Published in final edited form as:

J Clin Psychiatry. 2013 May ; 74(5): 482–491. doi:10.4088/JCP.12r07848.

Can Psychotherapists Function as Their Own Controls? Meta-Analysis of the “Crossed Therapist” Design in Comparative Psychotherapy Trials

Fredrik Falkenström, PhD,

Department of Behavioral Sciences, Linköping University, Linköping, Sweden Centre for Clinical Research Sörmland, Uppsala University, Uppsala, Sweden

John C. Markowitz [Professor],

New York State Psychiatric Institute, New York, NY, Columbia University College of Physicians & Surgeons, New York, NY

Hanske Jonker, MSc,

Faculty of Social Sciences, University of Utrecht, Utrecht, Holland, New York State Psychiatric Institute, New York, NY, Columbia University College of Physicians & Surgeons, New York, NY

Björn Philips, PhD, and

Department of Behavioral Sciences, Linköping University, Linköping, Sweden, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden, Center for Dependency Disorders, Stockholm County Council, Stockholm, Sweden

Rolf Holmqvist [Professor]

Department of Behavioral Sciences, Linköping University, Linköping, Sweden

Abstract

Objective—Clinical trials sometimes have the same therapists deliver more than one psychotherapy, ostensibly to control for therapist effects. This “crossed therapists” design makes controlling for therapist allegiance imperative, as therapists may prefer one treatment they deliver to the other(s). Research has established a strong relationship between principal investigators’ allegiances and treatment outcome. Study therapists’ allegiances probably also influence outcome, yet this moderating factor on outcome has never been studied.

Data Sources—English language abstracts in Psychinfo and MedLine from January 1985 to December 2011 were searched for keywords “psychotherapy” and “randomized trial.”

Study Selection—The search yielded 990 abstracts that were searched manually. Trials using the same therapists in more than one condition were included.

Data extraction—Thirty-nine studies fulfilled inclusion criteria. Meta-regression analyses assessed the influence of researchers’ allegiance on treatment outcome, testing the hypothesis that studies poorly controlling for therapist allegiance would show stronger influence of researcher

Corresponding author: Fredrik Falkenström, Lustigkullevägen 17, SE-616 33 Åby, Sweden, Fredrik.Falkenstrom@liu.se.

Financial disclosures: This study received no direct funding. Fredrik Falkenström receives salary from the Sörmland County Council. Dr. Markowitz gets salary support from New York State Psychiatric Institute; small book royalties from American Psychiatric Press, Basic Books, and Oxford U. Press; and an editorial stipend from Elsevier Press. He has grant support from NIMH (grant MH079078). Hanske Jonker is a student working part time as a secretary at HSK group and previously at Altrecht as secretary. Dr. Philips receives salary from Linköping University and Stockholm County Council, and small book royalties from Liber. He has grant support from the Swedish Council for Working Life and Social Research (grant 2007-0457) and from Karolinska Institutet. Dr. Holmqvist gets salary from Linköping University and a small book royalty from Liber. He has grant support from the Swedish Council for Working Life and Social Research (grant 2008-0149).

allegiance on outcome. A single-item measure assessed researchers' reported attempts to control for therapist allegiance.

Results—Only one (3%) of 39 studies measured therapist treatment allegiance. Another five (13%) mentioned controlling for, without formally assessing, therapist allegiance. Most publications (64%) did not even mention therapist allegiance. In studies not controlling for therapist allegiance, researcher allegiance strongly influenced outcome, whereas studies reporting control for therapist allegiance showed no differential researcher allegiance. Cognitive-behavioral trials less frequently described controlling for therapist allegiance.

Conclusions—The “crossed therapist” design is subject to bias due to differential psychotherapist allegiance. Worrisome results suggest that researchers strongly allied to a treatment may ignore therapist allegiance, potentially skewing outcomes. All clinical trials, and especially “crossed therapist” designs, should measure psychotherapist allegiance to evaluate this possible bias.

One of the sacrosanct assumptions of a client is that their therapist believes in the treatment being delivered.

-- Wampold, 2001 (1, p159)

Introduction

Some psychotherapy researchers have designed clinical trials to “cross” therapists with treatments: that is, the same therapists deliver more than one type of psychotherapy. The rationale is that using the same therapist across therapies controls for therapist effects on treatment outcome. The design has attracted surprisingly little critical attention in the scientific literature. This article reviews studies that have used this design and discusses important research implications.

Researchers agree that individual therapists may generate different outcomes with patients, some having consistently greater efficacy than others (2,3), although the size of this therapist effect has stirred debate (4,5). In randomized clinical trials (RCTs) where therapists are trained using a manual for a specific disorder, between-therapist variation is likely smaller than in general practice (6). Nonetheless, reviews of psychotherapy research recommend controlling for therapist effects by including the therapist factor as a random term in the statistical analysis of any clinical trial, regardless of whether the therapist effect reaches conventional significance levels. The reasoning is that even small therapist effects increase the type 1 error rate of the between-treatment comparison if the therapist factor is omitted from the statistical model (7). Crits-Christoph and Mintz (6) recommend using an alpha level of .2 to .3 (rather than the usual .05) as the criterion for deciding to exclude the therapist factor in the main outcome analysis of a clinical trial.

The RCT is considered the optimal design to ensure internal validity of clinical trials, reducing the influence of variables other than the one(s) experimentally manipulated (8,9). Medical and psychotherapy outcome research deems this design the ‘gold standard.’ Psychotherapy RCTs most commonly compare different teams of (ideally) similarly experienced, equally expert therapists providing the different study treatments -- a practical approach, as there might not be so many therapists who are able to deliver more than one therapy correctly and competently. (As far as we know this latter issue has never been studied but it is our personal experiences that it is difficult to learn more than one psychotherapy well.) Elkin (10) enumerated the difficulties of this modal design, which “nests” therapists within treatments: i.e., different teams of therapists provide the different treatments. This design potentially confounds treatment efficacy with therapist efficacy.

Elkin described the difficulties of ensuring similar characteristics, experience, competence, training, and supervision of the competing teams of therapists.

The alternative design “crosses” psychotherapists over treatments: the same (“crossed”) therapists provide both the experimental and comparison treatment(s) (6). Some advocates of this design have argued that it controls for therapist effects (11–13): the therapist provides his or her own control. Although this design might seem to disentangle therapists from treatments, the need persists to ensure that therapists have comparable training, experience, and competence in both experimental and comparison treatments. As several authors have noted, the “crossed” design raises additional problems: contamination between treatments, and differential therapist allegiance to the treatments (6,7,10,12).

Contamination involves therapists using techniques from one treatment in the comparison condition, which blurs distinctions between treatments and should skew results toward finding no treatment differences. Still more crucial, therapeutic allegiance reflects therapist belief that one psychotherapy works better than another. We have uncovered no empirical research studying the connection between *therapist* allegiance and outcome. Several reviews, however, show that *researcher* (principal investigator) allegiance to a treatment correlates strongly with study outcome (14–16). This is a correlational finding, but if the relationship is indeed causal, one way that researcher allegiance could affect outcome is through researchers selecting study therapists who share the researcher’s allegiance. There are other ways the researcher’s allegiance could affect outcome; for example, by selecting ineffective comparison treatments or choosing not to publish studies that run counter to their allegiance (14). Researcher allegiance and therapist allegiance are thus different but most likely interdependent concepts.

Study therapists who value one treatment above another would likely communicate their bias to their patients, affecting both technical and nonspecific aspects of treatment: both treatment delivery and the conviction with which they deliver each therapy. Therapists may have greater natural affinity for and skill in one treatment than another. Therefore, even absent empirical data, it is plausible to hypothesize that differential therapist allegiance in a “crossed” therapist design may artificially enlarge between-treatment differences. Differential allegiance and/or competence might also contribute to smaller between-treatment differences if therapists deliver an ineffective treatment more enthusiastically and competently than an inherently more potent treatment. The well established relationship between research allegiance and good outcome suggests this is rarer, however. Contamination and differential allegiance could still affect a study in which distinct teams of therapists provide the different treatments. Contamination of techniques from other therapeutic schools could occur if study therapists have previous training in other therapies, and differential allegiance could occur if one therapist team feels greater enthusiasm for its treatment model than another.

Therapist allegiance may thus constitute a moderating variable, comparable to variables such as therapist experience and training, which all outcome studies should consider. Nonetheless, the demanding arrangement of switching back and forth between competing models renders the “crossed” design particularly vulnerable to criticism. Indeed, having the same therapists deliver multiple modalities may obscure potential therapist bias. By using the “crossed design” to ostensibly control for therapist factors, researchers may ignore important allegiance differences that may receive greater scrutiny in studies comparing “nested” competing therapist teams, even though paradoxically the latter probably have less variability in therapeutic allegiance.

As the literature has never addressed this issue, we review studies that used the “crossed” therapist design to study how researchers have addressed therapist allegiance. A preliminary qualitative review (unpublished; available from first author upon request) revealed that therapist allegiance was seldom if ever measured. We therefore chose as our primary research question whether researchers reported having *controlled for* therapist allegiance, and whether not controlling for therapist allegiance affected treatment outcome. We hypothesized a stronger relationship between researcher allegiance and study outcome when therapist allegiance was poorly controlled. Because methodological quality of RCTs has evolved (17), we also explored whether control for therapist allegiance increased over time.

Method

Selection of studies

We searched Psycinfo and MedLine (EBSCO) for abstracts containing the terms “psychotherapy” and “randomized trial.” The search, limited to English language articles published in peer-reviewed journals between 1985 and 2011, yielded 990 hits (652 in Psycinfo, 338 in MedLine) that the first author (FF) scrutinized. He searched all abstracts and when necessary article method sections to determine whether the RCT used the same psychotherapists in more than one treatment condition. Requests sent to two psychotherapy research mailing lists sought leads on “crossed” studies. One was the email list of the Society for Psychotherapy Research, an international multidisciplinary organization devoted to psychotherapy research of diverse theoretical orientations. The second was the “Psychodynamic research” listserv, open to people actively associated with psychodynamic research in an academic or clinical practice setting. Tips from these two mailing lists yielded two studies (18,19). Incorporating our own and colleagues’ knowledge of crossed therapist studies added another three studies (20–22).

The search excluded trials of component and dismantling studies within the same general therapy orientation, as differential competence and allegiance would likely be less problematic in such studies. We included, however, component studies comparing models that are recognized as treatments in their own right: e.g., comparisons between cognitive and behavioral treatments, because although the two models are often combined as cognitive-behavioral therapy (CBT), both therapists and researchers may have strong allegiances in the continuing debate over the relative contributions of cognitive and behavioral techniques to CBT outcome. We excluded studies of the same therapy model delivered in different formats (e.g., individual versus group format) because allegiance likely differs more between competing theoretical models than between different delivery formats. Finally, we excluded studies published before 1985 because secular improvement in RCT methodological quality (17) suggested that older studies would lack comparability to more modern ones.

This search yielded 39 trials published between 1986 and 2011, which we included in the study (11–13, 18–54).

Measures

Control for therapist allegiance (CTA)—The first author conducted a qualitative review of the articles. This review provided the basis for developing a measure for rating control for therapist allegiance in RCTs, the Falkenström Allegiance Control for Therapists measure (FACT). The FACT comprises a single item measuring researcher control for therapist allegiance on a scale from 0 to 3:

- 0 The article does not mention the issue of potentially differing therapist allegiances. If the authors mention the “crossed” therapist design at all, they do so only as a blanket disclaimer of therapist effects.
- 1 The issue of potentially different allegiances is mentioned but not assessed or controlled for in any way.
- 2 The authors report making some attempt to control for therapist allegiance, such as trying to recruit therapists who informally deny therapy preferences. However, they make no attempt to measure therapist allegiance or to statistically control for this variable.
- 3 The authors measure therapist allegiance and, if necessary, apply relevant statistical control for this variable.

Researcher allegiance (RA)—We measured researcher allegiance using the methodology most prevalent in published studies (e.g., 14,55). The method entails scoring researcher allegiance from the way published papers are written, using manualized scoring criteria. Three raters (FF, HJ, and BP) read published articles, scrutinizing especially the Introduction and Methods sections. The scale for researcher allegiance was adapted from Gaffan et al. (55). They developed the following criteria, which should be viewed as indicators rather than absolute rules:

- 0 No allegiance:
 - a. No evidence is presented that the treatment is effective
 - b. The authors simply mention that the treatment will be included, without elaboration
- 1 Weak allegiance:
 - a. Authors describe previous research indicating effectiveness of the treatment, but the results are described as mixed or give no indication that the treatment should be effective for the population tested in the study
 - b. The authors state a hypothesis with mixed predictions: e.g., the treatment is expected to benefit some patients but not others
- 2 Moderate allegiance:
 - a. Reference is given to previous research indicating efficacy of the treatment relative to no treatment
 - b. Evidence is presented from the literature that the treatment is effective for this population of patients
 - c. Statements showing that the author(s) believes the treatment is effective or widely approved
 - d. A short but clear rationale for the treatment’s procedure.
- 3 Strong allegiance:
 - a. Reference to published research showing superiority of the treatment to some other treatment
 - b. A specific hypothesis is presented as to why the treatment should be superior to the other treatment(s) included in the study

- c. A detailed description (at least 10 lines) explains the treatment's procedures and aims
- d. The treatment was developed or first introduced by one of the authors (in our scoring this immediately gave an allegiance score of 3)

Following Gaffan et al. (55), we converted the ratings to relative allegiance scores by subtracting the lowest score from the higher ones within each study. For example, if treatment A and treatment B both had researcher allegiance scores of 3 (high allegiance), both would receive relative allegiance scores of 0. If a third treatment C in the same study had an absolute allegiance score of 1, the relative allegiance scores would be 2 for treatments A and B, and 0 for C.

Treatment dose and therapist characteristics—We rated whether researchers had ensured that all treatments were delivered in equal dosage, and whether therapists in all conditions had equal training, experience, and supervision. These were single-item variables with simple yes/no/not reported alternatives. Training was defined not only by the training the researchers provided for the study, but by all training therapists had had in the respective treatment modalities.

Effect sizes—The effect sizes were between-groups standardized mean change scores, Δ (56), calculated by first taking the mean pre-post difference divided by the pretreatment standard deviation (d) separately for treatment and control groups. The variance was then calculated as:

$$V_d = \left(\frac{1}{n} + \frac{d^2}{2n} \right) \cdot 2(1-r)$$

where n is the group size and r the correlation between pre- and post measurements. Because the latter was almost never reported, we assumed a pre-post correlation of .60 for all measurements in all studies. Both d and V_d were then corrected for sample size bias:

$$\begin{aligned} g &= J \cdot d \\ V_g &= J^2 \cdot V_d \\ J &= 1 - \frac{3}{4(df)-1} \end{aligned}$$

Finally the between-groups standardized mean change scores were calculated as:

$$\begin{aligned} \Delta &= g_T - g_C \\ V_{\Delta} &= V_{g_T} + V_{g_C} \end{aligned}$$

Nine studies reported some measures as proportions rather than continuous measures ($n=15$ measures of 94). In these cases Cohen's h was used (57). Four studies provided no information for calculating effect sizes for between-group contrasts that were statistically non-significant ($n=14$ measures of 27). In these cases we assumed zero difference between groups. This was unfortunate, as the effect sizes were likely not actually zero; yet omitting these non-significant measures would have further distorted our data.

Two studies had to be excluded from the meta-regression analyses: one (34) because the effect size used was time-to-event (hazard ratio) from a survival analysis, which was not

deemed comparable to Becker's Δ in the other studies. In the other study (48) patients crossed over between therapies at mid-treatment, confounding interpretation of effect sizes.

Procedure

The number of comparisons for each study was the number of treatment groups minus one. A study comparing two treatments had one comparison; a study comparing three treatments had two comparisons, etc., yielding one between-group effect size per comparison. Effect sizes were scored using the group scoring lowest on researcher allegiance as the comparison group. When allegiance did not differ between groups, experimental/comparison group status was arbitrary. The relative allegiance score for the comparison group, being redundant, was not used.

Reliability

The first author (FF) rated all variables for all studies. Two other raters (HJ and BP) rated all variables for 33 and 6 studies, respectively. Inter-rater agreement was calculated using intra-class correlation (two-way random effects model) for continuous variables, Spearman's ρ for ordinal variables, and Cohen's $Kappa$ for nominal variables.

Statistical analyses

A problem in meta-analysis is that some data are statistically non-independent. Most studies used more than one outcome measure, with the result that outcome measures used for the same patients were dependent on each other. Several studies compared more than two therapies, creating additional non-independence in data.

Fortunately, recent advances in statistical methodology provide solutions to this problem. Multi-level meta-analysis (58) enables explicit modeling of the co-variances on both within- and between-study levels. However, this method ordinarily requires knowledge of the within-study intercorrelations (i.e., correlations between different measures used), information seldom if ever available to a meta-analysis. To address this problem, Hedges et al. (59) created a theorem allowing computation of regression estimators in meta-regression that are robust for violations of the non-independence assumption. This theorem computes fixed regression estimators that take random effects (between-studies variation) into account. Being most concerned about within-study dependence of effect sizes, we chose the correlated effects model. The value of ρ , the mean intercorrelation of effect sizes, was first set to .5, deemed a reasonable guess of the values of ρ . We then ran sensitivity tests assigning different values to ρ between .1 and .9. Demonstrating the robustness of the approach, these different values of ρ had little effect on the estimated regression coefficients.

Computations used the Stata 11 macro *robumeta* to compute robust standard errors in meta-regression (60). Despite conducting multiple statistical significance tests, we chose not to control for family-wise type I error. This first study on this subject was by nature exploratory, and the multiple tests mostly tested different research questions. The use of corrections for family-wise type I error has been questioned (61).

Results

Table 1 lists all 39 included studies.

Reliability of ratings

Reliability estimates were good to excellent for the main predictor variables. For FACT, Spearman ρ was .91. For the meta-regression analyses, we collapsed the FACT into a

dichotomous scale where 0 meant no control for therapist allegiance (categories 0 and 1 from the full scale) and 1 meant some control for therapist allegiance (categories 2 and 3). The rationale for this decision was that mere mention that the same therapists do both treatments (category 1) probably does not influence the effect of researcher allegiance on outcome. Only one study met category 3, raising concern that this single trial would excessively influence the statistical analyses. The collapsed scale had reliability somewhat lower than the original, but still acceptable ($Kappa = .72$). Relative researcher allegiance had an Intraclass Correlation of .82 for the ratings of a single rater, which was considered good reliability. The ratings of the first rater (FF) were used in the final analyses.

For Treatment Dose, Therapist Training, and Supervision reliability was adequate to good ($Kappa = .82, .64, .77$, respectively). For Therapist Experience reliability was low ($Kappa = .53$), most likely because the initial scoring criteria were unclear. Because all disagreements arose between rater 1 and 2, rater 3 rated all disagreements and we used these ratings as the final score for these studies.

Descriptive statistics

No studies overlapped with the earlier meta-analysis by Luborsky et al. (14), and only three with Gaffan et al. (55). Median publication year was 2002. Thirty studies (83%) included a cognitive, behavioral, or cognitive-behavioral therapy as at least one treatment condition. The 39 studies yielded 48 treatment comparisons. Among these 48 comparisons, ten provided no way of ascertaining which treatment was intended as experimental and which as control treatment. Among the remaining 38 comparisons, the most common experimental treatment was CBT ($n = 18$). Including variants of CBT like Behavior Therapy (BT), Cognitive Therapy (CT), or self-management therapy increased that figure ($n = 29$). The most common comparison treatments were supportive ($n = 14$), client-centered ($n = 5$), and relaxation ($n = 5$) therapies.

Therapist allegiance

Figure 1 shows the FACT ratings for all 39 studies. Only one (3%) study (49) earned the maximal FACT rating, i.e. addressing measurement of therapist allegiance and its statistical control if needed. Another five studies (13%) reported having attempted to recruit therapists with balanced allegiances as a way of controlling for therapist allegiance, although none had actually measured therapist allegiance. The remaining studies either did not mention therapist allegiance at all ($n = 26$; 67%) or mentioned it only as a study limitation or to dismiss its importance ($n = 7$; 19%).

Comparing studies with researcher allegiance favoring CBT to those lacking such allegiance on the (collapsed) FACT scale indicated that CBT-allegiant researchers controlled for therapist allegiance less often than researchers without CBT allegiance (Fisher's exact test $p = .002$).

Researcher allegiance predicting outcome

We first ran a meta-regression analysis with RA as the sole predictor. This revealed a statistically significant effect of researcher allegiance on outcome ($b = .10$, $se = .02$, $t = 3.87$, $p < .001$, 95 % CI: .05, .15). Each point increase in relative allegiance increased the between-groups effect size by .10.

Control for therapist allegiance predicting outcome

We next tested the direct effect of CTA (collapsed FACT) on outcome. This effect was not significant ($b = -.07$, ns). Yet only five studies reported controlling for allegiance, limiting statistical power.

Control for therapist allegiance moderates relationship between researcher allegiance and outcome

Our primary hypothesis was that RA more strongly affects outcome in studies that do not control for therapist allegiance (CTA). We added an interaction term ($RA \times CTA$) to the statistical model that contained main effects of RA and CTA on outcome. This model proved impossible to run, however, as none of the studies that had attempted control for therapist allegiance had been scored as having any differential researcher allegiance. The interaction term was therefore almost completely collinear with the original CTA variable (i.e. $RA \times CTA = CTA$).

Exploring this further revealed a large difference in researcher allegiance between the studies that had not attempted to control for therapist allegiance ($n = 32$) and those reporting having attempted some control ($n = 6$). The former group had a mean relative allegiance score of 2.1 ($sd = 0.97$), whereas the latter group had a mean of 0 (with no variation). This difference was highly statistically significant: $t(32) = 12.2, p < .001$.

Publication year

Testing whether recent studies controlled for therapist allegiance more than older studies unexpectedly showed a significant relationship opposite from the expected direction: more recent studies controlled *less* for therapist allegiance ($\rho = -.33; p = .04$). Using the dichotomized therapist allegiance control variable yielded an almost identical result. Further exploration revealed that five of the six studies that controlled for therapist allegiance were published between 1987 and 1996 (the sixth was published in 2007).

Reporting of dose and therapist characteristics

Figure 2 shows reporting of equal therapy dose, therapist training, experience, and supervision. Most studies reported having ensured equal dose (85%) and amount of supervision (69%) in all treatments, but most studies inadequately reported therapist training and experience. About two-thirds of the studies did not report how much training and experience the therapists had in each of the treatments they delivered (64% and 72%, respectively). Six studies reported higher treatment dose, six more therapist experience, and five more training in the treatment scoring higher on researcher allegiance.

Discussion

Psychotherapy is a complex enterprise, involving many factors: patient, therapist, treatment factors, and complex interactions between all these. The randomized controlled trial design is ideally suited to isolating the treatment factor: when patients are randomized to treatments, confounding factors should be equally distributed between groups. However, because it is usually not feasible (and probably not desirable) to randomize therapists to treatments, the problem of confounding therapist and treatment effects remains.

Some researchers have advocated solving this problem by using the same psychotherapists to conduct more than one study treatment. Despite the salience of this design issue, no research has previously reviewed trials with the “crossed therapist” design. Our main finding was that the great preponderance of the 39 such studies we collected failed to report the key issue of therapist treatment allegiance. Despite the glaring need in a crossed design to ascertain therapist allegiance to the respective treatments, less than half (36%, $n = 14$) the trials reviewed even mentioned therapist or researcher allegiance. Still fewer (13%, $n = 5$) explicitly reported attempting to control for, and only one (3%) actually measured, therapist allegiance. About two thirds of the studies surveyed did not report therapists’ previous training or experience in the treatments they provided. Most studies included involved CBT

as a treatment (87%, $n=34$), and researchers with CBT allegiance more often used this design without controlling for therapist allegiance. Thoma et al. (62) found that the average CBT trial of depression scored at the lower range of adequate methodological quality and that CBT trials did not on average show better quality than psychodynamic trials

Although most researchers did not even mention therapist allegiance as a possible bias, some acknowledged that their study therapists believed the comparison treatment to be ineffective and that this may have biased results:

“The choice of the same therapists for both treatments created a twofold problem. On the positive side, the effect of the same therapist was constant across the two therapeutic conditions. On the negative side, all were CBT-oriented therapists and this may have biased the treatment in favor of CBT as the therapists were requested to use methods they judged as noneffective.” (11, p. 108)

Our analyses, corroborating earlier research (14,55) while surveying largely different outcome trials, showed that researcher allegiance clearly influenced treatment outcome. Our main research question was whether this effect diminished in studies that controlled for therapist allegiance. Interestingly, we found no differential researcher allegiance in any of the studies that had reportedly attempted some control for therapist allegiance. This sharply contrasted with the studies not attempting such control, for which mean researcher allegiance score was 2.1 on a scale ranging from 0 to 3. The stronger the researchers' allegiances to one psychotherapy, the more likely they appeared to ignore therapist allegiance. If this indeed indicates how research using the crossed therapist design is being conducted, it is remarkable and worrisome that researchers strongly allied to one treatment consciously or unconsciously overlook the potential bias of therapist allegiance.

Some researchers may argue that treatment integrity checks provide an objective assessment of whether therapists deliver treatment as it is meant to be delivered, hence integrity checks should reveal therapist allegiance effects as problems in the delivery of treatment. Although our review did not focus on integrity, we did note that almost seventy percent ($n=27$) of the included studies reported formal adherence analyses, and only one of these reported an adherence problem. Treatment integrity should ideally include ratings of therapist competence. Assessing competence is difficult, however, and a recent meta-analysis showed no significant relationship of competence ratings to outcome (63). We doubt that integrity analyses based on fairly coarse measures will detect more subtle effects of therapists who technically adhere to the treatment protocol but do not deliver treatment with the same enthusiasm as therapists who really believe in their model.

Our analyses suggested that controlling for therapist allegiance has fallen out of fashion: five of the six studies that controlled for therapist allegiance were published between 1987 and 1996, and only one since. Our impression is that the psychotherapy research community has seldom discussed the issue. Therapist allegiance remains a crucial unstudied factor in psychotherapy research. Strength of belief in a therapy may affect the therapist's comfort and authenticity in conducting treatment, the therapy's plausibility for the patient, and thus the strength of the therapeutic alliance. This matters when a therapist conducts a single therapy, and still more when the therapist delivers more than one modality.

If therapist allegiance influences therapeutic outcome, then the “crossed therapist” design may prove a double-cross: while ostensibly controlling for therapist factors between therapists, it may merely obscure them within the same therapist. Although grant reviewers frequently suggest controlling for therapist effects by “crossing” therapists, the research supporting this injunction is essentially non-existent. Does choosing a “crossed therapist” design solve the problem of confounding therapist effects with treatment effects? The

answer depends on what kinds of therapist characteristics affect outcome. If general personality characteristics (warmth, empathy, interpersonal skills, etc.) that could reasonably be considered independent of treatment type matter most for patient outcome, the crossed design might control for those effects. If therapist effects rather reflect competence in or enthusiasm for a particular treatment model, however, the alternative design of “amicably competitive” balanced therapist teams (64) is a better choice. In this design, equally experienced, competent, and adherent teams of psychotherapists compete in friendly rivalry, demonstrating in effect the optimal effects of a given intervention. Rather than these therapists attempting multiple competencies, their goal is expertise in their assigned modality, with variables such as experience and competence counterbalanced across groups.

One could argue that researcher allegiance will out, regardless of design: the alternative design structure of nested “amicably competitive” expert therapist teams (64) could be equally compromised if the principal investigator chooses less competent or experienced therapists for comparator conditions. We recommend that all clinical trials (“crossed” or “nested”) measure and report therapist allegiance, at least until more is known about the effect of therapist allegiance on the process and outcome of psychotherapy. To our knowledge, no instrument for measuring therapist allegiance has yet been published. In our own ongoing studies we have developed simple self-report instruments for this purpose.

Limitations

This is a new field of research. We sought but may not have located all recent studies using the “crossed therapist” design. How common this approach is relative to “nested” competitive therapy teams delivering single modalities is unclear; the latter seems more prevalent. The relatively small number of studies included limits the conclusions that can be drawn, especially about the effect of controlling for therapist allegiance as this occurred in only six of the studies. A key limitation (and key study finding) is the absence of therapist allegiance ratings in the psychotherapy trials under study.

Recommendations for future research

Researchers should rigorously report therapists’ backgrounds when using the crossed therapist design, clarifying the training and experience study therapists have in conducting each of the different treatments they use in the trial, and not just trial-related training but also previous training and experience. Researchers should report whether they have tried to statistically control for therapist effects (including allegiance) in their main outcome analyses; and if not, why not (i.e., what significance criterion they used in preliminary analyses to rule out therapist effects). A “crossed therapist” design, if employed, should control for therapist x treatment interaction effects, not just therapist main effects. (Some researchers infer an absence of therapist effects in their study when they find no differences in outcome between individual therapists (e.g., 20). Yet finding no therapist differences in outcome does not rule out the possibility that all therapists deliver one treatment better than the other either due to either greater competency or allegiance to that treatment.)

All clinical trials should routinely measure therapist allegiance and relate it to outcome. The FACT could be added to existing RCT quality measures (17) in reviews and meta-analyses of RCTs.

Given the potential problems of the “crossed therapists” design, why use it? We recommend that in standard efficacy trials, where the predominant question is differential efficacy between two treatments, the “crossed design” be considered a weaker alternative to the “amicably competitive” therapist teams approach. The “crossed therapists” design could have utility for studying interactions between therapist characteristics and treatment method

effects on outcome. Another indication for this design might be practical, where geography or other factors constrain patient access to multiple therapists practicing different treatments. We recommend that researchers avoid “crossed therapists” designs unless the research questions clearly focus on issues that this design facilitates, or unless practicalities such as geographic isolation preclude alternative designs. If used, “crossed therapist designs” must measure therapist allegiance.

Therapist allegiance warrants consideration equal to researcher allegiance in designing and interpreting the results of psychotherapy trials. Therapists in every clinical trial could rate on a simple, anchored Likert-type scale their belief in as well as their self-perceived skill in all the study’s therapies; how well each therapy fits their clinical perspectives; and their prediction of the prospective study outcome. Such evaluations would help to control for potential study biases and might yield interesting findings about psychotherapy outcome trials. Combining such data with psychotherapy process and outcome variables might elucidate the suitability of particular therapists for particular therapies.

Clinical points

- Therapist belief in treatment is likely to be a strong “non-specific” effect of psychotherapy, yet this factor has almost never been studied
- Researcher allegiance may influence study findings in part through selecting biased therapists in “crossed therapist” study designs

Acknowledgments

The authors are indebted to AC Del Re, PhD, for instructions on effect size calculations.

References

1. Wampold, BE. The great psychotherapy debate: Models, methods, and findings. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2001. p. 159
2. Kraus DR, Castonguay L, Boswell JF, Nordberg SS, Hayes JA. Therapist effectiveness: Implications for accountability and patient care. *Psychother Res.* 2011; 21:267–276. [PubMed: 21623550]
3. Wampold BE, Brown GS. Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *J Consult Clin Psychol.* 2005; 73:914–923. [PubMed: 16287391]
4. Elkin I, Falconnier L, Martinovich Z, Mahoney C. Therapist effects in the National Institute of Mental Health treatment of depression collaborative research program. *Psychother Res.* 2006; 16:144–160.
5. Kim D-M, Wampold BE, Bolt DM. Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health treatment of depression collaborative research program data. *Psychother Res.* 2006; 16:161–172.
6. Crits-Christoph P, Mintz J. Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *J Consult Clin Psychol.* 1991; 59:20–26. [PubMed: 2002139]
7. de Jong K, Moerbeek M, van der Leeden R. A priori power analysis in longitudinal three-level multilevel models: An example with therapist effects. *Psychother Res.* 2010; 20:273–84. [PubMed: 19946814]
8. Chambless DL, Hollon SD. Defining empirically supported therapies. *J Consult Clin Psychol.* 1998; 66:7–18. [PubMed: 9489259]
9. Shadish, WR.; Cook, TD.; Campbell, DT. Experimental and quasi-experimental designs for generalized causal inference. Boston, MA: Houghton, Mifflin and Co; 2002.

10. Elkin I. A major dilemma in psychotherapy outcome research: Disentangling therapists from therapies. *Clin Psychol: Sci Pract.* 1999; 6:10–32.
11. Cottraux J, Note I, Yao SN, de Mey-Guillard C, Bonasse F, Djamoussian D, Mollard E, Note B, Chen Y. Randomized controlled comparison of cognitive behavior therapy with Rogerian supportive therapy in chronic post-traumatic stress disorder: A 2-year follow-up. *Psychother Psychosom.* 2008; 77:101–110. [PubMed: 18230943]
12. Staines GL. Comparative outcome evaluations of psychotherapies: Guidelines for addressing eight limitations of the gold standard of causal inference. *Psychotherapy: Theory Res Pract Train.* 2007; 44:161–174.
13. Shapiro DA, Barkham M, Rees A, Hardy GE, Reynolds S, Startup M. Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *J Consult Clin Psychol.* 1994; 62:522–534. [PubMed: 8063978]
14. Luborsky L, Diguer L, Seligman DA, Rosenthal R, Krause ED, Johnson S, Halperin G, Bishop M, Berman JS, Schweizer E. The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clin Psychol: Sci Pract.* 1999; 6:95–106.
15. Luborsky L, Singer B, Luborsky L. Comparative studies of psychotherapies: Is it true that "everyone has won and all must have prizes"? *Arch Gen Psychiatry.* 1975; 32:995–1008. [PubMed: 239666]
16. Robinson LA, Berman JS, Neimeyer RA. Psychotherapy for the treatment of depression: A comprehensive review of controlled outcome research. *Psychol Bull.* 1990; 108:30–49. [PubMed: 2200072]
17. Kocsis JH, Gerber AJ, Milrod B, Roose SP, Barber J, Thase ME, Perkins P, Leon AC. A new scale for assessing the quality of randomized clinical trials of psychotherapy. *Compr Psychiatry.* 2010; 51:319–324. [PubMed: 20399343]
18. Barkham M, Rees A, Shapiro DA, Stiles WB, Agnew RM, Halstead J, et al. Outcomes of time-limited therapy in applied settings: Replicating the Second Sheffield Psychotherapy Project. *J Consult Clin Psychol.* 1996; 64(5):1079–85. [PubMed: 8916639]
19. Luty SE, Carter JD, McKenzie JM, Rae AM, Frampton CM, Mulder RT, et al. Randomised controlled trial of interpersonal psychotherapy and cognitive-behavioural therapy for depression. *Br J Psychiatry.* 2007; 190:496–502. [PubMed: 17541109]
20. Clark DM, Ehlers A, Hackmann A, McManus F, Fennell M, Grey N, et al. Cognitive therapy versus exposure and applied relaxation in social phobia: A randomized controlled trial. *J Consult Clin Psychol.* 2006; 74(3):568–78. [PubMed: 16822113]
21. McIntosh VVW, Jordan J, Carter FA, Luty SE, McKenzie JM, Bulik CM, et al. Three psychotherapies for anorexia nervosa: A randomized, controlled trial. *Am J Psychiatry.* 2005; 162(4):741–7. [PubMed: 15800147]
22. Taylor S, Thordarson DS, Maxfield L, Fedoroff IC, Lovell K, Ogrodniczuk J. Comparative efficacy, speed, and adverse effects of three PTSD treatments: Exposure therapy, EMDR, and relaxation training. *J Consult Clin Psychol.* 2003; 71(2):330–8. [PubMed: 12699027]
23. Agras W, Walsh T, Fairburn CG, Wilson G, Kraemer HC. A multicenter comparison of cognitive-behavioral therapy and interpersonal psychotherapy for bulimia nervosa. *Arch Gen Psychiatry.* 2000; 57(5):459–66. [PubMed: 10807486]
24. Barkham M, Shapiro DA, Hardy GE, Rees A. Psychotherapy in two-plus-one sessions: Outcomes of a randomized controlled trial of cognitive-behavioral and psychodynamic-interpersonal therapy for subsyndromal depression. *J Consult Clin Psychol.* 1999; 67(2):201–11. [PubMed: 10224730]
25. Beck A, Sokol L, Clark D, Berchick R, Wright F. A crossover study of focused cognitive therapy for panic disorder. *Am J Psychiatry.* 1992; 149(6):778–83. [PubMed: 1590494]
26. Bjornsson AS, Bidwell LC, Brosse AL, Carey G, Hauser M, Mackiewicz Seghete KL, et al. Cognitive-behavioral group therapy versus group psychotherapy for social anxiety disorder among college students: A randomized controlled trial. *Depress Anxiety.* 2011; 28(11):1034–42. [PubMed: 22076970]
27. Borkovec TD, Mathews AM, Chambers A, Ebrahimi S, Lytle R, Nelson R. The effects of relaxation training with cognitive or nondirective therapy and the role of relaxation-induced

- anxiety in the treatment of generalized anxiety. *J Consult Clin Psychol.* 1987; 55(6):883–8. [PubMed: 3320121]
28. Borkovec TD, Mathews AM. Treatment of nonphobic anxiety disorders: A comparison of nondirective, cognitive, and coping desensitization therapy. *J Consult Clin Psychol.* 1988; 56(6): 877–84. [PubMed: 2904928]
 29. Castelnovo G, Manzoni GM, Villa V, Cesa GL, Molinari E. Brief strategic therapy vs cognitive behavioral therapy for the inpatient/telephone-based outpatient treatment of binge eating disorder. The STRATOB randomized controlled clinical trial. *Clinical Practice and Epidemiology in Mental Health.* 2011;7.
 30. Cottraux J, Note ID, Boutitie F, Milliere M, Genouihlac V, Yao SN, et al. Cognitive therapy versus rogerian supportive therapy in borderline personality disorder: Two-year follow-up of a controlled pilot study. *Psychother Psychosom.* 2009; 78(5):307–16. [PubMed: 19628959]
 31. Dunn NJ, Rehm LP, Schillaci J, Soucek J, Mehta P, Ashton CM, et al. A randomized trial of self-management and psychoeducational group therapies for comorbid chronic posttraumatic stress disorder and depressive disorder. *J Trauma Stress.* 2007; 20(3):221–37. [PubMed: 17598141]
 32. Fairburn CG, Kirk J, O'Connor M, Cooper PJ. A comparison of two psychological treatments for bulimia nervosa. *Behav Res Ther.* 1986; 24(6):629–43. [PubMed: 3800834]
 33. Fairburn CG, Jones R, Peveler RC, Carr SJ, et al. Three psychological treatments for bulimia nervosa: A comparative trial. *Arch Gen Psychiatry.* 1991; 48(5):463–9. [PubMed: 2021299]
 34. Frank E, Kupfer DJ, Thase ME, Mallinger AG, Swartz HA, Fagioli AM, et al. Two-year outcomes for interpersonal and social rhythm therapy in individuals with bipolar I disorder. *Arch Gen Psychiatry.* 2005; 62(9):996–1004. [PubMed: 16143731]
 35. Goldman RN, Greenberg LS, Angus L. The effects of adding emotion-focused interventions to the client-centered relationship conditions in the treatment of depression. *Psychother Res.* 2006; 16(5): 537–49.
 36. Gosselin P, Ladouceur R, Morin CM, Dugas MJ, Baillargeon L. Benzodiazepine discontinuation among adults with GAD: A randomized trial of cognitive-behavioral therapy. *J Consult Clin Psychol.* 2006; 74(5):908–19. [PubMed: 17032095]
 37. le Grange D, Crosby RD, Rathouz PJ, Leventhal BL. A randomized controlled comparison of family-based treatment and supportive psychotherapy for adolescent bulimia nervosa. *Arch Gen Psychiatry.* 2007; 64(9):1049–56. [PubMed: 17768270]
 38. Greenberg L, Watson J. Experiential therapy of depression: Differential effects of client-centered relationship conditions and process experiential interventions. *Psychother Res.* 1998; 8(2):210–24.
 39. Guthrie E, Creed F, Dawson D, Tomenson B. Randomised controlled trial of psychotherapy in patients with refractory irritable bowel syndrome. *Br J Psychiatry.* 1993; 163:315–21. [PubMed: 8401959]
 40. Herbert JD, Gaudiano BA, Rheingold AA, Moitra E, Myers VH, Dalrymple KL, et al. Cognitive behavior therapy for generalized social anxiety disorder in adolescents: A randomized controlled trial. *J Anxiety Disord.* 2009; 23(2):167–77. [PubMed: 18653310]
 41. Hopko DR, Armento MEA, Robertson SMC, Ryba MM, Carvalho JP, Colman LK, et al. Brief behavioral activation and problem-solving therapy for depressed breast cancer patients: Randomized trial. *J Consult Clin Psychol.* 2011; 79(6):834–49. [PubMed: 21988544]
 42. Hudson JL, Rapee RM, Deveney C, Schniering CA, Lyneham HJ, Bovopoulos N. Cognitive-behavioral treatment versus an active control for children/adolescents with anxiety disorders. A randomized trial. *J Am Acad Child Adol Psychiatry.* 2009; 48(5):533–44. 10.1097/CHI.0b013e31819c2401
 43. Lipsitz JD, Gur M, Vermes D, Petkova E, Cheng J, Miller N, et al. A randomized trial of interpersonal therapy versus supportive therapy for social anxiety disorder. *Depress Anxiety.* 2008; 25(6):542–53. [PubMed: 17941096]
 44. Marchand A, Coutu MF, Dupuis G, Fleet R, Borgeat F, Todorov C, et al. Treatment of Panic Disorder with Agoraphobia: Randomized Placebo-Controlled Trial of Four Psychosocial Treatments Combined with Imipramine or Placebo. *Cog Behav Ther.* 2008; 37(3):146–59.

45. Marks I, Lovell K, Noshirvani H, Livanou M, Thrasher S. Treatment of posttraumatic stress disorder by exposure and/or cognitive restructuring: A controlled study. *Arch Gen Psychiatry*. 1998; 55(4):317–25. [PubMed: 9554427]
46. Masheb RM, Kerns RD, Lozano C, Minkin MJ, Richman S. A randomized clinical trial for women with vulvodynia: Cognitive-behavioral therapy vs. supportive psychotherapy. *Pain*. 2009; 141(1–2):31–40. [PubMed: 19022580]
47. Miklowitz DJ, Otto MW, Frank E, Reilly-Harrington NA, Wisniewski SR, Kogan JN, et al. Psychosocial treatments for bipolar depression: A 1-year randomized trial from the Systematic Treatment Enhancement Program. *Arch Gen Psychiatry*. 2007; 64(4):419–27. [PubMed: 17404119]
48. Shapiro DA, Firth J. Prescriptive v. exploratory psychotherapy: Outcomes of the Sheffield psychotherapy project. *Br J Psychiatry*. 1987; 151:790–9. [PubMed: 3502806]
49. Snyder DK, Wills RM. Behavioral versus insight-oriented marital therapy: Effects on individual and interspousal functioning. *J Consult Clin Psychol*. 1989; 57(1):39–46. [PubMed: 2925973]
50. Strauman TJ, Vieth AZ, Merrill KA, Kolden GG, Woods TE, Klein MH, et al. Self-system therapy as an intervention for self-regulatory dysfunction in depression: A randomized comparison with cognitive therapy. *J Consult Clin Psychol*. 2006; 74(2):367–76. [PubMed: 16649881]
51. Tarrier N, Pilgrim H, Sommerfield C, Faragher B, Reynolds M, Graham E, et al. A randomized trial of cognitive therapy and imaginal exposure in the treatment of chronic posttraumatic stress disorder. *J Consult Clin Psychol*. 1999; 67(1):13–8. [PubMed: 10028204]
52. Walsh BT, Wilson GT, Loeb KL, Devlin MJ, Pike KM, Roose SP, et al. Medication and psychotherapy in the treatment of bulimia nervosa. *Am J Psychiatry*. 1997; 154(4):523–31. [PubMed: 9090340]
53. Wilfley DE, Welch RR, Stein RI, Spurrell EB, Cohen LR, Saelens BE, et al. A randomized comparison of group cognitive-behavioral therapy and group interpersonal psychotherapy for the treatment of overweight individuals with binge-eating disorder. *Arch Gen Psychiatry*. 2002; 59(8): 713–21. [PubMed: 12150647]
54. Zettle RD, Rains JC. Group cognitive and contextual therapies in treatment of depression. *J Clin Psychol*. 1989; 45(3):436–45. [PubMed: 2745733]
55. Gaffan E, Tsaousis J, Kemp-Wheeler S. Researcher allegiance and meta-analysis: The case of cognitive therapy for depression. *J Consult Clin Psychol*. 1995; 63:966–980. [PubMed: 8543719]
56. Becker BJ. Synthesizing standardized mean-change measures. *Br J Math Stat Psychol*. 1988; 41:257–78.
57. Cohen J. A power primer. *Psychol Bull*. 1992; 112:155–159. [PubMed: 19565683]
58. Raudenbusch, SW.; Bryk, AS. Applications and data analysis methods. 2. Thousand Oaks, CA: Sage Publications; 2002. Hierarchical Linear Models.
59. Hedges LV, Tipton E, Johnson MC. Robust variance estimation in meta-regression with dependent effect size estimates. *Res Synth Methods*. 2010; 1:39–6.
60. Hedberg EC. ROBUMETA: Stata module to perform robust variance estimation in meta-regression with dependent effect size estimates. S457219 ed: Boston College Department of Economics; 2011.
61. O’Keefe DJ. Colloquy: Should Familywise Alpha Be Adjusted? *Human Communication Research*. 2003; 29:431–47.
62. Thoma NC, McKay D, Gerber AJ, Milrod BL, Milrod BL, Kocsis JH. A quality-based review of randomized controlled trials of cognitive-behavioral therapy for depression: An assessment and metaregression. *The American Journal of Psychiatry*. 2012; 169(1):22–30. [PubMed: 22193528]
63. Webb CA, DeRubeis RJ, Barber JP. Therapist adherence/competence and treatment outcome: A meta-analytic review. *J Consult Clin Psychol*. 2010; 78:200–211. [PubMed: 20350031]
64. Markowitz JC, Kocsis JH, Christos P, Bleiberg K, Carlin A. Pilot study of interpersonal psychotherapy versus supportive psychotherapy for dysthymic patients with secondary alcohol abuse or dependence. *J Nerv Ment Dis*. 2008; 196:468–74. [PubMed: 18552624]

Table 1

Studies using the “crossed therapist” design.

Author (reference)	Treatments	Condition	N	Researcher allegiance ^a	FACT rating
Agras, Walsh (23)	CBT vs IPT	Bulimia nervosa	220	CBT:1 IPT:0	0
Barkham, Rees (18)	Cognitive-Behavioral vs Psychodynamic-Interpersonal	Depression	36	CB:0 PI:0	2
Barkham, Shapiro (24)	Cognitive-Behavioral vs Psychodynamic Interpersonal	Sub-syndromal depression	116	CB:0 PI:0	2
Beck, Sokol (25)	CBT vs Client Centered supportive	Panic disorder	33	CBT:3 CC:0	1
Bjornsson, Bidwell (26)	Cognitive-Behavioral Group Therapy vs Group Psychotherapy	Social Anxiety Disorder	45	CBGT:2 GPT:0	0
Borkovec, Mathews (27)	CBT vs nondirective	Generalized anxiety disorder	30	CBT:3 ND:0	0
Borkovec and Mathews (28)	Cognitive vs Coping Desensitization vs Nondirective	Generalized anxiety disorder and panic disorder without agoraphobia	30	CT:3 CD:2 ND:0	0
Castelnovo, Manzoni (29)	Brief Strategic Therapy vs Cognitive Behavioral Therapy	Binge Eating Disorder	60	BST:1 CBT:0	0
Clark, Ehlers (20)	Cognitive Therapy vs Exposure + Applied Relaxation	Social Phobia	62	CT:2 EXP+AR:0	0
Cottraux, Note (11)	CBT vs Rogerian Supportive Therapy	Chronic PTSD	60	CBT:2 RST:0	1
Cottraux, Note (30)	CBT vs Rogerian Supportive Therapy	Borderline Personality Disorder	65	CBT:2 RST:0	1
Dunn, Rehm (31)	Self-management therapy vs psychoeducational group therapy	Chronic PTSD + Depression	101	SMT:3 PEGT:0	0
Fairburn, Kirk (32)	CBT vs short term focal psychotherapy	Bulimia Nervosa	24	CBT:2 STFP:0	0
Fairburn, Jones (33)	CBT vs IPT vs Behavior Therapy	Bulimia Nervosa	75	CBT:2 IPT:0 BT:0	1
Frank, Kupfer (34)	Interpersonal and Social Rhythm Therapy vs Intensive Clinical Management	Bipolar I Disorder	175	IPSRT:3 ICM:0	0
Goldman, Greenberg (35)	Process-experiential psychotherapy vs client-centered psychotherapy	Depression	38	PE:1 CC:0	1
Gosselin, Ladouceur (36)	CBT vs Nonspecific Supportive Therapy	Generalized Anxiety Disorder + Benzodiazepine dependency	61	CBT:3 NST:0	0
le Grange, Crosby (37)	Family Based Therapy vs Supportive Therapy	Anorexia Nervosa	80	FBT:3 ST:0	0
Greenberg and Watson (38)	Process-experiential psychotherapy vs client-centered psychotherapy	Depression	34	PE:1 CC:0	1
Guthrie, Creed (39)	Psychodynamic Therapy vs Supportive Listening	Irritable Bowel Syndrome	101	PDT:3 SL:0	1
Herbert, Gaudiano (40)	CBT vs Educational/Supportive Therapy	Social Anxiety Disorder	73	CBT:2 E/ST:0	0
Hopko, Armento (41)	Behavioral Activation vs Problem Solving Therapy	Depressed Breast Cancer Patients	80	BA:1 PST:0	0
Hudson, Rapee (42)	CBT vs Group Support and Attention	Children and adolescents with anxiety disorders	112	CBT:3 GSA:0	0

Author (reference)	Treatments	Condition	N	Researcher allegiance ^a	FACT rating
Lipsitz, Gur (43)	IPT vs Supportive therapy	Social anxiety disorder	70	IPT:2 ST:0	0
Luty, Carter (19)	CBT vs IPT	Major Depression	177	CBT:0 IPT:0	2
Marchand, Coutu (44)	Cognitive Therapy vs Exposure, CT+Exposure vs Supportive Therapy	Panic Disorder with Agoraphobia	137	CT:2 Exp:2 CT+Exp:3 ST:0	0
Marks, Lovell (45)	Exposure vs Cognitive restructuring vs Exposure combined with Cognitive restructuring vs Relaxation	PTSD	87	E:3 Cr:2 E+Cr:3 R:0	0
Masheb, Kerns (46)	CBT vs Supportive Therapy	Vulvodynia	50	CBT:3 ST:0	0
McIntosh, Jordan (21)	CBT vs IPT vs Nonspecific Supportive Clinical Management	Anorexia Nervosa	56	CBT:1 IPT:1 SCM:0	0
Miklowitz, Otto (47)	Family-focused therapy vs Interpersonal and Social Rhythm therapy vs CBT vs Collaborative Care	Bipolar Depression	293	FFT:3 SRT:3 CBT:3 CC:0	0
Shapiro and Firth (48)	Prescriptive vs Exploratory therapy	Mixed depression and anxiety (mostly depression)	40	Pr: 0 Exp:0	2
Shapiro, Barkham (13)	Cognitive-Behavioral vs Psychodynamic-Interpersonal	Depression	117	CB:0 PI:0	2
Snyder and Wills (49)	Behavioral vs Insight oriented Marital Therapy	Marital distress	79	BMT:0 IOMT:0	3
Strauman, Vieth (50)	Self-System Therapy vs Cognitive Therapy	Depression	45	SST:1 CT:0	1
Tarrier, Pilgrim (51)	Imaginal Exposure vs Cognitive Therapy	Chronic PTSD	72	IE:0 CT:0	0
Taylor, Thordarson (22)	CBT (prolonged exposure) vs EMDR vs Relaxation training	PTSD	60	CBT: 2 EMDR:1 R:0	0
Walsh, Wilson (52)	CBT vs Supportive Therapy	Bulimia Nervosa	120	CBT:3 ST:0	0
Wilfley, Welch (53)	Group CBT vs Group IPT	Binge Eating Disorder	162	CBT:0 IPT:0	0
Zetle and Rains (54)	Group Comprehensive Distancing vs group Cognitive Therapy vs group Partial Cognitive Therapy	Depression	31	CD:3 CT:2 PCT:0	0

Note:

^aRelative researcher allegiance score for each treatment condition included in study