5) and lower (n=5) quartiles of PWPs, determined from their PHQ-9/GAD-7 rank position. For clarity in the remainder of the report, these PWPs will be referred to by their rank position on this composite scale (e.g., PWP number 13 will be referred to as PWP-1 as this is their rank position).

Anonymised service codes for PWPs in the two quartiles are shown in Table 5 Codes are not shown for the middle group of PWPs in order to protect the identity of the single PWP practitioner from that service. The service codes illustrate that there were PWPs from a mix of services across quartiles and that two services had PWPs in both the upper and lower quartiles.

Table 5:

PWP Ranks of PHQ-9 and GAD-7 Scores

	PWP Identification numbers				
Rank	PHQ-9	GAD-7	Composite PHQ-9 & GAD-7	Service code (based on composite)	Recovery rate
1	13	13	13	A	10
2	10	10	10	B	14
3	14	14	14	C	11
4	21	17	21	D	13
5	6	21	17	D	6
6	11	5	6		17
7	17	11	11		18
8	1	6	5		1
9	7	7	1		7
10	5	1	7		21
11	18	3	18		5
12	19	18	19		9
13	9	9	3		20
14	4	19	9		16
15	3	4	4		19
16	12	12	12		3
17	20	20	20	D	4
18	2	16	2	E	2
19	16	2	16	C	15
20	15	15	15	C	8
21	8	8	8	F	12

3.4 Comparison of upper and lower quartiles of PWPs

The PWPs populating the upper and lower quartiles were compared using data from: (1) client outcome scores on the PHQ-9 and GAD-7, (2) self-rated and supervisor-rated measures of intuition, ego strength, and resilience, and (3) qualitative interview data analysed using template analysis. The results are presented below.

3.4.1 Comparison of quartiles based on client outcome scores

Table 6 shows that scores on both the PHQ-9 and GAD-7 were significantly lower at post-treatment for the upper quartile than the lower quartile (PHQ-9, $t(662) = 11.08$, $p < .001$, $d = .79$; GAD-7, $t(675) = 11.33$, $p < .001$, $d = .81$). Levene's test indicated unequal variances in both cases (PHQ-9, $F = 25.527$, $p < .001$; GAD-7, $F = 45.203$, $p < .001$) and the degrees of freedom were adjusted from 824 to 662 on the PHQ-9 and to 675 on the GAD-7. The change scores of PWPs in the upper quartile were significantly higher than those in lower quartile (PHQ-9,$t(553) = 9.81$, $p<.001$, $d = .72$; GAD-7 $t(535) = 9.13$, $p<.001$, $d = .68$). Again, Levene's test indicated unequal variances in both cases (PHQ-9, $F = 4.54$, $p = .033$; GAD-7, $F = 7.93$, $p = .005$) and the degrees of freedom were adjusted from 824 to 552 on the PHQ-9 and 534 on the GAD-7. PWPs in the upper quartile saw clients for significantly more sessions than those in the lower quartile ($t(824) = 4.06$, $p<.001$ $d = .30$).

As PWP-21 appeared to be a significant outlier based on visual inspection of Figures 1 and 2, t-tests were conducted again excluding this PWP to test the effect of PWP-21. Excluding these data did not change the significance of the results.

Table 6

Client Outcome Scores by PWP Quartiles

	PHQ-9 *M (SD)*			GAD-7 *M (SD)*			
Quartile	Post-score	Change score	Uncontrolled effect Size	Post- score	Change score	Uncontrolled effect size	Number of sessions
Upper	6.99 (5.85)	5.76 (6.17)	0.92	6.42 (5.13)	5.12 (5.82)	0.95	4.99 (2.85)
Lower	13.45 (6.48)	1.31 (6.12)	0.20	11.09 (6.42)	1.27 (5.56)	0.22	4.13 (2.90)
T-test	11.08**	9.81**		11.33**	9.126**		4.06**

**p<0.001

3.4.2 Comparison of quartiles based on PWP measures

The measures completed by PWPs (i.e., resilience, ego strength, and intuition) and by supervisors (i.e., intuition) were analysed using the Mann Whitney U test. Table 7 shows that resilience scores were significantly higher for PWPs in the upper quartile than in the lower quartile (U = 2.000, N_1 = 5, N_2 = 5, p = .03). Figure 3 shows the difference between the groups on resilience scores. All other comparisons between the groups of PWPs were non-significant. Supervisors differed significantly on their ratings of PWPs in the groups on experiential intuition (U = .500, N_1 = 5, N_2 = 5, p = .02). Differences between supervisors on the rational intuition scale were not significant. Again, these results were repeated excluding the PWP ranked 21 to determine whether this outlier had an effect on the results. This analysis did not change the significance of the results.

Figure 3: Box plots of PWP quartiles and resilience (CD-RISC) scores

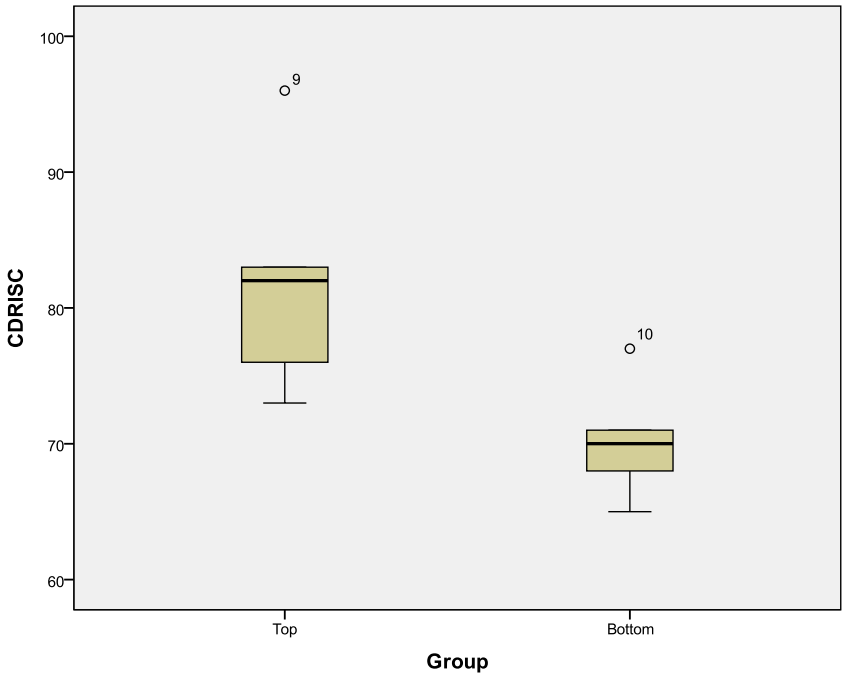


Table 7

PWP Age and Scores of Resilience, Ego Strength and Intuition by Quartiles

Quartile	PWP Age M (SD)	CD-RISC Resilience Score M (SD)	PIES Ego Strength Score M (SD)	REI Intuition Score M (SD): PWP rated		REI Intuition Score M (SD): Supervisor rated	
				Rational	Experiential	Rational	Experiential
Upper	33.00 (8.83)	82.00 (8.86)	274.80 (9.99)	64.00 (2.35)	58.40 (5.18)	62.40 (2.88)	50.80 (4.82)
Lower	25.60 (3.13)	70.20 (4.44)	268.00 (9.62)	66.00 (2.00)	57.20 (3.11)	61.25 (2.22)	59.75 (4.99)
Mann Whitney U	7.000	2.000*	8.000	4.000	12.000	7.500	0.500*

* $p<0.05$

3.5 Overview of quantitative results

Multi-level modelling showed that almost 9% of the variance in outcome was attributable to therapists. The upper quartile of PWPs had significantly higher self-rated levels of resilience than the lower quartile of PWPs and supervisors of the lower quartile of PWPs rated their supervisees as higher in levels of experiential intuition. The following section reports the results of qualitative methodology used to explore the differences between the most effective and less effective groups of PWPs further.

3.6 Qualitative Results

The results presented here are from the second stage of analysis and compare the more effective and less effective PWP groups. The focus on these two groups is aimed at highlighting the differences between more and less effective practitioners. Full templates of lower order themes are not included in the report due to space constraints but examples are presented in Appendix XXIII and full templates are available from the author.

High and lower order themes in the upper and lower quartiles for the PWPs and supervisors are mapped out in Figures 4 and 5 respectively. High order themes are presented on the left of the figures, with lower order themes of participants in the upper and lower quartiles on the right. The text in bold print relates to the lower order themes that emerged uniquely for either the upper or lower quartiles. All other lower order themes presented emerged in both quartiles of participants. Only differences between the groups are discussed further. Two high order themes were deleted due to lack of subthemes emerging amongst participants in the two groups: “how previous experience hindered” (PWP

question only) and “how CPD has influenced PWP practice” (supervisor and PWP question).

Quality control procedures for the first stage of analysis resulted in a 78% agreement between the two raters for the lower order themes, which was considered acceptable. The second stage of quality control resulted in two changes to the labelling of lower order themes. The initial label of “deeper processing” was changed to “process supervision” in the upper quartile “engaging in supervision to improve practice” high order theme. Additionally under the high order theme of “gaps in skills and knowledge”, the lower order theme of “specific types of intervention” for PWPs in the less effective quartile was changed to “specific types of presentation” as this fitted more accurately with the quotes from PWPs.

Illustrative quotes are used to highlight the relevant themes and are labelled by PWP letter and unique quote number. PWPs in the upper quartile were coded A - E, with PWPs in the lower quartile coded V - Z. For anonymity purposes, PWP letters do not relate directly to rank order and all quotes have been changed to refer to female PWPs, even if the PWPs interviewed were male.

Figure 4: High order and lower order sub-themes for PWP top and bottom quartiles

High order themes	*Lower order theme: PWP Upper Quartile*	*Lower order theme: PWP Lower Quartile*
How previous experience helped current practice	• Experience of talking to people in distress • **Knowledge of CBT principles**	• Experience of talking to people in distress • **Development of interpersonal skills**
Engagement with supervision to improve skills	• Checking decisions • **Being prepared and organised** • **Utilising supervisor's knowledge** • **Process supervision**	• Checking decisions • **Specific clinical questions**
Methods of improving practice	• Reading • Discussions with colleagues • **Observing others**	• Reading • Discussions with colleagues
Awareness of gaps in skills or knowledge	• **Knowledge of medication** • **Gaps go beyond remit of low intensity**	• **Specific skills** • **Knowledge of specific presentations**
Hallmarks of clinical practice	• Engage and relate well with others • Core clinical skills • **Communication skills** • **Adapting interventions to the individual**	• Engage and relate well with others • Core clinical skills
Effectiveness in working in the stepped care model	• **Understanding IAPT model and PWP role** • **Knowledge of when to step up**	• **Communication**

Figure 5: High order and lower order sub-themes for supervisors of PWPs in the top and bottom quartiles

Higher order themes	*Lower order themes: Supervisor Upper Quartile*	*Lower order themes: Supervisor Lower Quartile*
Engagement with supervision to improve skills	• Engages well • Good organisational skills • **Openness to discussing difficulties** • **Active supervision participant**	• Engages well • Good organisational skills • **Openness**
Methods of improving practice	• Observation • **Proactive in improving practice** • **Online research**	• Observation
Hallmarks of clinical practice	• Clinical skills • Knowledge and understanding • **Organisational skills**	• Clinical skills • Knowledge and understanding • **Interpersonal skills** • **Openness**
Effective components of therapeutic delivery	• Interpersonal skills • Clinical skills • **Knowledge and understanding**	• Interpersonal skills • Clinical skills

3.6.1 Upper quartile unique responses

Figure 4 shows that a total of 11 unique lower order themes were reported by PWPs in the upper quartile. Supervisor responses elicited 6 different unique lower order themes.

The more effective group of PWPs identified gaps in their skills and knowledge being limited to knowledge of medications. Reports indicated that they believed other gaps would go outside of the remit of PWP interventions:

"I'm aware that there are things that I'm not sure about but I think they're more high intensity work that needs doing" (PWP D:1)

This lack of gaps, other than medication, suggests that these PWPs were confident in their abilities to fulfil the PWP role. Previous experience with CBT appeared to help develop this confidence in their current role.

"I'd already done some face to face CBT work with less complex cases....so intervention wise I felt quite comfortable delivering interventions which are using a cognitive behavioural approach" (PWP B:2)

Confidence was also reflected in the PWPs description of their ability to adapt interventions to the individual. The flexibility in providing interventions suggests that PWPs had a good basic understanding of these treatments and enough confidence in delivering them to allow them to be flexible.

Having a good understanding of the IAPT model and PWP role was a factor that this group of PWPs said helped them to operate effectively in the stepped care model.

"I think knowing my own limitations, I think that's really important to the stepped care model that you stick to the clear steps" (PWP A:3)

Supervisors reported that a good knowledge and understanding was an effective component of their supervisee's therapeutic delivery. Additionally, supervisors reflected that PWPs were open to discussing any difficulties they had in their work:

"She's very happy to bring along examples of things that are going well, things that are going less well" (Supervisor A:4)

This suggests that effective PWPs felt comfortable in highlighting and discussing areas of difficulty.

This group of PWPs and their supervisors reported that PWPs used proactive means of developing their skills. This was reflected in a number of ways. Supervisors described their supervisees as being proactive in finding methods to improve their practice and used online research. PWPs described using observation of others in clinical practice to improve their own practice. Within supervision sessions, supervisors reported that these PWPs were active participants. This proactive stance was also reflected in PWPs reports of their engagement in supervision, describing how they get the most out of their

supervisors by “utilising their supervisors’ skills” and also engaging in supervision at a deeper, process level.

“With clinical supervision I try to get the most out of it through thinking of different ways we can use it, like by having case discussions, case presentations....role plays and things” (PWP A:5)

Quote 5 illustrates the meaning of this theme as using supervision in a creative way to develop skills at a deeper level than simply using it to check out clinical decisions. PWPs also reported being proactive for supervision by ensuring that they were prepared and organised.

“I usually [pull together information on] who I’m currently working with, recently assessed, recently discharged [other information including outcome scores] and present a copy of that every case management supervision to my supervisor and highlight which ones I’d like to discuss” (PWP B:6)

Quote 6 demonstrates the high level of organisation and preparation that this group reported making for their supervision sessions. Good organisational skills of these PWPs were also reported by their supervisors in relation to the high order theme of “hallmarks of clinical practice.” This illustrates that both PWPs themselves and supervisors felt that this group of more effective PWPs were very organised.

In addition to being organised and prepared, PWPs also discussed being thorough in their approach with clients, ensuring clarification in their communication:

"I do make efforts to be explicit with clients about exactly why it is that I'm talking to them about doing certain things, the rationale for it and how it's going to help them" (PWP D:7)

Quote 7 illustrates the lower order theme of communication, where effective PWPs reported taking time to thoroughly explain the interventions they were using with clients.

3.6.2 Lower quartile unique responses

Figure 5 shows that five unique lower order themes emerged from PWPs in the lower quartile. Three unique lower order themes emerged from their supervisors.

In contrast to PWPs in the upper quartile, this group of PWPs lower order themes appeared to reflect less confidence in their skills and abilities. For example, in terms of gaps in their skills and knowledge, the lower quartile group of PWPs felt that they needed further development in specific skills:

"There's a lot of emphasis on behavioural activation but I don't feel like I have very good skills in delivering that" (PWP W:8)

The reports of their approach to supervision also reflected less confidence, using supervision to ask specific clinical questions:

> *"I've got loads of guidance on who I shouldn't be seeing and who needs stepping up, and about things about the disorders we didn't look at uni"* (PWP Y:9)

PWPs in the lower quartile reported that the main way to be effective in the stepped care model was through communication.

> *"I make my best efforts to introduce myself to the GP so that I'm not just a name, so that they know I'm a presence and that I'm there to support their patients and try and open communication a little bit"* (PWP Z:10)

This account differs from PWPs in the top quartile as it does not reflect any specific factor associated with IAPT or the PWP role. In quote 10, the focus is on communication and not on communication specifically about step two interventions or about the IAPT model generally.

Openness was a quality that supervisors of PWPs in the lower quartile reported across two higher order themes for example *"If she's not sure of something she will ask"* (PWP V:11). However, this differs from openness reported by supervisors of upper quartile PWPs as it does not reflect an openness to difficulties.

3.6.3 Overview of qualitative results

The results from the TA analysis show lower order themes from PWPs in the upper quartile appeared to be related to these PWPs being confident in their interventions, feeling that they have few gaps in their skills and knowledge and taking a proactive, organised and thorough approach to supervision and their clinical work. In contrast, PWPs in the lower quartile appeared less confident in their skills and abilities, with more focus on gaps of specific skills and knowledge of different presentations.

4. Discussion

4.1 Discussion Overview

The aim of the present research was twofold: first, to test for the presence of therapist effects in a sample of psychological wellbeing practitioners (PWPs) across a variety of IAPT services, and second, to establish key factors that were characteristic of and differentiated between more effective and less effective PWPs. Results indicated that therapist effects accounted for approximately 9% of the variance in client outcomes and that more effective PWPs reported significantly higher levels of resilience than less effective PWPs. More effective PWPs also appeared to be more confident in their skills and took a more proactive, organised and thorough approach to their work than less effective PWPs. These results will be discussed in relation to the two key aims of the research: first, establishing the extent to which therapist effects are present in this PWP sample and, second, identifying the key components associated with more and less effective therapists. In addition, the wider implications for research and practice in the delivery of effective psychological therapies will be considered.

4.2 Therapist effects

The outcomes of this study can be seen as broadly representative of PWPs more generally. The present sample comprised 21 PWPs and although small, the overall recovery rates are similar to those of a larger sample of PWPs in the one year audit of national IAPT roll-out sites (Glover et al., 2010). When focussing specifically on the outcomes of low intensity interventions, Glover et al. (2010) reported recovery rates of 37.5 per cent, which is similar to the recovery rate of 35.4 per cent found in the present study.

Notwithstanding the limited experience of the sample of PWPs, all were delivering NICE-approved cognitive-behavioural interventions. Although adherence to the CBT interventions was not specifically assessed, all PWPs had been specifically trained in their standardised interventions and should have been receiving standard supervision within the IAPT services as recommended by Turpin and Wheeler (2011). In this context, the first aim of the research was to determine – given the standardisation of intervention approach – whether and to what extent therapist effects were present.

Utilising the most appropriate statistical analysis for nested data yielded a therapist effect of almost 9 per cent. This finding is comparable to previous studies of therapist effects in routine outpatient settings. Where the literature has reported the presence of therapist effects, the percentage of variance attributable to therapists has ranged from 5 – 8% (Lutz et al., 2007; Wampold & Brown, 2005). The current reported rate of approaching 9% is above the upper end of reported rates. This finding is important because the phenomenon of therapist effects has not been previously examined in the population of PWPs.

This finding of a therapist effect in a PWP sample challenges the notion that standardised, evidence based interventions are not affected by the provider of the intervention. NICE guidance is one such system that recommends specific interventions for particular diagnostic presentations, and the present findings suggest a need to place a greater emphasis of the role of the provider than is currently the case. The current finding adds to a growing body of literature suggesting that variation between practitioners is a feature of service delivery systems (e.g., Lutz et al., 2007 and Wampold & Brown, 2005). The variation found suggests that it is not only *what* intervention clients receive that has an effect on their progress, but also *who* provides the treatment.

The clinical implication of the finding of a therapist effect in a PWP sample is that some PWPs will be less effective than others. Research has shown that when provided with feedback, therapists' clinical outcomes can improve (Lambert, Whipple, Vermeersch, Smart, Hawkins, Nielsen & Goates, 2002). Clinicians and supervisors should therefore regularly review clinical outcomes in order to ensure that treatment is effective and if it is not, take action to remedy the situation (Kraus et al., 2011).

As this study utilised a small sample of PWPs, future studies could build on the current research by utilising a larger PWP sample to determine whether these findings are replicated. In total, 29 PWPs were recruited for this study. However, datasets for 8 PWPs could not be included due to technical difficulties within the IAPT services in terms of their data management, storage and retrieval systems. These difficulties are a challenge to the routine retrieval of data that is at the heart of the IAPT model and philosophy (CSIP Choice & Access Team,

2008). Subsequently, the sample utilised was smaller than the 30 practitioners recommended for use in MLM by Soldz (2006). However, this provided a naturalistic test to the 30 by 30 rule and found the presence of a therapist effect despite the smaller numbers. This finding suggests that using a sample smaller than 30 therapists can still yield reliable results. The results of this study were comparable to those of a study utilising a larger sample of therapists (Saxon & Barkham, submitted), in that the distribution of therapist residuals resulted in residual plots similar to the shape of those found in this smaller sample. The main difference between the two sets of findings is that the smaller sample of PWPs utilised resulted in larger confidence intervals of residuals than those found with a bigger sample of therapists. Notwithstanding the difference in confidence intervals, the similarities between the two studies suggest that a smaller sample of therapists can be reliably used in MLM analysis. One consequence may be that the guidance of a minimum of 30 therapists by 30 clients may need revision, as these findings suggest that MLM procedures appear robust down to approximately 20 therapists.

The number of clients on each PWP's caseload ranged from 8 to 197 clients per PWP. This range was influenced by a number of factors. The cohort of PWPs sampled were those who completed their training in 2010, however the start dates for the courses varied from September 2009 to April 2010 meaning that PWPs from these courses had differing amounts of opportunity to have individual client contact. Individual services also varied on their policy about how long PWPs had to be in the service before they could independently begin working with clients, as opposed to shadowing more experienced staff. The status of individual services also impacted on the number of clients seen, as

some services were well established with high referral rates, whilst others were newly developed meaning that they had lower rates of referral and less well established pathways. Services also differed in their expectations of the PWP role, with some PWPs being involved in more service development work and running groups, leaving them less time to work with clients on an individual basis. Individual PWPs also varied in their employment status (i.e. full time versus part time) and their sickness leave. Notwithstanding the differences in PWP caseload, variations were seen in their effectiveness, even when PWPs had seen a small number of clients.

Additionally, no relationship was found between number of clients on PWP caseload and change scores or post-treatment scores indicating that caseload was not related to effectiveness. However, the two PWPs ranked as least effective had the highest caseloads, which were more than double the amount of clients seen by the PWP with the next highest caseload. This may therefore suggest that extremely high caseloads are not advantageous for successful clinical outcomes. However, future studies with larger sample sizes would be needed to explore this relationship further.

4.3 Differences between more and less effective PWPs

Notwithstanding the small sample sizes of PWPs in the upper and lower quartiles, significant differences were found between the groups on a number of variables. Client outcome scores of the two groups confirmed the effectiveness ratings assigned from MLM rankings, as the more effective PWPs had superior outcomes to their less effective colleagues. Uncontrolled effect sizes produced by the more effective PWPs were large whereas those produced by the less

effective PWPs was small, and more effective PWPs had significantly higher change scores than their less effective peers, indicating higher levels of effectiveness in the more effective PWP group. The most effective PWPs also saw their clients on average for a longer amount of time. This is inconsistent with previous findings in the therapist effects literature, where effective therapists were also efficient in their use of time (Okiishi et al., 2003). The inconsistency with previous literature may reflect differences in the content and the overall short-term approach of PWP work.

When addressing the question of what factors are associated with effectiveness, resilience appeared to be the most important factor, with more effective PWPs reporting higher self-rated levels of resilience. The concept of resilience relates to the ability to cope with adversity or stress (Rutter, 1993). The PWP role can be seen to be stressful as it involves carrying and managing a high caseload of clients with associated risk issues at times, which creates a large amount of administrative work as well as a considerable volume of clinical work. Additionally, PWPs undertake an intensive 12-month training programme where they receive feedback on their clinical and academic skills. In this context, therefore, it is unsurprising that PWPs need to have a level of resilience to manage all this work and to do so effectively. It may be that PWPs who have higher levels of resilience are able to cope with and manage the stressful training process and PWP role more successfully than those who are low in resilience.

As previously stated, resilience is associated with intrapersonal and environmental characteristics (Tusai & Dyer, 2004). The use of an overall self-

report measure of resilience did not capture these concepts to examine whether such characteristics of resilience are associated with effective therapists. Future studies may benefit from exploring these factors in relation to effective therapists. Additionally, supervisor ratings of their PWPs resilience utilising validated, other-rated measures, could also be beneficial for future studies, to determine if supervisors corroborate the finding that more effective PWPs exhibit higher levels of resilience.

In this study, although supervisor accounts of PWP resilience was not quantitatively examined, supervisor's qualitative accounts of the more effective PWPs provided some indication that these PWPs appeared resilient as the PWPs were open to discussing difficulties in their work. Openness to difficulties was not a factor discussed by supervisors of less effective PWPs. This openness could suggest that the more effective PWPs were able to effectively handle feedback on cases that were not going well and therefore did not shy away from bringing these to supervision. In turn, discussing difficulties may have helped to develop skills in dealing with clinical challenges, thus helping to increase effectiveness.

Experiential intuition is defined as information processing that is "preconscious, rapid, automatic, holistic, primarily nonverbal and immediately associated with affect" (Pacini & Epstein, 1999, p. 972). The use of such a processing style was rated significantly higher by supervisors of less effective PWPs than those of the more effective PWPs. This rating suggests that less effective PWPs are seen by their supervisors to rely more on affect driven information processing than more effective PWPs. This finding is limited by the fact that the REI

questionnaire used to measure this construct is not validated for “other-rated” use. Accordingly, these results should be interpreted with caution. Despite this, the finding of a significant difference between the groups as rated by supervisors, warrants further exploration of the role of experiential intuition in the effectiveness of PWPs.

The use of less affect driven processing styles in their work was also reflected in the qualitative accounts of the more effective PWPs. These PWPs reported that they were more proactive, prepared and organised, which may have led supervisors to observe this group of PWPs as using less emotional processing and primarily being rational and analytic. More effective PWPs also reported an overall sense of being more confident in their skills and abilities. This confidence may be a result of the more effective PWPs being at a more advanced stage in their clinical development than less effective PWPs. For example, more effective PWPs reported a good understanding of the interventions, the IAPT model and PWP role, the ability to flexibly adapt their interventions and used supervision in more creative and process driven ways. These PWPs also reported the only gap in their skills was on knowledge of medication and that other gaps would go outside of their remit of low intensity interventions, for example they had gaps in their knowledge of detailed high intensity interventions.

In contrast, less effective PWPs reported gaps in their skills and knowledge of the standard interventions delivered by PWPs, for example in behavioural activation. This lack of confidence in their skills may have been a factor in their

using supervision to ask specific clinical questions, rather than the creative and use of process supervision reported by more effective PWPs.

The use of supervision in different ways by trainee therapists at different stages of their development has been reported in the supervision literature (Stoltenberg & McNeill, 1997). Developmental models of supervision suggest that trainee therapists utilise supervision in different ways depending on their developmental needs. Stoltenberg and McNeill (1997) define level 1 trainees as relying on supervisors to provide specific guidance as they learn and develop new skills. When trainees progress to level 2, they become more confident having developed the basic skills and become more open to discussing personal issues of self-awareness (Stolenberg & McNeill, 1997). Part of the findings of the present analysis may reflect these developmental stages, with PWPs in the lower quartile perhaps still being at level 1 in their development and the upper quartile of PWPs may have progressed to level 2.

Taking all the findings relating to more and less effective PWPs together suggest that more effective PWPs create greater change in their clients and use more sessions to do so. More effective PWPs are more resilient, appear to supervisors to use less experiential intuition in their work, are proactive and organised, more confident and perhaps are at a more advanced developmental stage in relation to their clinical work than less effective PWPs.

4.4 Methodological Strengths and Limitations

The main strength of this study was the use of methodological pluralism, utilising a combination of client outcome measures, PWP self-report measures,

a supervisor rated measure and interview data. This mixed method approach allowed for triangulation of different types of data, in order to thoroughly examine PWP effectiveness. The blinding procedures used in the design of the research were also a strength of the design. Such procedures minimised any potential bias in analysing the effectiveness of PWPs, as their rank position was not known whilst MLM and qualitative analysis were being carried out.

Although the findings from the multi-level modelling can be generalised to a wider population of PWPs due to the treatment of PWPs as a random factor in the analysis, the results of the comparison of upper and lower quartiles should be interpreted with caution due to the small numbers of PWPs in each group (n=5). The small sample therefore appears robust for using MLM, but means that small numbers of PWPs in the upper and lower groups for comparison may have reduced the reliability of these comparisons. Future studies should aim to use a larger sample size in order to increase the number of practitioners in the most and least effective groups, therefore increasing the reliability of the findings.

4.5 Conclusions

This study has found that there is variability between the effectiveness of PWPs and that therapist effects exist in this population. Almost 9 per cent of the variance in outcomes was found to be attributable to PWPs, which is consistent with the broader literature on therapist effects. Variation exists in a PWP sample despite the standardised training and supervision of this therapist group, and the finding adds to a growing body of literature indicating that therapists provide

a significant contribution to outcomes. The treatment of PWPs as a random factor in the MLM analysis means that this finding can be generalised to PWPs more widely.

More effective PWPs were found to have higher levels of resilience and reported confidence in their skills and abilities, and engaged in their work in a proactive, organised and thorough manner. The results of qualitative analysis suggested that the more effective PWPs may have been at a more advanced stage of their clinical development than less effective PWPs. Supervisors of less effective PWPs reported more use of experiential intuition in their supervisees than supervisors of more effective PWPs.

This is the first study to examine the role of PWPs in this way. The findings suggest that there are differences between the most and least effective PWPs, particularly in their levels of resilience and future studies should aim to explore the relationship between resilience and effective practitioners further.

References

Barkham, M., Stiles, W.B., Connell, J., Twigg, E., Leach, C., Lucock, M., Mellor-Clark, J., Bower, P., King, M., Shapiro, D.A., Hardy, G.E., Greenberg, L., & Angus, L. (2008). Effects of psychological therapies in randomized trials and practice-based studies. *British Journal of Clinical Psychology, 47*, 397-415.

Brown, G.S., Lambert, M.J., Jones, E.R. & Minami, T. (2005). Identifying highly effective psychotherapists in a managed care environment. *American Journal of Managed Care, 11*, 513-520.

Clark, D.M., Layard, R., Smithies, R., Richards, D.A., Suckling, R. & Wright, B. (2009). Improving access to psychological therapy: Initial evaluation of two UK demonstration sites. *Behaviour Research and Therapy,* 47, 910-920.

Connor, K.M. & Davidson, J.R.T. (2003). Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety*, *18*, 6-82.

Crits-Christoph, P., & Gallop, R. (2006). Therapist effects in the National Institute of Mental Health Treatment of Depression Collaborative Research Program and other psychotherapy studies. *Psychotherapy Research, 16*, 178-181.

Care Services and Improvement Partnership Choice and Access Team (2008) *Improving Access to Psychological Therapies (IAPT) Commissioning Toolkit.* London: Department of Health.

Dinger, U., Strack, M., Leichsenring, F., Wilmers, F., & Schauenburg, H. (2006). Therapist effects on outcome and alliance in inpatient psychotherapy. *Journal of Clinical Psychology, 64,* 344-354.

Elkin, I., Falconnier, L., Martinovich, Z., & Mahoney, C. (2006a). Therapist effects in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Psychotherapy Research, 16,* 144-160.

Elkin, I., Falconnier, L., Martinovich, Z., & Mahoney, C. (2006b). Rejoinder to commentaries by Stephen Soldz and Paul Crits-Christoph on therapist effects. *Psychotherapy Research, 16*, 182-183.

Elkin, I., Shea, M.T., Watkins, J.T., Imber, S.D., Stotsky, S.M., Collins, J.F.& Parloff, M.B. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archive of General Psychiatry, 46*, 971-982.

English, I. (1993). Intuition as a function of the expert nurse: A critique of Benner's novice to expert model. *Journal of Advanced Nursing, 18*, 387-393.

Faust, D. & Zlotnick, C. (1995). Another dodo verdict? Revisiting the comparative effectiveness of professional and paraprofessional therapists. *Clinical Psychology and Psychotherapy, 2,* 157-167.

Gilbody, S., Richards, D., Brearley, S., & Hewitt, C. (2007). Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. *Journal of General Internal Medicine, 22*, 1596-1602.

Glover, G., Webb, M., & Evison, F. (2010). Improving Access to Psychological Therapies: A review of the progress made by sites in the first roll-out year. *North East Public Health Observatory.*

Handley, S.J., Newstead, S.E., & Wright, H. (2000). Rational and experiential thinking: A study of the REI in *International Perspectives on Individual Differences Volume 1 Cognitive styles*. R.J. Riding & S.G. Rayner. (Eds) Ablex Publishing Corporation, USA Stamford.

Hubble, M.A., Duncan, B.L., & Miller, S.D. (1999). The heart and soul of change: Delivering what works in therapy. Washington DC: American Psychological Association.

Huppert, J.D., Bufka, L.F., Barlow, D.H., Gorman, J.M., Shear, K.M, & Woods, S.W. (2001). Therapists, therapist variables, and Cognitive-Behavioral Therapy Outcome in a multicenter trial for panic disorder. *Journal of Consulting and Clinical Psychology, 69*, 747-755.

Improving Access to Psychological Therapies (2008). *Implementation Plan: Curriculum for Low-Intensity Therapies Workers.* London: Department of Health.

Jacobson, N.S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 2-19.

Jennings, L. & Skovholt, T.M. (1999). The cognitive, emotional and relational characteristics of master therapists. *Journal of Counseling Psychology, 46*, 3-11

Kim, D., Wampold, B.E., & Bolt, D.M. (2006). Therapist effects in psychotherapy: A random-effects modelling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research, 16*, 161-172.

King, N. (1998). Template analysis. In G.Symon & C. Cassel, C. (Eds). *Qualitative methods and analysis in organisational research: A practical guide.* London: Sage Publications.

King, N. (2004). Using templates in the thematic analysis of text. In C. Cassell & G. Symon (Eds). *Essential guide to qualitative methods in organizational research.* London: Sage Publications.

King, N., Thomas, K., Bell, D., & Bowes, N. (2003). *Evaluation of the Calderdale and Kirklees out of hours protocol for palliative care: Final report.*

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*, 267-276.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*, 606-613.

Lambert, M.J., Whipple, J.L., Vermeersch, D.A., Smart, D.W., Hawkins, E.J., Nielsen, S.L., & Goates, M. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology and Psychotherapy, 9*, 91-103.

Layard, R. (2006). *The depression report: A new deal for depression and anxiety disorders.* Centre for Economic Performance's Mental Health Policy Group.

Luborsky, L.L., McLellan, A.T., Woody, G.E., O'Brien, C.P. & Auerbach, A. (1985). Therapist success and its determinants. *Archives of General Psychiatry, 42*, 602-611.

Lutz, W., Scott, L., Martinovich, Z., Lyons, J.S., & Stiles, W. B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology, 54*, 32-39.

Markstrom, C.A., & Marshall, S.K. (2007). The psychosocial inventory of ego strengths: Examination of theory and psychometric properties. *Journal of Adolescence, 30*, 63-79.

Markstrom, C.A., Sabino, V.M., Turner, B.J. & Berman, R.C. (1997) The Psychosocial Inventory of Ego Strengths: Development and validation of a new Eriksonian measure. *Journal of Youth and Adolescence, 26*, 705-732.

Miller, S. D., Hubble, M. & Duncan, B. (2008). Supershrinks: What is the secret of their success? *Psychotherapy in Australia, 14*, 14-22.

Najavits, L.M., & Strupp, H. H. (1994). Differences in the effectiveness of psychodynamic therapists: A process-outcome study. *Psychotherapy, 31*, 114-123.

Okiishi, J.C., Lambert, M.J., Nielsen, S.L., & Ogles, B.M. (2003). Waiting for supershrink: Empirical analysis of therapist effects. *Clinical Psychology and Psychotherapy, 10*, 361-373.

Pacini, R. & Epstein, S. (1999). The relation of rational and experieintial information processing styles to personality basic beliefs and the ratio-bias phenomenon. *Journal of Personality and Social Psychology, 76*, 972-987.

Pretz, J.E. & Totz, K.S. (2007). Measuring individual differences in affective, heuristic and holistic intuition. *Personality and Individual Differences, 43,* 1247-1257.

Rabash, J., Charlton, C., Browne, W.J., Healy, M., & Cameron, B. (2009). *MLwiN Version 2.11. Centre for Multilevel Modelling.* University of Bristol.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierachical linear models: Applications and data analysis methods. 2*nd Edition. Newbury Park, CA: Sage.

Rea, B. (2001). Finding our balance: The investigation and clinical application of intuition. *Psychotherapy, 38*, 97-106.

Richards, D. & Whyte, M. (2009). Reach Out: National programme student materials to support the delivery of training for Psychological Wellbeing Practitioners delivering low intensity interventions. 2nd Edition. Rethink, UK.

Richardson, G., & Richards, D. (2009). *Psychological Wellbeing Practitioners: Playing a key role in maintaining the nation's wellbeing. Best practice guide.* Improving Access to Psychological Therapies.

Roth, A. & Fonagy, P. (2005) *What Works for Whom? A Critical Review of Psychotherapy Research.* New York: The Guilford Press.

Rutter, M. (1993). Resilience: Some conceptual considerations. *Journal of Adolescent Health, 14,* 626-631.

Saxon, D., & Barkham, M. (submitted). The shape of therapist effects: The contribution of patient severity and risk to therapist variation in effectiveness.

Soldz, S. (2006). Models and meanings: Therapist effects and the stories we tell. *Psychotherapy Research, 16,* 173-177.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Lowe, B. (2006). A brief measure for assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine. 166*, 1092-1097.

Stoltenberg, C.D., & McNeill, B.W. (1997). *Clinical supervision from a developmental perspective: Research and practice.* In C.E. Watkins Jr. (Ed.) Handbook of psychotherapy supervision. Pp. 184-202. New York: Wiley.

Turpin, G. & Wheeler, S. (2011). IAPT supervision guidance. Improving Access to Psychological Therapies.

Tusai, K. & Dyer, J. (2004). Resilience: A historical review of the construct. *Holistic Nursing Practice, 18*, 3-10.

Wampold, B.E. (2001). *The Great Psychotherapy Debate: Models, Methods and Findings.* NJ: Lawrence Erlbaum Associates

Wampold, B.E., & Brown, G.S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73*, 914-923.

Weinberger, J. (1993). Common factors in psychotherapy. In G. Stricker & J.R. Gold (Eds.), *Comprehensive Handbook of Psychotherapy Integration.* New York: Plenum Press.

Welling, H. (2005). The intuitive process: The case of psychotherapy. *Journal of Psychotherapy Integration, 15*, 19-47.

Appendices

Appendix I – University Journal Approval Letter

The University Of Sheffield.

Department Of Psychology.
Clinical Psychology Unit.

Doctor of Clinical Psychology (DClin Psy) Programme
Clinical supervision training and NHS research training & consultancy.

Clinical Psychology Unit
Department of Psychology
University of Sheffield
Western Bank
Sheffield S10 2TP UK

Telephone: 0114 2226550
Fax: 0114 2226610
Email: c.harrison@sheffield.ac.uk

14th March 2011

Helen Green
Third year trainee
Clinical Psychology Unit
University of Sheffield

Dear Helen

I am writing to indicate our approval of the journal(s) you have nominated for publishing work contained in your research thesis.

Literature Review: Journal of Consulting & Clinical Psychology

Research Report: Behaviour, Research & Therapy

Please ensure that you bind this letter and copies of the relevant Instructions to Authors into an appendix in your thesis.

Yours sincerely

Dr Rebecca Knowles
Research Tutor

Appendix II – Instructions for Authors, Journal of Consulting and Clinical Psychology

Journal of Consulting and Clinical Psychology – Instructions for Authors

Please consult APA's Instructions for All Authors for information regarding

- Manuscript Preparation
- Submitting Supplemental Materials
- References
- Figures
- Permissions
- Publication Policies
- Ethical Principles

Submission

Prior to submission, please review the submission guidelines detailed below. Starting in 2011, the completion of a Manuscript Submission Checklist (PDF, 35KB) that signifies that authors have read this material and agree to adhere to the guidelines is now required. The checklist should follow the cover letter as part of the submission.

Manuscripts that do not conform to the submission guidelines may be returned without review.

Please submit manuscripts electronically, either using Microsoft Word (.doc) or Rich Text Format (.rtf) via the Manuscript Submission Portal.

If you encounter difficulties with submission, please e-mail Sara Young or call 202-336-5859.

General correspondence may be directed to the Editorial Office via e-mail.

Masked Review

This journal uses a masked reviewing system for all submissions. The first page of the manuscript should omit the authors' names and affiliations but should include the title of the manuscript and the date it is submitted.

Footnotes containing information pertaining to the authors' identities or affiliations should not be included in the manuscript, but may be provided after a manuscript is accepted.

Make every effort to see that the manuscript itself contains no clues to the authors' identities.

Keep a copy of the manuscript to guard against loss.

Cover Letter

The cover letter accompanying the manuscript submission must include all authors' names and affiliations to avoid potential conflicts of interest in the review process. Addresses and phone numbers, as well as electronic mail addresses and fax numbers, if available, should be provided for all authors for possible use by the editorial office and later by the production office.

Length and Style of Manuscripts

Full-length manuscripts should not exceed 35 pages total (including cover page, abstract, text, references, tables, and figures), with margins of at least 1 inch on all sides and a standard font (e.g., Times New Roman) of 12 points (no smaller). The entire paper (text, references, tables, etc.) must be double spaced.

Instructions on preparing tables, figures, references, metrics, and abstracts appear in the *Publication Manual of the American Psychological Association* (6th edition).

Authors submitting manuscripts that report new data collection, especially randomized clinical trials (RCTs), should comply with the newly developed APA Journal Article Reporting Standards (PDF: 98 KB) (JARS; see *American Psychologist*, 2008, *63*, 839–851 or Appendix in the *APA Publication Manual*).

For papers that exceed 35 pages, authors must justify the extended length in their cover letter (e.g., reporting of multiple studies), and in no case should the paper exceed 45 pages total. Papers that do not conform to these guidelines may be returned without review.

The References section should immediately follow a page break.

Brief Reports

In addition to full-length manuscripts, the *JCCP* will consider Brief Reports of research studies in clinical psychology. The Brief Report format may be appropriate for empirically sound studies that are limited in scope, contain novel or provocative findings that need further replication, or represent replications and extensions of prior published work.

Brief Reports are intended to permit the publication of soundly designed studies of specialized interest that cannot be accepted as regular articles because of lack of space.

Brief Reports must be prepared according to the following specifications: Use 12-point Times New Roman type and 1-inch (2.54-cm) margins, and do not exceed 265 lines of text including references. These limits do not include the title page, abstract, author note, footnotes, tables, or figures.

An author who submits a Brief Report must agree not to submit the full report to another journal of general circulation. The Brief Report should give a clear, condensed summary of the procedure of the study and as full an account of the results as space permits.

Commentaries

JCCP now publishes papers that are commentaries of previously published articles in this journal. Two types of commentaries will be considered:

Brief Comment

A Brief Comment would be written in response to a single article previously published in *JCCP*. The primary purpose would be to provide a meaningful insight, concern, alternative interpretation, clarification, or critical analysis. It is not intended to be pedestrian in nature (e.g., simply highlighting that a given study is statistically underpowered). Rather, its publication would provide for a richer and more comprehensive understanding of a methodological, conceptual, or professional issue that significantly adds to the literature.

Similar to a Brief Report, Brief Comments should not exceed 265 lines of text including references. This limit does not include the title page, abstract, or author notes. The title of a Brief Comment should include a subtitle reflecting the actual title and year of publication of the article that engendered the comment. For example—"The Importance of Focusing on External Validity: A Brief Comment on *Testing the Efficacy of Two Differing Types of Stress Management Interventions for the Treatment of Essential Hypertension* (Jones & Smith, 2012)."

Brief Comments should be submitted in a timely manner, no later than 9 months after publication of the original article. Upon acceptance of a Brief Comment, the author(s) of the original paper would be invited to submit a response, whereupon, if acceptable, both the Brief Comment and Response would be published together. Such Responses to a Brief Comment should also not exceed 265 lines of text including references.

Extended Comment

The purpose of this type of article is essentially similar to that of a Brief Comment (i.e., to provide a meaningful insight, concern, alternative interpretation, clarification, or critical analysis), but would be written in response to a series of articles previously published in *JCCP* or that involves a more extensive and far-reaching conceptual or methodological issue. An example might include describing and analyzing the limitations of a particular statistical or methodological procedure used in several studies previously published in *JCCP*, provided along with meaningful recommendations.

This type of article should not exceed approximately one half the length of the original paper (note that 1 journal page equals approximately 3–3.5 manuscript pages). Unless permission from the editor is received, no Extended Comment should exceed 20 manuscript pages inclusive of all references, tables, and figures.

Similar to a Brief Comment, where and when appropriate, if such a paper is accepted, the author(s) of the original article(s) will be contacted to write a response, whereupon, if acceptable, both the Extended Comment and Response would be published together. This Invited Response should not exceed approximately one half the length of the Extended Comment.

The title of this type of article need not include a subtitle representing the original article(s). One important review criteria involves the timeliness of the topic and its potential contribution to the scientific literature base relevant to the scope of *JCCP* content.

Conceptual/Theoretical Papers

Whereas the majority of papers published in *JCCP* will involve descriptions of quantitatively-based investigations, this journal also considers conceptual articles on topics of broad theoretical, methodological, or practical interest that advance the field of clinical psychology. Examples might include describing a new methodological or statistical procedure, delineating methods of enhancing dissemination of research findings from the lab to real-world settings, or advocating the need to increase the profession's research efforts regarding a traditionally underserved population.

Similar formatting guidelines for submitting a full length research article would apply for these types of papers.

Title of Manuscript

The title of a manuscript should be accurate, fully explanatory, and preferably no longer then 12 words. The title should reflect the content and population studied (e.g., "treatment of generalized anxiety disorders in adults").

If the paper reports a randomized clinical trial (RCT), this should be indicated in the title. Note that JARS criteria must be used for reporting purposes.

Abstract and Keywords

Starting in 2010, all manuscripts published in the *Journal of Consulting and Clinical Psychology* will include a structured abstract of up to 250 words.

For studies that report randomized clinical trials or meta-analyses, the abstract also must be consistent with the guidelines set forth by JARS or MARS (Meta-Analysis Reporting Standards) guidelines, respectively. Thus, in preparing a manuscript, please ensure that it is consistent with the guidelines stated below.

Please include an Abstract of up to 250 words, presented in paragraph form. The Abstract should be typed on a separate page (page 2 of the manuscript), and must include each of the following sections:

- **Objective:** A brief statement of the purpose of the study
- **Method:** A detailed summary of the participants (*N*, age, gender, ethnicity) as well as descriptions of the study design, measures (including names of measures), and procedures
- **Results:** A detailed summary of the primary findings that clearly articulate comparison groups (if relevant), and that indicate significance or confidence intervals for the main findings
- **Conclusions:** A description of the research and clinical implications of the findings

After the abstract, please supply up to five keywords or short phrases.

Participants: Description and Informed Consent

The Method section of each empirical report must contain a detailed description of the study participants, including (but not limited to) the following: age, gender, ethnicity, SES, clinical diagnoses and comorbidities (as appropriate), and any other relevant demographics.

In the Discussion section of the manuscript, authors should discuss the diversity of their study samples and the generalizability of their findings.

The Method section also must include a statement describing how informed consent was obtained from the participants (or their parents/guardians) and indicate that the study was conducted in compliance with an appropriate Internal Review Board.

Measures

The Method section of empirical reports must contain a sufficiently detailed description of the measures used so that the reader understands the item content, scoring procedures, and total scores or subscales. Evidence of reliability and validity with similar populations should be provided.

Statistical Reporting of Clinical Significance

JCCP requires the statistical reporting of measures that convey clinical significance. Authors should report means and standard deviations for all continuous study variables and the effect sizes for the primary study findings. (If effect sizes are not available for a particular test, authors should convey this in their cover letter at the time of submission.)

JCCP also requires authors to report confidence intervals for any effect sizes involving principal outcomes (see Fidler et al., *Journal of Consulting and Clinical Psychology*, 2005, pp. 136–143 and Odgaard & Fowler, *Journal of Consulting and Clinical Psychology*, 2010, pp.287–297).

In addition, when reporting the results of interventions, authors should include indicators of clinically significant change. Authors may use one of several approaches that have been recommended for capturing clinical significance, including (but not limited to) the reliable change index (i.e., whether the amount of change displayed by a treated individual is large enough to be meaningful; see Jacobson et al., *Journal of Consulting and Clinical Psychology*, 1999), the extent to which dysfunctional individuals show movement into the functional distribution (see Jacobson & Truax, *Journal of Consulting and Clinical Psychology*, 1991), or other normative comparisons (see Kendall et al., *Journal of Consulting and Clinical Psychology*, 1999).

The special section of *JCCP* on "Clinical Significance" (*Journal of Consulting and Clinical Psychology*, 1999, pp. 283–339) contains detailed discussions of clinical significance and its measurement and should be a useful resource (see also Atkins et al., *Journal of Consulting and Clinical Psychology*, 2005, pp. 982–989).

Discussion of Clinical Implications

Articles must include a discussion of the clinical implications of the study findings or analytic review. The Discussion section should contain a clear statement of the extent of clinical application of the current assessment, prevention, or treatment methods. The extent of application to clinical practice may range from suggestions that the data are too preliminary to support widespread dissemination to descriptions of existing manuals available from the authors or archived materials that would allow full implementation at present.

Randomized Clinical Trials: Use of JARS Guidelines

JCCP requires the use of JARS guidelines for randomized clinical trials, consistent with the recommendations and policies established by the Publications and Communications Board of the American Psychological Association. JARS offers a standard way to improve the quality of such reports, and to ensure that readers have the information necessary to evaluate the quality of a clinical trial.

Manuscripts that report randomized clinical trials are required to include a flow diagram of the progress through the phases of the trial. When a study is not fully consistent with JARS guidelines, the limitations should be acknowledged and discussed in the text of the manuscript.

For follow-up studies of previously published clinical trials, authors should submit a flow diagram of the progress through the phases of the trial and follow-up. The above checklist information should be completed to the extent possible, especially for the Results and Discussion sections of the manuscript.

Authors of RCTs should also describe procedures to assess for treatment fidelity (also known as treatment integrity), including both therapist adherence and competence. Where possible, results should be reported regarding the relationship between fidelity and outcome found in the investigation.

View the JARS guidelines (PDF: 98 KB)

Meta-Analyses of Randomized Clinical Trials: Use of MARS Guidelines

JCCP requires the use of the APA MARS guidelines for meta-analyses of randomized clinical trials. MARS offers a standard way to improve the quality of such reports, and to ensure that readers have the information necessary to evaluate the quality of a meta-analysis.

Manuscripts that report meta-analyses of randomized clinical trials are required to include a flow diagram of the progress through the stages of the meta-analysis. When a study is not fully consistent with MARS, the limitations should be acknowledged and discussed in the text of the manuscript.

MARS guidelines are included in the JARS guidelines (PDF: 98 KB)

Nonrandomized Trials

For nonrandomized designs that often are used in public health and mental-health interventions, *JCCP* requires compliance with JARS.

Failure to comply with JARS or MARS can result in the return of manuscripts without review.

Appendix III

Instructions for Authors – Behaviour, Research and Therapy

Preparation

Article structure

Subdivision - unnumbered sections
Divide your article into clearly defined sections. Each subsection is given a brief heading. Each heading should appear on its own separate line. Subsections should be used as much as possible when cross-referencing text: refer to the subsection by heading as opposed to simply "the text".

Appendices
If there is more than one appendix, they should be identified as A, B, etc. Formulae and equations in appendices should be given separate numbering: Eq. (A.1), Eq. (A.2), etc.; in a subsequent appendix, Eq. (B.1) and so on. Similarly for tables and figures: Table A.1; Fig. A.1, etc.

Essential title page information

• ***Title.*** Concise and informative. Titles are often used in information-retrieval systems. Avoid abbreviations and formulae where possible.
• ***Author names and affiliations.*** Where the family name may be ambiguous (e.g., a double name), please indicate this clearly. Present the authors' affiliation addresses (where the actual work was done) below the names. Indicate all affiliations with a lower-case superscript letter immediately after the author's name and in front of the appropriate address. Provide the full postal address of each affiliation, including the country name, and, if available, the e-mail address of each author.
• ***Corresponding author.*** Clearly indicate who will handle correspondence at all stages of refereeing and publication, also post-publication. **Ensure that telephone and fax numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address. Contact details must be kept up to date by the corresponding author.**
• ***Present/permanent address.*** If an author has moved since the work described in the article was done, or was visiting at the time, a "Present address" (or "Permanent address") may be indicated as a footnote to that author's name. The address at which the author actually did the work must be retained as the main, affiliation address. Superscript Arabic numerals are used for such footnotes.

Abstract
A concise and factual abstract is required with a maximum length of 200 words. The abstract should state briefly the purpose of the research, the principal results and major conclusions. An abstract is often presented separately from the article, so it must be able to stand alone. For this reason, References should be avoided, but if essential, then cite the author(s) and year(s). Also, non-standard or uncommon abbreviations should be avoided, but if essential they must be defined at their first mention in the abstract itself.

Graphical abstract

A Graphical abstract is optional and should summarize the contents of the article in a concise, pictorial form designed to capture the attention of a wide readership online. Authors must provide images that clearly represent the work described in the article. Graphical abstracts should be submitted as a separate file in the online submission system. Image size: Please provide an image with a minimum of 531 × 1328 pixels (h × w) or proportionally more. The image should be readable at a size of 5 × 13 cm using a regular screen resolution of 96 dpi. Preferred file types: TIFF, EPS, PDF or MS Office files. See http://www.elsevier.com/graphicalabstracts for examples.

Highlights

Highlights are mandatory for this journal. They consist of a short collection of bullet points that convey the core findings of the article and should be submitted in a separate file in the online submission system. Please use 'Highlights' in the file name and include 2 to 5 bullet points (maximum 85 characters, including spaces, per bullet point). See http://www.elsevier.com/highlights for examples.

Keywords
Immediately after the abstract, provide a maximum of 6 keywords, to be chosen from the APA list of index descriptors. These keywords will be used for indexing purposes.

Abbreviations

Define abbreviations that are not standard in this field in a footnote to be placed on the first page of the article. Such abbreviations that are unavoidable in the abstract must be defined at their first mention there, as well as in the footnote. Ensure consistency of abbreviations throughout the article.

Acknowledgements

Collate acknowledgements in a separate section at the end of the article before the references and do not, therefore, include them on the title page, as a footnote to the title or otherwise. List here those individuals who provided help during the research (e.g., providing language help, writing assistance or proof reading the article, etc.).

Shorter communications
This option is designed to allow publication of research reports that are not suitable for publication as regular articles. Shorter Communications are appropriate for articles with a specialized focus or of particular didactic value. Manuscripts should be between 3000-5000 words, and must not exceed the upper word limit. This limit includes the abstract, text, and references, but not the title page, tables and figures.

Artwork

Electronic artwork
General points
• Make sure you use uniform lettering and sizing of your original artwork.
• Save text in illustrations as "graphics" or enclose the font.
• Only use the following fonts in your illustrations: Arial, Courier, Times, Symbol.
• Number the illustrations according to their sequence in the text.
• Use a logical naming convention for your artwork files.
• Provide captions to illustrations separately.
• Produce images near to the desired size of the printed version.
• Submit each figure as a separate file.

A detailed guide on electronic artwork is available on our website:
http://www.elsevier.com/artworkinstructions
You are urged to visit this site; some excerpts from the detailed information are given here.
Formats
Regardless of the application used, when your electronic artwork is finalised, please "save as" or convert the images to one of the following formats (note the resolution requirements for line drawings, halftones, and line/halftone combinations given below):
EPS: Vector drawings. Embed the font or save the text as "graphics".
TIFF: color or grayscale photographs (halftones): always use a minimum of 300 dpi.
TIFF: Bitmapped line drawings: use a minimum of 1000 dpi.
TIFF: Combinations bitmapped line/half-tone (color or grayscale): a minimum of 500 dpi is required.
If your electronic artwork is created in a Microsoft Office application (Word, PowerPoint, Excel) then please supply "as is".

Please do not:
• Supply files that are optimised for screen use (like GIF, BMP, PICT, WPG); the resolution is too low;
• Supply files that are too low in resolution;
• Submit graphics that are disproportionately large for the content.

Tables

Number tables consecutively in accordance with their appearance in the text. Place footnotes to tables below the table body and indicate them with superscript lowercase letters. Avoid vertical rules. Be sparing in the use of tables and ensure that the data presented in tables do not duplicate results described elsewhere in the article.

References

Citation in text
Please ensure that every reference cited in the text is also present in the reference list (and vice versa). Any references cited in the abstract must be given in full. Unpublished results and personal communications are not recommended in the reference list, but may be mentioned in the text. If these references are included in the reference list they should follow the standard reference style of the journal and should include a substitution of the publication date with either "Unpublished results" or "Personal communication" Citation of a reference as "in press" implies that the item has been accepted for publication.

Web references
As a minimum, the full URL should be given and the date when the reference was last accessed. Any further information, if known (DOI, author names, dates, reference to a source publication, etc.), should also be given. Web references can be listed separately (e.g., after the reference list) under a different heading if desired, or can be included in the reference list.

Reference management software
This journal has standard templates available in key reference management packages EndNote (http://www.endnote.com/support/enstyles.asp) and Reference Manager (http://refman.com/support/rmstyles.asp). Using plug-ins to wordprocessing packages, authors only need to select the appropriate journal template when preparing their article and the list of references and citations to these will be formatted according to the journal style which is described below.

Reference style
Text: Citations in the text should follow the referencing style used by the American Psychological Association. You are referred to the Publication Manual of the American Psychological Association, Sixth Edition, ISBN 978-1-4338-0561-5, copies of which may be ordered from http://books.apa.org/books.cfm?id=4200067 or APA Order Dept., P.O.B. 2710, Hyattsville, MD 20784, USA or APA, 3 Henrietta Street, London, WC3E 8LU, UK. Details concerning this referencing style can also be found at http://linguistics.byu.edu/faculty/henrichsenl/apa/apa01.html.
List: references should be arranged first alphabetically and then further sorted chronologically if necessary. More than one reference from the same author(s) in the same year must be identified by the letters "a", "b", "c", etc., placed after the year of publication.
Examples:
Reference to a journal publication:
Van der Geer, J., Hanraads, J. A. J., & Lupton, R. A. (2000). The art of writing a scientific article. *Journal of Scientific Communications, 163*, 51–59.
Reference to a book:
Strunk, W., Jr., & White, E. B. (1979). *The elements of style.* (3rd ed.). New York: Macmillan, (Chapter 4).
Reference to a chapter in an edited book:
Mettam, G. R., & Adams, L. B. (1994). How to prepare an electronic version of your article. In B. S. Jones, & R. Z. Smith (Eds.), *Introduction to the electronic age* (pp. 281–304). New York: E-Publishing Inc.

Video data

Elsevier accepts video material and animation sequences to support and enhance your scientific research. Authors who have video or animation files that they wish to submit with their article are strongly encouraged to include these within the body of the article. This can be done in the same way as a figure or table by referring to the video or animation content and noting in the body text where it should be placed. All submitted files should be properly labeled so that they directly relate to the video file's content. In order to ensure that your video or animation material is directly usable, please provide the files in one of our recommended file formats with a preferred maximum size of 50 MB. Video and animation files supplied will be published online in the electronic version of your article in Elsevier Web products, including ScienceDirect: http://www.sciencedirect.com. Please supply 'stills' with your files: you can choose any frame from the video or animation or make a separate image. These will be used instead of standard icons and will personalize the link to your video data. For more detailed instructions please visit our video instruction pages at http://www.elsevier.com/artworkinstructions. Note: since video and animation cannot be embedded in the print version of the journal, please provide text for both the electronic and the print version for the portions of the article that refer to this content.

Supplementary data

Elsevier accepts electronic supplementary material to support and enhance your scientific research. Supplementary files offer the author additional possibilities to publish supporting applications, high-resolution images, background datasets, sound clips and more. Supplementary files supplied will be published online alongside the electronic version of your article in Elsevier Web products, including ScienceDirect: http://www.sciencedirect.com. In order to ensure that your submitted material is directly usable, please provide the data in one of our recommended file formats. Authors should submit the material in electronic format together with the article and supply a concise and descriptive caption for each file. For more detailed instructions please visit our artwork instruction pages at http://www.elsevier.com/artworkinstructions.

Submission checklist

The following list will be useful during the final checking of an article prior to sending it to the journal for review. Please consult this Guide for Authors for further details of any item.
Ensure that the following items are present:
One Author designated as corresponding Author:
• E-mail address
• Full postal address
• Telephone and fax numbers
All necessary files have been uploaded
• Keywords
• All figure captions
• All tables (including title, description, footnotes)
Further considerations
• Manuscript has been "spellchecked" and "grammar-checked"
• References are in the correct format for this journal
• All references mentioned in the Reference list are cited in the text, and vice versa
• Permission has been obtained for use of copyrighted material from other sources (including the Web)
• Color figures are clearly marked as being intended for color reproduction on the Web (free of charge) and in print or to be reproduced in color on the Web (free of charge) and in black-and-white in print
• If only color on the Web is required, black and white versions of the figures are also supplied for printing purposes
For any further information please visit our customer support site at http://support.elsevier.com.

Appendix IV – Downs and Black Quality Checklist

The Downs and Black checklist has not been included for copyright purposes

The University Of Sheffield.

Department Of Psychology.
Clinical Psychology Unit.

Doctor of Clinical Psychology (DClin Psy) Programme
Clinical supervision training and NHS research training & consultancy.

Clinical Psychology Unit
Department of Psychology
University of Sheffield
Western Bank
Sheffield S10 2TN UK

Participant Information Sheet – PWP Version

Key components of effective Psychological Wellbeing practitioners: Evidence from routine practice

You are being invited to take part in a research project. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part. Thank you for reading this.

What is the project's purpose?
Previous research has found that regardless of the psychological model being used in therapy, there are practitioner effects that contribute to the success of effective therapy. The aim of this study is to identify those factors exhibited by Psychological Wellbeing Practitioners (PWPs) that make them effective in their work. This will involve asking you to complete some questionnaires. We will also ask a sample of PWPs and their supervisors to take part in telephone interviews about what they believe makes an effective PWP. Supervisors will also be asked to complete a mirror set of questionnaires about how you go about your work in everyday practice.

You have been chosen to take part because you are due to complete your Low Intensity Worker PG Cert training in 2010 at the University of Sheffield, Nottingham or York.

Do I have to take part?
It is up to you to decide whether or not to take part. If you do decide to take part you will be asked to sign a consent form. You can still withdraw at any time without it affecting any part of your training or employment; you do not have to give a reason. To withdraw please contact Helen Green by email at pcp08hlg@shef.ac.uk

What will happen to me if I take part?
If you do decide to take part, the research team will be given access to your client outcome scores on the IAPT minimum dataset through training and when you start in services from September 2010 to April 2011. You will be sent a questionnaire pack and freepost return, and those wishing to take part will complete and return them. Those who do wish to take part may then be chosen to take part in an individual telephone interview in Sept 2010 – May 2011. Supervisors of interviewed PWPs will be contacted in early 2011.

If you decide to take part in the interviews, these will be carried out over the telephone. These interviews will be recorded and you will be informed of this at the beginning of the interview. The

audio recordings of the interviews made during this research will be used only for analysis. No other use will be made of them without your written permission and no one outside the project will be allowed access to the original recordings. Once the research is complete, the tapes will be destroyed.

What are the possible risks and benefits of taking part?
There are no foreseeable risks involved in taking part in this study, however should you experience any discomfort, disadvantage or risk you should contact the lead researcher (Helen Green) immediately. Whilst there are no immediate benefits for those people participating in the project, it is hoped that this work will inform future training programmes and understanding of the components of effective therapists.

What happens if the research study stops earlier than expected?
If for any unseen reason the research study stops earlier than expected, you will be contacted with the option of the data from your questionnaires and interview tapes and transcripts being destroyed and not used in any other research. However, you will also be given the option of this information being used in other research.

What will happen to the results of the research project?
All the information that we collect about you during the course of the research will be kept strictly confidential between members of the research team. You will not be able to be identified in any reports or publications. The results of the research will form the basis for the lead researcher's Doctor of Clinical Psychology thesis due for completion in July 2011. It is intended that from this the results will be written up for publication in the following year. This research is being funded by the Department of Clinical Psychology, University of Sheffield.

Who has ethically reviewed the project?
All research in the NHS is looked at by an independent group of people called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by the South Yorkshire Research Ethics Committee.

What if I wish to complain about the way in which this study has been conducted?
If you have *any* cause to complain about the way in which you have been approached or treated during the course of this study, the normal NHS complaints mechanisms are available to you and are not compromised in any way because you have taken part in a research study.

If you have any complaints or concerns please contact the project co-ordinator, Helen Green at pcp08hlg@sheffield.ac.uk. Otherwise you can use the normal University complaints procedure and contact the following person: Dr David Fletcher, Registrar and Secretary's Office, University of Sheffield, Firth Court, Western Bank, Sheffield, S10 2TN.

Alternatively you can contact the Patient Advice Liaison Service (PALS) at: Faye Mellors, PALS Officer, Sheffield Health and Social Care NHS Foundation Trust, Fulwood House, Old Fulwood Road, Sheffield, S10 3TH or telephone on 0114 271 8768.

Contact for further information
If you would like further information, please email Helen Green at: pcp08hlg@sheffield.ac.uk.

Thank you for reading this information sheet.

The University Of Sheffield.

Department Of Psychology.
Clinical Psychology Unit.

Doctor of Clinical Psychology (DClin Psy) Programme
Clinical supervision training and NHS research training & consultancy.

Clinical Psychology Unit
Department of Psychology
University of Sheffield
Western Bank
Sheffield S10 2TN UK

Participant Consent Form (PWP Version)

Title of Project: Key features of effective Psychological Wellbeing Practitioners: Evidence from routine practice.

Name of Researcher: Helen Green

Participant Identification Number for this project:

Please initial box

1. I confirm that I have read and understand the information sheet for the above project and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. ☐
3. I understand that my responses will be anonymised before analysis. I give permission for members of the research team to have access to my anonymised responses. ☐
4. I consent for my supervisor to be contact if I am selected to take part in the telephone interviews. ☐
5. I agree to take part in the above research project. ☐

Name of Participant	Date	Signature
Lead Researcher	Date	Signature

V1 – 22nd June 2010

The University Of Sheffield.

Department Of Psychology.
Clinical Psychology Unit.

Doctor of Clinical Psychology (DClin Psy) Programme
Clinical supervision training and NHS research training & consultancy.

Clinical Psychology Unit
Department of Psychology
University of Sheffield
Western Bank
Sheffield S10 2TN UK

Participant Information Sheet – Supervisor Version

Key components of effective Psychological Wellbeing practitioners: Evidence from routine practice

You are being invited to take part in a research project. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part. Thank you for reading this.

What is the project's purpose?
Previous research has found that regardless of the psychological model being used in therapy, there are practitioner effects that contribute to the success of effective therapy. The aim of this study is to identify those factors exhibited by Psychological Wellbeing Practitioners (PWPs) that make them effective in their work. This will involve asking you to complete some questionnaires and taking part in a short telephone interview about what you believe makes an effective PWP.

You have been chosen to take part because you are the supervisor of someone due to complete their Low Intensity Worker PG Cert training in 2010 at the University of Sheffield, Nottingham or York.

Do I have to take part?
It is up to you to decide whether or not to take part. If you do decide to take part you will be asked to sign a consent form. You can still withdraw at any time without it affecting any part of your employment; you do not have to give a reason. To withdraw please contact Helen Green by email at pcp08hlg@shef.ac.uk

What will happen to me if I take part?
If you do decide to take part, you will be sent a questionnaire pack which you should complete ***about your supervisee,*** and not about yourself. Your supervisee will have already completed the same set of questionnaires about themselves and they have given their consent for you to be approached to take part in this study.

You will also be contacted to take part in a telephone interview. These interviews will be recorded and you will be informed of this at the beginning of the interview. The audio recordings of the interviews made during this research will be used only for analysis. No other use will be

made of them without your written permission and no one outside the project will be allowed access to the original recordings. Once the research is complete, the tapes will be destroyed.

What are the possible risks and benefits of taking part?
There are no foreseeable risks involved in taking part in this study, however should you experience any discomfort, disadvantage or risk you should contact the lead researcher (Helen Green) immediately. Whilst there are no immediate benefits for those people participating in the project, it is hoped that this work will inform future training programmes and understanding of the components of effective therapists.

What happens if the research study stops earlier than expected?
If for any unseen reason the research study stops earlier than expected, you will be contacted with the option of the data from your questionnaires and interview tapes and transcripts being destroyed and not used in any other research. However, you will also be given the option of this information being used in other research.

What will happen to the results of the research project?
All the information that we collect about you during the course of the research will be kept strictly confidential between members of the research team. You will not be able to be identified in any reports or publications. The results of the research will form the basis for the lead researcher's Doctor of Clinical Psychology thesis due for completion in July 2011. It is intended that from this the results will be written up for publication in the following year. This research is being funded by the Department of Clinical Psychology, University of Sheffield.

Who has ethically reviewed the project?
All research in the NHS is looked at by an independent group of people called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by the South Yorkshire Research Ethics Committee.

What if I wish to complain about the way in which this study has been conducted?
If you have *any* cause to complain about the way in which you have been approached or treated during the course of this study, the normal NHS complaints mechanisms are available to you and are not compromised in any way because you have taken part in a research study.

If you have any complaints or concerns please contact the project co-ordinator, Helen Green at pcp08hlg@sheffield.ac.uk. Otherwise you can use the normal University complaints procedure and contact the following person: Dr David Fletcher, Registrar and Secretary's Office, University of Sheffield, Firth Court, Western Bank, Sheffield, S10 2TN.

Alternatively you can contact the Patient Advice Liaison Service (PALS) at: Faye Mellors, PALS Officer, Sheffield Health and Social Care NHS Foundation Trust, Fulwood House, Old Fulwood Road, Sheffield, S10 3TH or telephone on 0114 271 8768.

Contact for further information
If you would like further information, please email Helen Green at: pcp08hlg@sheffield.ac.uk.

Thank you for reading this information sheet.

Appendix VIII – Participant Consent Form: Supervisor Version

The University Of Sheffield.

Department Of Psychology.
Clinical Psychology Unit.

Doctor of Clinical Psychology (DClin Psy) Programme
Clinical supervision training and NHS research training & consultancy.

Clinical Psychology Unit
Department of Psychology
University of Sheffield
Western Bank
Sheffield S10 2TN UK

Participant Consent Form (Supervisor Version)

Title of Project: Key features of effective Psychological Wellbeing Practitioners: Evidence from routine practice.

Name of Researcher: Helen Green

Participant Identification Number for this project:

Please initial box

1. I confirm that I have read and understand the information sheet for the above project and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. ☐
3. I understand that my responses will be anonymised before analysis. I give permission for members of the research team to have access to my anonymised responses. ☐
4. I agree to take part in the above research project. ☐

Name of Participant	Date	Signature
Lead Researcher	Date	Signature

V1 – 22nd June 2010

Appendix IX – PWP Demographic Information Sheet

Psychological Wellbeing Practitioner

Demographic Information Sheet

Name:.....

Gender: Female Male

Age:.....

Supervisor's Name:.....

Supervisor's Email:.....

Supervisor's Telephone Number:.....

Please outline any previous experience you have working in mental health prior to completing your PG Cert Low Intensity Worker and the length of this experience.

Please outline any mental health training you had prior to completing your PG Cert Low Intensity Worker and the length of this training.

Appendix X – Psychosocial Inventory of Ego Strengths (PIES)

The PIES has not been included for copyright purposes

Appendix XI – Rational Experiential Inventory (REI): PWP Version

The REI has not been included for copyright purposes

Appendix XII – Connor-Davidson Resilience Scale (CD-RISC)

The CD-RISC has not been included for copyright purposes

Interview Schedule – PWP Version

[Start tape recorder]

As previously stated in the information sheet this telephone interview will be recorded. Is that ok?

If participant says yes, begin the interview

1. What previous experience did you have before becoming a psychological wellbeing practitioner?
2. Since finishing the course, how much CPD have you engaged in?
3. How important to you is supervision?
4. What distinguishes a good PWP from a great PWP?
5. What do you think are the characteristics of an effective PWP?
6. Given two equally experienced PWPs, why would one be more effective than the other?
7. What is particularly therapeutic about you?

Check consent once the interview is complete.

Thank you for participating.

Appendix XIV – PWP Interview Schedule: Amended Version

Amended Interview Schedule: PWP Version

[Start tape recorder]

As previously stated in the information sheet this telephone interview will be recorded. Is that ok?

If participant says yes, begin the interview

There will be seven questions altogether and the interview should take about 15-20 minutes.

1. What previous experience did you have before becoming a PWP and how does it help or hinder your current practice?
2. Since finishing the PWP course, how much continuing professional development or training have you engaged in and how has it influenced your PWP practice?
3. How do you use and engage with clinical and case management supervision to improve your skills as a PWP?
4. What other methods do you use to improve your practice?
5. Are you aware of any gaps in your skills or knowledge might make you less effective as a PWP at times?
6. What do you think are the hallmarks of your clinical practice as a PWP.
7. What do you feel makes you effective in your role as a PWP in terms of working in the stepped care model?

That's the end of the interview. Is there anything else that you want to add?

Now that the interview has finished, are you still happy for the results to be used in the analysis?

Thank you for participating.

Appendix XV – Rational Experiential Inventory (REI): Supervisor version

The REI supervisor version has not been included for copyright purposes

Appendix XVI – Supervisor Interview schedule: Original Piloted Version

Interview Schedule – Supervisor Version

[Start tape recorder]

As previously stated in the information sheet this telephone interview will be recorded. Is that ok?

If participant says yes, begin the interview

1. How long have you been supervising X?
2. Since you have been supervising x, how much CPD have they engaged in?
3. How important to x is supervision?
4. How prepared is x for your supervision sessions?
5. What distinguishes a good PWP from a great PWP?
6. What do you think are the characteristics of an effective PWP?
7. Given two equally experienced PWPs, why would one be more effective than the other?
8. What is particularly therapeutic about x?

Check consent once the interview is complete.

Thank you for participating.

Amended Interview Schedule – Supervisor Version

[Start tape recorder]

As previously stated in the information sheet this telephone interview will be recorded. Is that ok?

If participant says yes, begin the interview

There will be eight questions altogether and the interview should take about 15-20 minutes.

1. How long have you been supervising X?
2. Since you have been supervising x, how much continuing professional development or training have they engaged in and how has it influenced their skill levels?
3. How well does X engage in clinical and case management supervision to improve their skills as a PWP?
4. What other methods does x use to improve their practice?
5. What do you think distinguishes an average PWP from a really effective PWP?
6. What do you think are the hallmark clinical and organisational skills of X?
7. Given two equally experienced PWPs, why would one be more effective than the other in your experience?
8. What features of their therapeutic delivery make x effective?

That's the end of the interview. Is there anything else that you want to add?

Now that the interview has finished, are you still happy for the results to be used in the analysis?

Thank you for participating.

Appendix XVIII – Patient Health Questionnaire - 9 (PHQ-9)

The PHQ-9 has not been included for copyright purposes

Appendix XIX – Generalised Anxiety Disorder – 7 (GAD-7)

The GAD-7 has not been included for copyright purposes

NHS

National Research Ethics Service

South Yorkshire Research Ethics Committee
Millside
Mill Pond Road
Meanwood
Leeds
LS6 4RA

Telephone: 0113 3050166
Facsimile:

7 October 2010

Amended "Favourable Opinion with Conditions" Letter (which supersedes letter dated 4 August 2010)

Ms Helen Green
Trainee Clinical Psychologist
Clinical Psychology Unit
University of Sheffield
Western Bank, Sheffield
S10 2TP

Dear Ms Green

Study Title:	**Key components of effective psychological wellbeing practitioners: evidence from routine practice**
REC reference number:	**10/H1310/56**
Protocol number:	**n/a**

The Research Ethics Committee reviewed the above application at the meeting held on 29 July 2010. Thank you for attending to discuss the study.

Ethical opinion

> *In the meeting you confirmed that personal data would be stored be destroyed after 6 months but anonymous data would be stored for 5 years. The Committee was happy with this.*

The members of the Committee present gave a **favourable ethical opinion** of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

For NHS research sites only, management permission for research ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk. Where the only involvement of the NHS organisation is as a Participant Identification Centre, management permission for research is not required but the R&D office should be notified of the study. Guidance should be sought from the R&D office where necessary.

Sponsors are not required to notify the Committee of approvals from host organisations.

1. **The information sheet should refer to South Yorkshire REC, not "The REC".**

It is responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

You should notify the REC in writing once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised documentation with updated version numbers.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Investigator CV	H Green	20 June 2010
Investigator CV	M Barkham	16 June 2010
Protocol	3	
REC application		25 June 2010
Interview Schedule/Topic Guides	PWP V1	22 June 2010
Interview Schedules/Topic Guides	Supervisor V1	22 June 2010
Participant Information Sheet PWP	1	22 June 2010
Participant Information Sheet: Supervisor	1	22 June 2010
Participant Consent Form: Supervisor	1	22 June 2010
Participant Consent Form: PWP	1	22 June 2010
Questionnaire: Rational Experimental Inventory		
Questionnaire: Psychosocial Inventory of Ego Strengths		
Covering letter on headed paper		22 June 2010

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Service website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

10/H1310/56	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project

Yours sincerely

Joan Brown

Joan Brown
REC Co-ordinator
On Behalf of
Miss Jo Abbott
Chair

Email: joan.brown@leedspft.nhs.uk

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*
"After ethical review – guidance for researchers"

Copy to: *Lauren Smaller*
Academic Service
University of Sheffield
New Spring House
231 Glossop Road
Sheffield S10 2GW

Appendix XXI – Multi-level Model 1: PHQ-9

Copy

	PHQ-9 Single Level Regression	S.E.	PHQ-9 Single Level Regression with GAD-7 Prescore	S.E.	PHQ-9 Single Level Regression with Interaction	S.E.	PHQ-9 Random Intercept	S.E.	PHQ-9 Random Slopes	S.E.
Response	PHQlast		PHQlast		PHQlast		PHQlast		PHQlast	
Fixed Part										
cons	9.832	0.177	9.832	0.177	9.395	0.215	8.592	0.439	8.664	0.437
PHQfirst-gm)	0.618	0.028	0.563	0.038	0.564	0.038	0.559	0.036	0.549	0.046
GADfirst-gm)			0.090	0.044	0.113	0.045	0.115	0.042	0.117	0.041
GADfirst-gm)					0.018	0.005	0.014	0.005	0.012	0.005
Random Pa										
Level: cons										
Level: CLIEI										
cons/cons	35.318	1.491	35.187	1.491	34.792	1.469	29.771	1.268	29.122	1.250
Level: PWP										
cons/cons							2.818	1.147	2.779	1.131
rst-gm)/cons									0.074	0.074
PHQfirst-gm)									0.013	0.009
oglikelihood:	7183.336		7179.191		7166.516		7024.593		7010.868	
DIC:										
Units: cons	1		1		1					
Units: CLIEI	1122		1122		1122		1122		1122	
Units: PWPi							21		21	

Appendix XXII – Multi-level Model 2: GAD-7

Copy

	GAD7 Single Level Regression	S.E.	GAD7 Single Regression with PHQ prescore	S.E.	GAD7 Single Regression with Interaction	S.E.	GAD7 Random Intercept	S.E.	GAD7 Random Slope	S.E.
Response	GADlast		GADlast		GADlast		GADlast		GADlast	
Fixed Part										
cons	8.990	0.160	8.990	0.159	8.522	0.192	7.754	0.393	7.805	0.393
GADfirst-gm)	0.596	0.029	0.490	0.040	0.514	0.040	0.514	0.037	0.478	0.048
PHQfirst-gm)			0.132	0.035	0.132	0.034	0.129	0.032	0.131	0.032
PHQfirst-gm)					0.019	0.004	0.015	0.004	0.014	0.004
Random Pai										
Level: cons										
Level: CLIEI										
cons/cons	28.817	1.217	28.447	1.201	27.995	1.182	23.988	1.022	23.436	1.006
Level: PWP										
cons/cons							2.256	0.919	2.262	0.917
rst-gm)/cons									0.112	0.073
GADfirst-gm)									0.015	0.010
glikelihood:	6955.095		6940.593		6922.627		6782.139		6765.856	
DIC:										
Units: cons	1		1		1					
Units: CLIEI	1122		1122		1122		1122		1122	
Units: PWPi							21		21	

Appendix XXIII – Examples of High Order and Lower Order Themes of PWPs and Supervisors in the Upper and Lower Quartiles

Section 1: PWP Responses

High order theme	Lower order theme	Quote
ENGAGEMENT WITH SUPERVISION TO IMPROVE SKILLS		
Upper Quartile	Checking decisions (n=3)	"I find that so helpful and valuable, going through cases every week [in case management] to begin with it was more asking that the clinical decisions were correct....whereas now I feel more confident" (A) "Case management is really useful to talk through everyone that you're seeing and get some reassurance that you're doing the right thing and there's nothing you've missed" (C) "My supervisor has got a lot more years experience than me so I tend to use her for cases where I'm not so sure....saving the time for the cases where I'm not so sure" (E)
	Being prepared and organised (n=2)	"I think you've got to be organised for it" (B) "I usually [pull together information on] who I'm currently working with, recently assessed, recently discharged [other information including outcome scores] and present a copy of that every case management supervision to my supervisor and highlight which ones I'd like to discuss" (B) "With the case management it's just a question of being organised with the format of the PWP supervision, taking cases in that order" (E)

	Process supervision (n=3)	"With clinical supervision I try to get the most out of it through thinking of different ways we can use it, like by having case discussions, case presentations....role plays and things" (A) "Clinical [supervision] is quite useful [for] if there's a particular case I want to discuss in depth" (B) "[Clinical supervision] has been useful in helping me become more reflective and to work as effectively as I can" (B) "Highlighting the people who aren't engaging and discussing the reasons for that....is useful in case management" (C)
	Utilising supervisor's knowledge (n=2)	"I try and take her advice on board, because she's quite experienced and will sometimes suggest things I've not thought of and I'll think oh that's a good suggestion" (B) "Getting her high intensity perspective on it and taking it down to low intensity" (E)
Lower Quartile	Checking decisions (n=3)	"It's good to bring up even stuff you are comfortable with just to check if there's anything I've missed, or there's another way of looking at a case or formulation" (Z) "[case management] is about getting through as many patients as possible in a short amount of time, which is basically just about checking out risk issues and checking you're offering the right treatments" (X) "I think [case management] is really really helpful just because I obviously get to speak to a more experienced practitioner and its helpful to use them as a sounding board" (V)

	Specific clinical questions (n=3)	"I've got loads of guidance on who I shouldn't be seeing and who needs stepping up, and about things about the disorders we didn't look at uni" (Y) "there are sometimes when you can feel more like this patients got really complex problems and I'm not sure where to send them so supervision's been really helpful for that" (X) "If I do have questions about what I'm doing I can talk to my case management supervisor" (W) "it's especially helpful what you've got patients who you might not have encountered before and need some extra help with" (V) "I ask specific questions about patients with their treatment" (V)

Section 2: Supervisors responses

High order theme	Lower order theme	Quote
EFFECTIVE COMPONENTS OF THERAPEUTIC DELIVERY		
Upper Quartile	Interpersonal skills (n=3)	"She's got a very rapid ability to engage with patients" (C) "I know this from listening to recordings and observed practice, she's got an ability to engage with patients but at the same time keep the professional distance in from patients, she encourages....people to share information with her" (C) "She's got a clear ability to engage with a variety of patients from all different walks of life as well as an ability to engage with senior officers, GPs, etc which is obviously important" (B) "She's very focussed on the patient, they can pick that up from what she's talking about." (D)
	Clinical Skills (n=3)	"[she's good at] the listening and reflecting back, you know all the basic communication skills for someone who's doing that job." (A) "She's got very good listening skills, she's good at empathising with clients, she's good at making summaries and feeding back to clients and checking out information with them" (B) "She's very good at history taking and getting information out of people and uses open ended questions than other people that I might supervise" (B) "She's very good at setting goals with them and realistic goals" (D)

	Knowledge and Understanding (n=2)	"With her being analytic, having a good knowledge of theory and a good knowledge of the reach out materials, they're the key things that help her" (B) "Her knowledge base [makes her effective]" (E)
Lower Quartile	Interpersonal Skills (n=3)	"she is good at putting people at ease and trying to think from the other person's perspective" (Z) "Being down to earth in a sense.....she builds [good] rapports with clients so there's no sense of hierarchy or judgemental behaviours at all, she builds rapport with clients of different ages and backgrounds" (Y) "She has a very good therapeutic alliance" (X)
	Clinical Skills (n=3)	"She does use quite an empathic approach" (Z) "[she is good at] putting things in a language that the other person understands" (Z) "[she has] a lot of empathy" (X) "She thinks about his intervention and assessment and the detail, she doesn't jump in with both feet" (V)