
Article:
Power, N., Noble, L., Simmonds-Buckley, M. et al. (4 more authors) (2022) Associations between treatment adherence-competence-integrity (ACI) and adult psychotherapy outcomes : a systematic review and meta-analysis. Journal of Consulting and Clinical Psychology, 90 (5). pp. 427-445. ISSN 0022-006X

https://doi.org/10.1037/ccp0000736

Associations between treatment adherence-competence-integrity (ACI) and adult psychotherapy outcomes: a systematic review and meta-analysis

Niall Power[1,2*]

Louise Noble[1]

Melanie Simmonds-Buckley[1]

Stephen Kellett[1,3]

Corrie Stockton[1]

Nick Firth[4]

&

Jaime Delgadillo[1]

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

[2] Derbyshire Community Health Services NHS Foundation Trust, UK

[3] Rotherham, Doncaster and South Humber NHS Foundation Trust, UK

[4] School of Health and Related Research, University of Sheffield, UK

* Correspondence to Niall Power, Ash Green Specialist Learning Disability Service, Ashgate Road, Ashgate, Chesterfield, S42 7JE, UK; email: niall.power4@nhs.net, Tel: 01246 565000.

Abstract

Objective: To provide a comprehensive assessment of the association between psychological treatment adherence/competence/integrity (ACI) and clinical outcomes. ***Method:*** The review protocol was pre-registered (CRD42020193889). Studies that assessed ACI-outcome relationships for adult psychotherapy were searched across three databases (Scopus, PsycINFO, MEDLINE). Random effects meta-analyses were conducted on adherence-outcome, competence-outcome and integrity-outcome relationships. Separate analyses were performed for studies with hierarchical (i.e. patients nested within therapist) versus non-hierarchical study designs. Moderator analyses were performed according to pre-defined clinical and methodological features. GRADE assessments rated the quality of each meta-analytic comparison. ***Results:*** The review identified 62 studies suitable for inclusion (45 adherence-outcome, 39 competence-outcome and 7 integrity-outcome effect sizes; N=8,210 across all analyses). No significant adherence-outcome association was found. A small significant positive association was found only in non-hierarchical studies between competence and outcome (r = 0.17, 95% CI [0.07-0.26], $p < 0.001$, ~d = .34, GRADE = moderate). Small-to-moderate significant positive associations between integrity and outcome were found for both non-hierarchical (r = 0.15, 95% CI [0.06 – 0.23], $p < 0.001$, ~d = .30, GRADE = high) and hierarchical study designs (r = 0.23, 95% CI [0.01,0.43], $p < 0.044$, ~d = .47, GRADE = low). Diagnosis, treatment modality and year of publication significantly moderated the strength of ACI-outcome correlations. ***Conclusions:*** Competence and integrity are significantly associated with clinical outcome, with a magnitude comparable to wider common factors. Further research is required to study these process-outcome associations with greater precision in routine-care settings and to understand the role of moderating variables.

Public Health Significance Statement

This contemporary and comprehensive review of the body of evidence on the relationship between adherence/competence/integrity and clinical outcomes in adult psychotherapy indicates that competence and integrity are significantly related to clinical outcomes. Treatment integrity is where there is evidence of adherence to treatment procedures (with a clear differentiation from other approaches) in the context of skilful delivery of clinical methods. This meta-analysis offers empirical support for the routine monitoring of psychological interventions to audit that treatments are delivered in a theoretically consistent manner (therapist adherence), but with due attention to therapist interpersonal and relational skills (therapist competence). The integrity construct requires further conceptual clarification and more well conducted studies in naturalistic settings.

Keywords

Adherence; Competence; Integrity; Fidelity; Outcome; Psychotherapy.

A vast body of evidence attests to the effectiveness of empirically supported therapies (EST) for a wide range of psychological problems (Chambless & Hollon, 1998; Barkham & Lambert, 2021). However, there are still gaps in understanding how psychotherapy leads to therapeutic change (Kazdin, 2007). Psychotherapy process research focuses on phenomena occurring between therapists and clients (Orlinsky & Howard, 1986). These phenomena include interpersonal skills that are common to many forms of psychotherapy and also the implementation of treatment-specific, theoretically-informed strategies and/or techniques. The examination of therapist interpersonal skills (e.g., alliance-building, collaboration, empathy, credibility, etc.) has a long tradition in the field of process research and an established evidence base supporting their relevance to treatment outcomes (Heinonen & Nissen-Lie, 2020; Norcross & Lambert, 2019; Wampold, 2015). However, the relative importance of applying specific strategies – which is the central focus of the present study – has been a matter of controversy and debate in the field. Over the last decade, this dispute has been characterised by polarised views that either prescribe (e.g., Roth & Pilling, 2008) or contest (e.g., Wampold & Imel, 2015) the requirement to implement specific therapeutic models or strategies.

Adherence, Competence and Integrity (ACI)

The relevance of treatment-specific strategies has been examined in a number of ways; often with resort to the related concepts of treatment adherence, competence and integrity (ACI). *Treatment adherence* refers to the extent to which the methods or techniques prescribed in a therapy model/protocol then actually occur in treatment (Waltz et al., 1993). Hence, the concept of adherence measures a therapist's relative compliance with a pre-specified set of methods, which are theoretically-informed and grounded in a particular therapeutic model. In short, whether treatment is delivered as intended (Southam-Gerow et al., 2016). *Treatment competence* refers to the knowledge, skill and appropriateness (Waltz et al., 1993) with which these model-specific methods are delivered to the standards required to achieve expected outcomes (McLeod et al., 2018). As such, rating adherence to

some degree is a necessary pre-condition to rating competence. *Treatment integrity* – also referred to as *fidelity* (hereafter referred to as 'integrity'; Miller & Binder, 2002) – is a more recently developed and less well-defined construct compared to adherence and competence. Assessment of treatment integrity should encompass three related aspects: [1] therapist adherence, [2] therapist competence, and [3] treatment differentiation (Waltz et al., 1993; Perepletchikova et al., 2007). Measures of therapist adherence are fit for purpose for evaluating treatment differentiation when they detail both the proscribed (i.e., methods to avoid which may reduce intervention purity) and prescribed methods (Perepletchikova et al., 2007). This latter point, in particular, has led to a number of conceptual issues when attempting to assess, evaluate and report ACI in psychotherapy outcome studies (see Leichsenring et al., 2011 for a more comprehensive discussion). Some measures of adherence only contain prescribed methods (e.g., Cognitive-Behavioral Therapy Adherence Scale; Weck et al., 2014) and others set out both prescribed and proscribed methods (e.g., Collaborative Study Psychotherapy Rating Scale; Hill et al., 1992). Conceptual distinctions between adherence, competence and integrity are important although are arguably not as clear and consistently studied as other process constructs, such as the therapeutic alliance (e.g., Flückiger et al., 2018).

The gold standard method to measure ACI involves the regular and systematic rating of recorded therapy sessions, using psychometrically validated and structured rating scales with detailed associated scoring manuals, ratings being performed by more than one reviewer, and across a random sample of therapy sessions of the phases of a therapy (Hogue et al., 1996). Whilst ACI constructs share conceptual similarities, they are sufficiently distinct to be considered separately in process-outcome research and meta-analyses (e.g., Collyer et al., 2019).

Are ACI Domains Important in Psychotherapy Practice?

Monitoring ACI has been proposed to be relevant in clinical trials, training programmes and routine practice. In clinical trials, assessing adherence to the experimental intervention is necessary to draw valid conclusions concerning its efficacy (Perepletchikova et al., 2007). Moreover, assessing

competence is necessary to determine if clinical outcomes may vary as a function of competent delivery (Kazdin, 1986). Consistent with this logic, successful qualification in some clinical training courses requires trainees to demonstrate adherence and competence guided by pre-specified treatment protocols and rating scales (e.g., Branson, Shafran, & Myles, 2015). There have also been efforts towards developing empirically derived "competency frameworks" to support the implementation of ESTs into routine care (Roth & Pilling, 2008; Lemma et al., 2008; Parry et al., 2021). Consequently, some services have established service policies and guidelines requiring therapists to demonstrate ACI as part of their clinical practice and supervision (e.g., Gearing et al., 2011; National Collaborating Centre for Mental Health, 2020). Perepletchikova and Kazdin (2005) argued that the successful implementation of ESTs in routine services was dependant on therapists delivering ESTs with high adherence and competence. The maintenance of adherence and competence is therefore considered a key aspect of maintaining quality psychological care in routine settings (McLeod et al., 2019). The emphasis on ACI across these contexts is therefore based on the assumption that outcomes are dependent on their presence. But is this assumption supported by empirical evidence?

Meta-analyses of the relationship between ACI and treatment outcomes

In an early meta-analytic review of adherence/competence in adult psychotherapy studies, Webb et al., (2010) analysed data from 32 samples (within 36 eligible studies). They reported correlations between adherence–outcome (r = .02) and competence–outcome (r = .07) which were not statistically significant. Sub-group analyses indicated that studies controlling for therapeutic alliance had a significantly smaller competence-outcome association, although this was not the case for adherence. Zarafonitis- Müller et al., (2014) conducted a meta-analysis of the relationship between adherence/competence and outcome in seven studies of cognitive behavioural therapy (CBT). They found a significant competence-outcome correlation (r = 0.24) across multiple outcome domains, and a larger correlation for depression outcomes (r = 0.38); however non-significant findings were reported for adherence-outcome. In a more recent meta-analysis including 29 samples from child psychotherapy

studies, Collyer et al., (2019) reported a significant adherence-outcome correlation (r = 0.096), however non-significant findings were reported for the small subset of studies investigating competence (n=9) and fidelity/integrity (n=5).

A pattern observed from these meta-analyses, is that the earliest review indicated negligible and non-significant associations between ACI and treatment outcomes, whereas the more recent reviews found significant, albeit small, effect sizes favouring some ACI constructs, but not others. One possible explanation for this is that the earlier original studies from the 1980s and 1990s did not specifically measure adherence or competence using validated measures, but instead used proxy indicators. For example, counts of specific treatment actions (e.g., transference interpretations; Marziali et al., 1984) or generic checklists of various possible techniques that therapists could choose to use, rather than techniques that were expected to be adhered to or delivered skilfully. Another related explanation is that more recent studies have adopted newer more reliable measures of ACI constructs, and have also improved their methods to rate these constructs reliably (e.g., random sampling, double-rating). Furthermore, newer studies also report data from protocols that have been designed to treat specific conditions by targeting key maintenance aspects (e.g., CBT for social anxiety disorder; Ginzburg et al., 2012). Hence, results from earlier studies could be diluted by conceptual and measurement limitations, or it could be that the newer studies capture data from those treatment protocols appropriately refined to treat specific clinical conditions.

A further source of potential bias concerns the methodological design of ACI-outcome studies. Many of these studies, particularly earlier ones, ignore the fact that psychotherapy data has a hierarchical structure, where patients are nested or clustered ‘underneath’ therapists. Evidence strongly suggests that therapists can systematically vary in terms of their effectiveness, particularly in routine practice (Johns et al., 2019). This phenomenon, known as “therapist effects”, is thought to be related to therapist interpersonal skills and attitudes, rather than differential delivery of treatment-specific competences or technical skills (Heinonen & Nissen-Lie, 2020; Wampold & Owen, 2021). Therefore,

in order to draw inferences about ACI-outcome associations, the hierarchical structure of data should be taken into account and therapist effects should be controlled for using random effects in a multilevel modelling framework (Baldwin & Imel, 2013). To date, meta-analyses of ACI-outcome associations have not accounted for hierarchical analyses (i.e., multilevel modelling).

The Current Study

Overall, the pattern of results described above warrants a comprehensive update of the ACI-outcome literature in adult psychotherapy, particularly considering that the last meta-analytic review was published over a decade ago. Based on this rationale, we conducted a comprehensive systematic review and meta-analysis of adult psychotherapy ACI-outcome studies. The study was guided by the following objectives: (1) quantify the strength of the relationship between ACI constructs and treatment outcomes, (2) examine these relationships across hierarchical and non-hierarchical study designs, (3) explore theoretically-informed effect size moderators and (4) compare results with the last meta-analytic review conducted.

Method

Registration and Selection of Studies

The review protocol was pre-registered with PROSPERO (ID CRD42020193889). Online databases Scopus, PsycINFO and MEDLINE were searched in line with pre-defined inclusion/exclusion criteria. No publication date limits were used. The specific search terms and Boolean logic used are included in the online supplementary materials. Searches were conducted between 18th and 22nd August 2020. Forward and reverse citation searches were carried out on all included studies and reference lists of previous reviews were hand-searched. Unpublished literature was searched by contacting primary authors of all included studies and searching Grey Matters, Ethos and ProQuest: Dissertations and Theses databases. Authors (N=52) were contacted via email and given a period of 2-weeks to reply. Although some were identified through hand searches, no unpublished papers were received in response to the email request. Study inclusion criteria were as follows: (1)

adult (18+ years old) clinical population, (2) delivery of face-to-face individual or group psychotherapy, (3) utilised trained researcher/expert-rated measures of adherence/competence/integrity drawn from videotaped, audiotaped, or transcribed therapy sessions, (4) quantitative validated measures of treatment outcome, (5) measured treatment outcome after psychotherapy had completed, (6) conducted statistical analyses on the relationship between adherence/competence/integrity and treatment outcome and finally (7) were written in English. Study exclusion criteria were: (1) case studies (all other methodological designs were considered for inclusion), (2) therapist or patient-rated competence, (3) internet or telephone-based treatment and (4) competence ratings based on proxy evaluations (e.g., role plays).

In terms of the third inclusion criterion above, it is important to note that contemporary ACI operational definitions and associated measures have developed since early process-outcome studies, propelled by the move towards generating ESTs that use well defined treatment protocols. Studies that used an expert rater using a validated measure to explicitly measure the extent of method/technique use and quantitatively associated this with clinical outcome were included as adherence-outcome studies. Similarly, studies that used an expert rater using a validated measure of the skilfulness of the therapist in delivering specific methods/techniques were included as competence-outcome studies (e.g., measuring the clinical skills of psychotherapists; Chevron et al., 1983). In almost all other cases, studies were screened and included in the review and analyses based on whether authors described the construct under investigation as adherence, competence, or integrity.

A sub-sample of titles and abstracts (20%) and full texts (20%) were independently screened by a blind secondary reviewer (MSB). From the total number of titles/abstracts to be screened from bibliographic databases (7,098), the secondary reviewer randomly selected 20% (1,420) to screen based on the eligibility criteria. Screening decisions were then compared to the primary reviewer's (NP) results. Similarly, from the total number of full texts from bibliographic databases to be screened (248), the secondary reviewer picked 20% (50) at random to screen and these decisions were compared to the

primary reviewers' decisions. Inter-rater reliability indicated 'substantial agreement' (kappa = 0.77) in title/abstract screening and 'perfect agreement' (kappa = 1.0) in full text screening (Landis & Koch, 1977).

Data Extraction

Data were extracted by NP using a bespoke tool that was piloted and found to be fit-for-purpose. Data extracted included the study's primary author, year of publication, relevant effect size statistics and whether competence, adherence, integrity or fidelity was measured. Additional extracted information included the type of therapy investigated, disorder treated, patient and therapist sample size and demographic information (percentage female and mean age) and the names of the outcome and ACI measures used. The percentage of sessions rated was extracted, whether the whole or a subscale of the ACI measure was used and whether a generic or therapy-specific ACI measure was used. A sub-sample (10%) of included studies were extracted by an independent, blind reviewer (CS) and informally compared to the primary reviewers' (NP) extractions. This verified the reliability of the extraction tool.

Quality Assessments

An adapted version of the Newcastle-Ottawa scale (NOS): cohort studies (Wells et al., 2000) assessed risk of bias (RoB). The original version of the NOS: cohort studies scale contains eight items across three sections. In section one ('Selection'), there are four items: [1] Representativeness of the exposed cohort. [2] Selection of the non-exposed cohort, [3] Ascertainment of exposure, and [4] Demonstration that outcome of interest was not present at start of study. In section two ('Comparability'), there is one item: [5] Comparability of cohorts on the basis of the design or analysis. In Section three ('Outcome), there are three items: [6] Assessment of outcome, [7] Was follow-up long enough for outcomes to occur, and [8] Adequacy of follow-up cohorts. In the original version of the NOS, one star is awarded to the study being rated if an aspect of the study is deemed less biased (two

stars could be awarded for item 5: ‘Comparability’), which results in a maximum RoB score of nine stars.

In the adapted version used in the current review, item 2 (‘selection of the non-exposed cohort’) and item 5 (comparability of cohorts on the basis of the design or analysis) were removed as they were concerned with control conditions, which were not analysed by any of the studies. Item 8 (‘adequacy of follow-up cohorts’) was modified so that the follow-up adequacy referred to only the sample for which ACI-ratings were reported.

The RoB scoring system was also adapted for the current review. Studies were assessed on a scale of 0-6 stars (i.e. higher scores suggested lower RoB) and categorised as low RoB (score of 3 for ‘Selection’ and 2-3 for Outcome), moderate RoB (2 for Selection and 2-3 for Outcome) or high RoB (0-1 for Selection and 0-1 for Outcome). If a study met the criteria for one rating in one section (e.g., ‘Moderate’ for ‘Selection’) but met the criteria for a lower rating in another section (e.g., ‘High’ for ‘Outcome’) they would be awarded the lower of the two ratings.

A randomly selected subset of included studies was blind rated in two pairs (pair 1; CS & NP, N=18, 29.03% and pair 2; MSB & SK, N=4, 6.45%). The Kappa agreement was ‘substantial’ in pair 1 (0.68) and 'perfect' in pair 2 (1.0; Landis & Koch, 1977). The quality of evidence for each meta-analytic comparison was assessed with the GRADE tool (Atkins et al., 2005). Three reviewers (NP, SK and MSB) used five criteria; limitations of included studies, level of imprecision in estimates, amount of unexplained heterogeneity, indirectness of the evidence and publication bias. Downgrading or upgrading the quality of the evidence was achieved via consensus (rated as very low, low, moderate or high).

Effect Size Extraction and Calculation

For studies that reported a bivariate correlation coefficient to represent the association between ACI and treatment outcome, this effect size was directly extracted. In cases where a study reported an alternative effect size statistic, a correlation coefficient (*r*) was calculated using the Lenhard and

Lenhard (2016) effect size conversion calculator. In cases where a correlation coefficient could not be calculated, a partial correlation was calculated using the relevant standardised regression coefficient (Beta), standard error and number of observations from a regression model. If the standard error was not reported, 95% confidence intervals of Beta were used to calculate the standard error. If the statistics reported in the article were not sufficient for a correlation, nor a partial correlation coefficient to be calculated, the author was contacted to request sufficient data (16 authors were contacted and 2 responded with the requested data). If authors did not respond within two weeks, the study was excluded. Where necessary, the direction of the effect size was switched so that a positive correlation coefficient represented a positive association between ACI and outcome. Where multiple effect sizes were reported, a preference hierarchy was used so that each sample only contributed one effect size. If several effect sizes representing ACI-outcome associations were reported due to several different outcome measures being used, an outcome measure hierarchy was applied. The outcome hierarchy was as follows: (1) independent/clinician-rated outcome measures were given preference over self-report outcome measures; (2) problem-specific outcome measures relevant to the target problem were given preference over generic outcome measures; (3) validated self-report outcome measures were given preference over study specific outcomes (e.g. number of self-harm incidents over the past week). When several effect sizes were reported due to ACI being measured at numerous time-points, the time-point closest to the cessation of therapy was given preference. If several effect sizes were reported due to several sub-scales of the ACI measure being reported, the total ACI association was extracted. If no total was reported, the sub-scale most relevant to the type of therapy delivered was extracted. Finally, if no sub-scale had greater relevance to the type of therapy than any other, the mean correlation coefficient of the numerous sub-scale effect sizes was calculated to represent the overall association between ACI and outcome.

Meta-Analytic Strategy and Publication Bias

Effect sizes extracted from original studies were grouped according to whether the ACI-outcome analysis was based on a 'hierarchical' analysis or 'non-hierarchical' analysis. Effect sizes were grouped according to whether they were a competence-outcome, adherence-outcome or integrity-outcome association. Random-effects models were used due to the methodological differences between the studies (Borenstein et al., 2011) and meta-analyses were only carried out on groupings of at least two studies (Valentine et al., 2010). All analyses were carried out using Meta-Essentials Workbook 5 (correlational data), version 1.5 (Suurmond et al., 2017). Plots were produced in R (version 3.6.1) using the packages *forestplot* and *metafor* (Gordon & Lumley, 2021; Viechtbauer, 2010). Study heterogeneity was estimated using the 'Q' and 'I^2' statistics. As 'Q' is susceptible to the number of studies in an analysis (Higgins et al., 2003), I^2 was also calculated to indicate the extent of variability across studies due to heterogeneity rather than chance (Higgins et al., 2003). Percentage I^2 was 0% indicated no heterogeneity and 25%, 50% and 75% indicating low, medium and high heterogeneity respectively (Higgins et al., 2003). Publication bias was explored via visual symmetry analysis of funnel plots (Light & Pillemer, 1984) and examined statistically using Egger's regression test (Egger et al., 1997).

Moderator Analyses

Moderator analyses were carried out to explore heterogeneity across studies (Borenstein et al., 2011). Categorical and continuous study characteristics were analysed. For categorical moderator analyses, a subgroup analysis was conducted when at least 10 studies were eligible for inclusion and there were at least two studies in each subgroup (Deeks et al., 2019). In moderator analyses where the moderator was continuous, a meta-regression was conducted when at least six studies were eligible for inclusion in the specific meta-regression. The subgroup analyses specified a priori were (1) the disorder treated, (2) the treatment modality, (3) treatment format (i.e., individual or group), (4) timing of the ACI measure (i.e., early, middle, or late) and (5) alliance-confound (i.e., whether or not the study statistically controlled for therapeutic alliance). In the timing of the ACI subgroup analysis, if

various/random sessions were rated, then the study was excluded from the analysis. The meta-regression analyses tested whether effect sizes in each meta-analysis varied as a function of (1) the year of publication, (2) percentage coverage of ACI ratings and (3) risk of bias assessment. To extract the relevant data to calculate the percentage coverage of ACI ratings, a preference hierarchy was used. In studies where a standardised number of therapy sessions were delivered, the percentage was simply the number of sessions from which the ACI-ratings were taken. If the therapy durations varied across patients, the mean number of sessions delivered in the study was used. If therapy durations varied and no mean therapy-duration was reported, the maximum number of therapy sessions provided was used to calculate the percentage coverage. If no data was reported regarding therapy duration, the study was excluded from the 'percentage coverage' meta-regression. The level of significance used in moderator analyses was adjusted using Bonferroni corrections for the subgroup and meta-regression analyses separately.

Results

Study Selection

The PRISMA summary (Moher et al., 2009) is displayed in Figure S5 (Supplementary Information). Electronic database searches identified 7,098 records after duplicates were removed. A further 32 records were located via hand searching and the unpublished literature search. A total of 7,130 records were therefore screened by title and abstract; 6,850 were removed (96.07%). The remaining 280 full texts were screened against eligibility criteria; 218 were removed. In total, 62 individual articles were eligible for inclusion containing 91 effect size statistics, consisting of 39 competence-outcome, 45 adherence-outcome, and 7 integrity-outcome effect sizes.

Summary Study Characteristics

The characteristics of the non-hierarchical and hierarchical studies are summarised in Tables 1 and 2 respectively. Individual study characteristics are reported in the supplementary materials. In

terms of study location, 45.05% (41 samples) were conducted in the USA, with remaining samples drawn from Germany (12, 13.19%), United Kingdom (12, 13.19%) and Canada (13, 14.29%). The average patient sample size was 90.22 (SD=157.36), treated by an average of 15.81 (SD=28.14) therapists. The most common treatment was CBT (27, 29.67%) and depression (26, 28.57%) was the most common disorder treated. Only a minority of studies controlled for therapeutic alliance (11, 12.09%). Mean coverage of the proportion of therapy sessions rated was 31.98% (SD=34.21).

Risk of Bias and GRADE Assessments

The risk of bias (RoB) summaries are also reported in Tables 1 and 2. For non-hierarchical studies, the majority were rated as low RoB (64.29%, 59.26% and 100.00% for adherence, competence and integrity samples, respectively), though a considerable proportion were rated as moderate (between 32.14%-40.74%) and a small proportion as high RoB (between 0%-3.57%). Studies were generally of lower RoB across all adherence-outcome samples (64.29%-85.71% rated low RoB), compared to all competence-outcome samples (57.14%-59.26% rated low RoB). All four non-hierarchical integrity-outcome studies were rated low RoB. For the hierarchical studies, four competence-outcome studies were rated as low RoB and three rated as at high RoB. Mean (*SD*) RoB scores for non-hierarchical studies were 4.82 (0.82), 4.67 (0.62) and 5.33 (0.58) for adherence, competence and integrity studies respectively. Hierarchical studies had ratings of 4.86 (0.38), 4.43 (1.13) and 4.67 (0.58) for adherence, competence and integrity studies respectively. A consistent methodological strength was the representativeness of samples and a consistent weakness lack of independent/blind outcome assessment. RoB assessments of individual studies can be viewed in the supplementary materials. The initial quality of the six meta-analytic comparisons assessed using the GRADE was set at 'high' quality as most studies involved rigorous assessments of ACI (e.g., were observer-rated using validated tools and involved inter-rater reliability checks). No comparisons were downgraded on limitations of included studies or indirectness of evidence criteria. Unexplained heterogeneity, imprecision in

estimates and publication bias generated a mixture of gradings (see each comparison for assigned GRADE rating).

Meta-Analyses of Non-Hierarchical Studies

The non-hierarchical adherence-outcome meta-analysis from n = 2,015 patients across 28 studies (k = 37 independent samples) is displayed in Figure 1. The adherence-outcome association was small (r = 0.06, 95% CI [-0.03,0.16], p = 0.091, ~d = .12, GRADE = moderate) indicating no significant association between adherence and outcome. A down-grading of quality was made for unexplained heterogeneity, as there was evidence of substantial heterogeneity (I^2 = 71.28%; Q(df = 36) = 146.23, p < 0.001). Visual inspection of the funnel plot (see Figure S3a in Supplementary Information) suggested some asymmetry and thus risk of publication bias, but Egger's regression was non-significant (p = 0.195). The non-hierarchical competence-outcome meta-analysis from 4,082 patients across 27 studies (k = 32 independent samples) is displayed in Figure 2. The overall association (r = 0.17, 95% CI [0.07,0.26], p < 0.001, ~d = .34, GRADE = moderate), indicated a small but significant association between competence and outcome. A down-grading of quality was given for unexplained heterogeneity, as there was evidence of substantial heterogeneity (I^2 = 78.23%; Q(df = 31) = 146.98, p < 0.001). Visual inspection of the funnel plot (see Figure S3b in Supplementary Information) suggested risk of publication bias, but Egger's regression was non-significant (p = 0.960). The meta-analysis summarising integrity-outcome associations from 538 patients across 3 studies (k = 4 independent samples) is displayed in Figure 3. The overall association was small, but significant (r = 0.15, 95% CI [0.06,0.23], p < 0.001, ~d = .30, GRADE = high). There were no down-grading regarding quality. There was little evidence of heterogeneity (I^2 = 0%; Q(df = 3) = 1.07, p = 0.783). Visual inspection of the funnel plot (see Figure S3c in Supplementary Information) suggested risk of publication bias, but Egger's regression was non-significant (p = 0.778).

Meta-Analyses of Hierarchical Studies

The hierarchical adherence-outcome meta-analysis from 653 patients across 7 studies (k = 8 independent samples) is displayed in Figure 4. The adherence-outcome association was small (r = 0.04, 95% CI [-0.18,0.25], p = 0.339, ~d = .08, GRADE = very low), indicating no significant association between adherence and outcome. Quality down-gradings were given for unexplained heterogeneity (I^2 = 76.04%; Q(df = 7) = 29.22, p < 0.001), imprecision in estimates and publication bias suggested via visual inspection of the funnel plot (see Figure S4a in Supplementary Information) and significant Egger's regression (p = 0.046). The meta-analysis summarising competence-outcome associations from 517 patients across 7 studies (k = 7 independent samples) is displayed in Figure 4. The overall association was small (r = 0.03, 95% CI [-0.15,0.21], p = 0.353, ~d = .06, GRADE = low), indicating no significant association between competence and outcome. Quality down-gradings were given for unexplained heterogeneity (I^2 = 62.44%; Q(df = 6) = 15.97, p < 0.014) and imprecision in estimates. Visual inspection of the funnel plot (see Figure S4b in Supplementary Information) suggested risk of publication bias, but Egger's regression was non-significant (p = 0.212). The meta-analysis summarising integrity-outcome associations from 405 patients across 3 studies (k = 3 independent samples) is displayed in Figure 4. The overall association was moderate, but significant (r = 0.23, 95% CI [0.01,0.43], p < 0.044, ~d = .47, GRADE = low). Quality down-gradings were given for unexplained heterogeneity (I^2 = 77.03%; Q(df = 2) = 8.71, p < 0.013) and imprecision in estimates. Visual inspection of the funnel plot (see Figure S4c in Supplementary Information) suggested no asymmetry and Egger's regression was also non-significant (p = 0.357).

Sensitivity Analyses

As there were a number of studies that were included in Webb et al. (2010) that were not included in the current review's main analyses, sensitivity analyses were therefore performed which included the adherence/competence-outcome correlation coefficients of the additional studies. Seven adherence-outcome and two competence-outcome effect sizes were added (both non-hierarchical analyses). Forest plots from these sensitivity analyses are included in Supplementary Information

(Figure S1 and S2), in addition to a separate screening table. This states the original exclusion reasons for any studies included in Webb et al. (2010) and not in the current review, in addition to whether each study was included/excluded in the sensitivity analysis (Table S9). The sensitivity analyses showed that the non-hierarchical adherence-outcome meta-analysis remained non-significant (r = 0.05, 95% CI [-0.03,0.13], p = 0.101, ~d = .10) and non-hierarchical competence-outcome meta-analysis remained significant (r = 0.17, 95% CI [0.08,0.26], p <.001, ~d = .35).

Moderator Analyses

A total of 18 moderator analyses (6 categorical and 12 continuous) were viable. Only moderators in the non-hierarchical adherence-outcome and competence-outcome effect sizes were possible. The three subgroup analyses (disorder treated, treatment modality and alliance confound) are reported in Table 3. No significant variation in effects was found for adherence-outcome studies. For competence-outcome studies, significant differences were found for disorder treated and treatment modality. The strongest positive association between competence and outcome was during treating of anxiety disorders; substance-use/addiction interventions had a negative competence-outcome association. In terms of treatment modality, cognitive therapy had the strongest positive competence-outcome association; emotion focused therapy had a negative association. The meta-regressions investigating continuous moderators for adherence-outcome and competence-outcome associations in the non-hierarchical and hierarchical studies are reported in Table S10 (Supplementary Information). In hierarchical adherence-outcome studies, more recent studies were significantly associated with a larger positive adherence-outcome association.

Discussion

This systematic review and meta-analysis synthesised data from published and non-published studies investigating associations between ACI constructs and clinical outcomes during adult psychotherapy. Statistically significant competence-outcome and integrity-outcome associations were

found in conventional studies using non-hierarchical designs. For studies using hierarchical designs that controlled for between-therapist variability in treatment outcomes, only the integrity-outcome association was statistically significant. As such, and in contrast to an earlier review (Webb et al., 2010), this review presents new evidence that supports the notion that the skilled implementation of model specific treatment strategies is associated with better clinical outcomes. The findings concerning treatment integrity were replicated in both hierarchical and non-hierarchical study designs. This specific finding is novel and noteworthy because hierarchical designs better control for non-specific therapist effects, which are likely to reflect therapist interpersonal skills and attitudes (Wampold & Owen, 2021). As such, after partialling out the variance that may be due to such unspecific factors, technical skills reflected by the integrity construct are relevant to effective treatment with a moderate effect size. In terms of caution, it is important to also note the relatively small number of effect sizes included in both the non-hierarchical (four) and hierarchical (three) integrity-outcome meta-analyses. In addition, the definitional and operationalisation inconsistencies found across studies of the integrity/fidelity construct should be noted, and then caution should be used when interpreting the integrity-outcome findings (discussed further in limitations below).

The above findings were derived from a set of studies with generally low risk of bias, and risk of bias ratings were not found to moderate the observed ACI-outcome correlations. The applicability and construct validity of the RoB scale used, however, may have been limited (see Strengths and Limitations below) which may have reduced the validity of this moderator. GRADE assessments indicated that the meta-analyses in non-hierarchical designs had moderate to high reliability, whereas analyses in hierarchical designs had very low to low reliability (although this mainly reflects the small number of studies examining this newer concept).

Significant Moderators of Correlations Across Studies

It is important to note that no direct comparisons or post-hoc tests were conducted in the meta-analytic moderator analyses, and thus only moderators of the relative strength and direction of

associations can be discussed. The strongest positive competence-outcome correlations were found in studies of psychotherapies for depression and anxiety. This is consistent with prior meta-analyses that indicated stronger competence-outcome correlations in depression studies (Webb et al., 2010; Zarafonitis-Müller et al., 2014). Because depression and anxiety are common mental health problems, it is plausible that therapists may get more exposure to working with these problems, and therefore sustained practice and feedback may help to enhance competence. Consistent with prior evidence (Webb et al., 2010), CT and CBT had the strongest positive competence-outcome correlation, with EFT having a negative association. As CT and CBT studies tended to treat anxiety disorders and depression, while EFT was mostly used for psychological trauma, it is possible that the presenting problems may have partly confounded this moderator analysis. Hence, a more plausible and cautious interpretation of the data would suggest that competence is seemingly more important in CBT-oriented interventions, perhaps due to the number and complexity of techniques that are expected to be delivered by CBT therapists. For instance, supporting a patient to make effective use of cognitive restructuring while maintaining a collaborative stance and expressing empathy is perhaps more challenging than simply expressing empathy as the primary task in a therapy session.

Strengths and Limitations

The present study adhered to contemporary good-practice guidelines on systematic reviews of quantitative evidence including: the pre-registration and adherence to a review protocol, searches across multiple databases including grey literature, backward and forward citation searches, risk of bias assessments for primary studies, GRADE assessments of the meta-analyses and the investigation of publication bias and sources of heterogeneity. Despite these strengths, the results should be seen in light of several limitations.

Most limitations were due to the design of primary studies. Studies included in the meta-analyses were primarily grouped according to whether they were classed/defined as measuring ACI constructs relying on the authors' definitions. As some conceptual inconsistency exists in this research

area (e.g., the definitions and demarcation between ACI constructs), the groupings used in the meta-analyses may not have been highly homogenous. Furthermore, no independent verification regarding which ACI construct was being measured in each study was possible due to lack of access to primary study data.

Of particular concern was the inconsistently operationalised construct of treatment integrity, which appeared to be defined and measured in a variety of different ways across primary studies. Some studies, for instance, measured adherence and competence as a composite (combined) measure and did not refer to this construct as integrity (e.g., Haug et al., 2016; Haug et al., 2016), whilst others did refer to a similar composite measure (adherence and competence together) as measuring integrity (e.g., Holder et al., 2018). There were also inconsistencies regarding whether treatment differentiation was actually measured when assessing integrity. Some studies explicitly measured differentiation from other specific modalities (e.g., Gaume et al., 2018 used measures of therapist fidelity that differentiated between two treatment models used in the original trial: Motivational Enhancement Therapy and Social Behavior and Network Therapy). Others generally measured the therapist behaviours that were explicitly non-adherent to the intended model, although did not then specify the alternative modalities being differentiated from (e.g., Spohr et al., 2016). Contrastingly, other authors did not explicitly measure differentiation from other models, although did refer to the construct of interest as integrity (e.g., Haug et al., 2016; Holder et al., 2016).

As for other limitations of included studies, some did not apply gold standard methods to measure ACI constructs (i.e., lack of double-rating, sessions not randomly sampled, lengthy time-lag between rated sessions and time-point of outcome measurement, etc.). Certain ACI rating scales used in these studies had a restricted range, potentially attenuating the measurement of variability in ACI between therapists. Many samples were derived from either clinical trials or trainee psychotherapists, so the generalizability of results to routine practice settings is unclear. As many of the clinical trials applied rigorous methodological procedures to reduce ACI variability amongst therapists (i.e.,

enhanced training, monitoring, supervision, independent rating of recordings), it is likely that there is greater ACI variability in routine care settings due to the frequent absence of these features (Baldwin & Imel, 2013). Hence, the ACI-outcome correlation could in fact be stronger and more important in routine clinical care. An implication of these methodological sources of uncertainty is that the correlations reported in this meta-analysis could be smaller than the "true effect" of ACI, particularly in routine care.

Some limitations were also due to methodological aspects of the meta-analytic review. Primary studies may have controlled for therapeutic alliance in the ACI-outcome analyses, although if they also reported an unadjusted correlation between ACI-outcome, the latter was extracted for meta-analysis, to ensure that we could measure the specific correlation of interest with greater precision. However, not all studies controlling for the therapeutic alliance (or other confounders) also reported unadjusted correlations, which limited the precision of correlations extracted from some samples. The outcome measure time-point closest to the cessation of therapy was given preference over later time-points. The results of the review are therefore specific to exploring short-term (i.e., post-treatment) psychotherapy outcome, as longer-term follow-up outcomes were sparser. Furthermore, it was not possible to consistently extract or to analyse data on duration or number of therapy sessions and so it is still unknown as to whether these factors had a moderating effect on ACI-outcome associations. It is also important to highlight that the NOS RoB scale used was originally designed to evaluate epidemiological/medical studies and not psychotherapy studies. The measure is therefore likely to have had poor construct validity for the studies included in the review. More suitable RoB tools exist that were not chosen. Finally, a large number of moderator analyses (subgroup and meta-regression analyses) were conducted, increasing the risk of type 1 errors (Mascha, 2015) and so we limited our interpretation to the models with lowest risk of bias.

To address these shortcomings, future studies of ACI-outcome associations should: (1) aim to improve consistency in definitions and operationalisations of therapist adherence, competence and

integrity/fidelity (e.g., specifically by avoiding use of the term 'treatment integrity/fidelity' unless all three aspects of adherence, competence and differentiation are being measured together), (2) measure ACI constructs using the methodological standards proposed by Hogue et al. (1996) and listed above, (3) use hierarchical study designs with adequate sample sizes at each level and apply multilevel modelling, (4) control for common factors such the alliance and report adjusted/unadjusted results, (5) always report treatment duration/number of sessions delivered and (6) test if treatment duration or treatment modality (in studies with more than one treatment) moderate the ACI-outcome correlation. As it is plausible that the ACI-outcome correlation may be larger in routine-care settings, large sample naturalistic ACI-outcome studies using a range of psychotherapies with sufficient follow-up will be highly valuable in future.

Clinical and Theoretical Implications

It has been argued that small effect sizes attributable to ACI – such as those found in this review – are negligible, at least relative to broader relational and common factors (Wampold, 2015). However, the upper boundary of these effects (e.g., $r = 0.23$ for integrity) in hierarchical studies (which better control for therapist effects) equates to approximately 5.3% explained variance. This explained variance is equivalent to the estimated magnitude of therapist effects (Johns et al., 2019; 5%). Similarly, this estimate is comparable to the explained variance attributable to common factors such as the alliance (7.7%; Flückiger et al., 2018), goal consensus (5.7%; Tyron et al., 2018), empathy (7.8%; Elliott et al., 2018), positive regard (3.1%; Farber et al., 2018); congruence/genuineness (5.3%; Kolden et al., 2018) and perceived credibility (1.4%; Constantino et al., 2018). From this viewpoint, treatment integrity could potentially be as important as these wider therapist interpersonal skills. However, more replications of the integrity-outcome association would need to occur before integrity can take its place alongside some of the more established process-outcome associations (e.g., the therapeutic alliance). Additionally, it may be that integrity is a construct that is more suited to formal comparative outcome studies (e.g., RCTs) as opposed to routine clinical practice settings. This is because differentiation can

distinguish between treatments in a trial, but may be less relevant when differentiating between the possibly wide number of modalities that therapists may use in routine practice. Assessing and monitoring treatment integrity has clear value in clinical trials, including enhancing construct validity, internal/external validity and validity of statistical conclusions (Leichsenring et al., 2011). A potentially more relevant and clinically useful construct in routine practice maybe therapist competence. To date, competency appears to have been more clearly conceptualised and has been more consistently positively associated with patient outcome (Zarafonitis-Müller et al., 2014). The specific competence-outcome correlation in the treatment of anxiety problems ($r = 0.34$, ~d = .72, in Table 3) was considerably larger (explained variance = 11.6%) than all of the therapist relational skills listed above. As others have argued (e.g., Anderson et al., 2016), a balanced and realistic conclusion is that relational and technical skills are interwoven, and they reciprocally influence and potentiate each other. The provision of a convincing and coherent rationale for a treatment (a specific/technical skill) is likely to amplify perceived credibility and to foster goal consensus (common factors). Conversely, the genuine demonstration of empathy and positive regard by the therapist (common factors) will likely support, for example, anxious patients to engage with a challenging treatment procedure such as exposure (a specific/technical skill) more successfully. Exploring how therapists deliver treatment and the relationship to the outcomes produced is still a relatively new area of investigation and there remains much to be uncovered regarding these phenomena. For instance, a linear relationship between adherence and outcome has not been found (including in the current study). A curvilinear relationship has been found where medium adherence to the treatment manual may lead to optimal clinical outcomes compared to low or high adherence (Barber et al., 2006). Furthermore, an adherent therapist may be more impactful on patient outcome when the alliance is rated as low, as compared to when alliance is rated as high (Barber et al., 2008).

We argue that competence/integrity-outcome associations increase through the integration of (1) detailed awareness of psychological treatment protocols, (2) routine outcome monitoring, (3)

generation of regular and routine self and supervisor ACI ratings using valid and reliable measures, (4) provision of corrective and specific supervisor feedback and (5) deliberate and sustained practice of micro-skills. Research in the fields of clinical supervision/training (Bradley & Becker, 2021; Milne et al., 2011) and deliberate practice (Chow et al., 2015) indicate that the above methods optimise learning transfer and lead to measurable improvements in clinical skills and thus therapist competence. The present findings lend support to clinical guidelines that require practicing therapists to formally monitor and demonstrate competence and/or integrity as part of their accreditation and employment requirements (e.g., National Collaborating Centre for Mental Health, 2020). This quality assurance policy seems especially relevant for the treatment of depression and anxiety – where the competence/integrity-outcome association is seemingly most important.

Conclusions

Therapist competence is associated with better clinical outcomes for adults who have received psychotherapy. This finding is a replication of a previous meta-analysis of CBT process-outcome studies specifically (Zarafonitis- Müller et al., 2014). Furthermore, treatment integrity, as defined by the skilful implementation of evidence-based treatment strategies with high differentiation may be associated with better psychotherapy outcomes, particularly during the treatment of depression and anxiety. Although the magnitude of competence-outcome and integrity-outcome associations is small-to-moderate in statistical terms, these may be comparable to effect sizes for wider relational skills. The integrity-outcome findings are preliminary and conceptual clarification and replication of this association is needed in future. Some uncertainties remain about potential moderators of ACI-outcome relationships and its true magnitude in routine-practice settings. Further research is clearly required.

Data Transparency Statement

This manuscript is a review of the literature and thus data were derived from existing resources. The data in this review have not been published previously. The methods used in this review and meta-analysis are available in the Method and within the Supplementary Materials, including the search

strategy, databases searched, eligibility criteria, characteristics of included studies and associated effect sizes. Any additional data is available from the corresponding author upon reasonable request.

References

References marked with an asterisk indicate studies included in the main meta-analyses. Two asterisks indicate studies included in the sensitivity analyses.

*Abel, A., Hayes, A. M., Henley, W., & Kuyken, W. (2016). Sudden Gains in Cognitive-Behavior Therapy for Treatment-Resistant Depression: Processes of Change. *Journal of Consulting and Clinical Psychology, 84*(8), 726–737. https://doi.org/10.1037/ccp0000101

*Ablon, J. S., Levy, R. A., & Katzenstein, T. (2006). Beyond brand names of psychotherapy: Identifying empirically supported change processes. *Psychotherapy, 43*(2), 216–231. https://doi.org/10.1037/0033-3204.43.2.216

Anderson, T., McClintock, A. S., Himawan, L., Song, X., & Patterson, C. L. (2016). A prospective study of therapist facilitative interpersonal skills as a predictor of treatment outcome. *Journal of Consulting and Clinical Psychology, 84*(1), 57–66. https://doi.org/10.1037/ccp0000060

Atkins, D., Briss, P. A., Eccles, M., Flottorp, S., Guyatt, G. H., Harbour, R. T., ... & Williams, J. W. (2005). Systems for grading the quality of evidence and the strength of recommendations II: pilot study of a new system. *BMC Health Services Research*, *5*(1), 1-12. https://doi.org/10.1186/1472-6963-5-25

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

*Barber, J. P., Crits-Christoph, P., & Luborsky, L. (1996). Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. *Journal of Consulting and Clinical Psychology, 64*(3), 619–622. https://doi.org/10.1037/0022-006X.64.3.619

*Barber, J. P., Gallop, R., Crits-Christoph, P., Barrett, M. S., Klostermann, S., McCarthy, K. S., & Sharpless, B.A. (2008). The role of the alliance and techniques in predicting outcome of

supportive-expressive dynamic therapy for cocaine dependence. *Psychoanalytic Psychology, 25*(3), 461–482. https://doi.org/10.1037/0736-9735.25.3.461

**Barber, J. P., Gallop, R., Crits-Christoph, P., Frank, A., Thase, M. E., Weiss, R. D., & Gibbons, M. B. (2006). The role of therapist adherence, therapist competence, and the alliance in predicting outcome of individual drug counseling: Results from the NIDA Collaborative Cocaine Treatment Study. *Psychotherapy Research, 16,* 229–240. https://doi.org/10.1080/10503300500288951

Barber, J. P., Gallop, R., Crits-Christoph, P., Frank, A., Thase, M. E., Weiss, R. D., & Beth Connolly Gibbons, M. (2006). The role of therapist adherence, therapist competence, and alliance in predicting outcome of individual drug counseling: Results from the National Institute Drug Abuse Collaborative Cocaine Treatment Study. *Psychotherapy Research, 16*(02), 229-240. https://doi.org/10.1080/10503300500288951

Barkham, M., Lambert, M.J. (2021). The Efficacy and Effectiveness of the Psychological Therapies. In M. Barkham, W. Lutz, L.G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change, 7th ed*. Wiley.

*Bisseling, E. M., Schellekens, M. P. J., Spinhoven, P., Compen, F. R., Speckens, A. E. M., & Lee, M. L. van der. (2019). Therapeutic alliance—not therapist competence or group cohesion—contributes to reduction of psychological distress in group-based mindfulness-based cognitive therapy for cancer patients. *Clinical Psychology and Psychotherapy*, *26*(3), 309–318. https://doi.org/10.1002/cpp.2352

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

*Boyle, K., Deisenhofer, A.-K., Rubel, J. A., Bennemann, B., Weinmann-Lutz, B., & Lutz, W. (2020). Assessing treatment integrity in personalized cbt: The inventory of therapeutic interventions

and skills. *Cognitive Behaviour Therapy, 49*(3), 210-227. https://doi.org/10.1080/16506073.2019.1625945

Bradley, W. J., & Becker, K. D. (2021). Clinical supervision of mental health services: a systematic review of supervision characteristics and practices associated with formative and restorative outcomes. *The Clinical Supervisor, 40*(1), 88-111. https://doi.org/10.1080/07325223.2021.1904312

*Branson, A., Shafran, R., & Myles, P. (2015). Investigating the relationship between competence and patient outcome with CBT. *Behaviour Research and Therapy, 68*, 19–26. https://doi.org/10.1016/j.brat.2015.03.002

Breitenstein, S. M., Gross, D., Garvey, C. A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in Nursing & Health, 33*(2), 164-173. https://doi.org/10.1002/nur.20373

Branson, M., Shafran, R. & Myles, P. (2015). Investigating the relationship between competence and patient outcome with CBT. *Behaviour Research and Therapy, 68,* 19-26. https://doi.org/10.1016/j.brat.2015.03.002

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *66*(1), 7-18. https://doi.org/10.1037/0022-006X.66.1.7

*Chevron, E. S., & Rounsaville, B. J. (1983). Evaluating the clinical skills of psychotherapists. A comparison of techniques. *Archives of General Psychiatry, 40*(10), 1129–1132. https://doi.org/10.1001/archpsyc.1983.01790090091014

Chow, D. L., Miller, S. D., Seidel, J. A., Kane, R. T., Thornton, J. A., & Andrews, W. P. (2015). The role of deliberate practice in the development of highly effective psychotherapists. *Psychotherapy, 52*(3), 337–345. https://doi.org/10.1037/pst0000015

Collyer, H., Eisler, I., & Woolgar, M. (2019). Systematic literature review and meta-analysis of the relationship between adherence, competence and outcome in psychotherapy for children and

adolescents. *European Child & Adolescent Psychiatry, 29.* 417-431. https://doi.org/10.1007/s00787-018-1265-2

Constantino, M. J., Coyne, A. E., Boswell, J. F., Iles, B. R., & Vîslă, A. (2018). A meta-analysis of the association between patients' early perception of treatment credibility and their posttreatment outcomes. *Psychotherapy, 55*(4), 486–495. https://doi.org/10.1037/pst0000168

Deeks, J. J., Higgins, J. P., Altman, D. G. (2019). Analysing data and undertaking meta-analyses. In J.P. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M.J. Page, V.A. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions* (pp. 241-284). John Wiley & Sons. https://doi.org/10.1002/9781119536604.ch10

*Despland, J.-N., Roten, Y. de, Drapeau, M., Currat, T., Beretta, V., & Kramer, U. (2009). The role of alliance in the relationship between therapist competence and outcome in brief psychodynamic psychotherapy. *The Journal of Nervous and Mental Disease, 197*(5), 362–367. https://doi.org/10.1097/NMD.0b013e3181a20849

*Easden, M. H., & Fletcher, R. B. (2018). Therapist competence in case conceptualization and outcome in cbt for depression. *Psychotherapy Research, 30*(2), 151-169. https://doi.org/10.1080/10503307.2018.1540895

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

Elliott, R., Bohart, A. C., Watson, J. C., & Murphy, D. (2018). Therapist empathy and client outcome: An updated meta-analysis. *Psychotherapy, 55*(4), 399–410. https://doi.org/10.1037/pst0000175

Farber, B. A., Suzuki, J. Y., & Lynch, D. A. (2018). Positive regard and psychotherapy outcome: A meta-analytic review. *Psychotherapy, 55*(4), 411–423. https://doi.org/10.1037/pst0000171

*Farmer, C. C., Mitchell, K. S., Parker-Guilbert, K., & Galovski, T. E. (2016). Fidelity to the cognitive processing therapy protocol: Evaluation of critical elements. *Behavior Therapy, 48*(2), 195-206. https://doi.org/10.1016/j.beth.2016.02.009

*Feeley, M., DeRubeis, R. J., & Gelfand, L. A. (1999). The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 67*(4), 578–582. https://doi.org/10.1037/0022-006X.67.4.578

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy, 55*(4), 316–340. https://doi.org/10.1037/pst0000172

*Folke, S., Daniel, S. I. F., Gondan, M., Lunn, S., Takker, L., & Poulsen, S. (2017). Therapist adherence is associated with outcome in cognitive-behavioral therapy for bulimia nervosa. *Psychotherapy, 54*(2), 195–200. https://doi.org/10.1037/pst0000107

**Gaston, L., Piper, W. E., Debbane, E. G., Bienvenu, J. P., & Garant, J. (1994). Alliance and technique for predicting outcome in short and long term analytic psychotherapy. *Psychotherapy Research, 4*, 121–135. https://doi.org/10.1080/10503309412331333952

*Gaston, L., & Ring, J. M. (1992). Preliminary results on the inventory of therapeutic strategies. *The Journal of Psychotherapy Practice and Research, 1*(2), 135.

**Gaston, L., Thompson, L., Gallagher, D., Cournoyer, L., & Gagnon, R. (1998). Alliance, technique, and their interactions in predicting outcome of behavioral, cognitive, and brief dynamic therapy. *Psychotherapy Research, 8*, 190–209. https://doi.org/10.1080/10503309812331332307

*Gaume, J., Gmel, G., Faouzi, M., & Daeppen, J.-B. (2009). Counselor skill influences outcomes of brief motivational interventions. *Journal of Substance Abuse Treatment, 37*(2), 151–159. https://doi.org/10.1016/j.jsat.2008.12.001

*Gaume, J., Heather, N., Tober, G., & McCambridge, J. (2018a). A mediation analysis of treatment processes in the UK alcohol treatment trial. *Journal of Consulting and Clinical Psychology, 86*(4), 321–329. https://doi.org/10.1037/ccp0000287

*Gaume, J., Heather, N., Tober, G., & McCambridge, J. (2018b). A mediation analysis of treatment processes in the UK alcohol treatment trial. *Journal of Consulting and Clinical Psychology, 86*(4), 321–329. https://doi.org/10.1037/ccp0000287

Gearing, R. E., El-Bassel, N., Ghesquiere, A., Baldwin, S., Gillies, J., & Ngeow, E. (2011). Major ingredients of fidelity: a review and scientific guide to improving quality of intervention research implementation. *Clinical Psychology Review*, *31*(1), 79-88. https://doi.org/10.1016/j.cpr.2010.09.007

*Gibbons, C. J., Nich, C., Steinberg, K., Roffman, R. A., Corvino, J., Babor, T. F., & Carroll, K. M. (2010). Treatment process, alliance and outcome in brief versus extended treatments for marijuana dependence. *Addiction, 105*(10), 1799–1808. https://doi.org/10.1111/j.1360-0443.2010.03047.x

Ginzburg, D. M., Bohn, C., Höfling, V., Weck, F.,

Clark, D. M., & Stangier, U. (2012). Treatment specific competence predicts outcome in cognitive therapy for social anxiety disorder. *Behaviour Research and Therapy*, *50*(12), 747-752. https://doi.org/10.1016/j.brat.2012.09.001

*Goldman, G. A., & Gregory, R. J. (2009). Preliminary relationships between adherence and outcome in dynamic deconstructive psychotherapy. *Psychotherapy: Theory, Research, Practice, Training, 46*(4), 480–485. https://doi.org/10.1037/a0017947

*Gower, P.J. (2011). *Therapist competence, case conceptualisation and therapy outcome in cognitive behavioural therapy* [Unpublished doctoral thesis]. College of Life and Environmental Sciences, University of Exeter.

*Guydish, J., Campbell, B. K., Manuel, J. K., Delucchi, K. L., Le, T., Peavy, K. M., & McCarty, J. (2014). Does treatment fidelity predict client outcomes in 12-Step Facilitation for stimulant abuse? *Drug and Alcohol Dependence, 134*, 330–336. https://doi.org/10.1016/j.drugalcdep.2013.10.020

*Hall, I.E. (2007). *Therapist relationship and technical skills in two versions of emotion focused trauma therapy* [Unpublished doctoral thesis]. Faculty of Graduate Studies, University of Windsor.

*Haug, T., Nordgreen, T., Ost, L.G., Tangen, T., Kvale, G., Hovland, O. J., Heiervang, E. R., & Havik, O. E. (2016). Working alliance and competence as predictors of outcome in cognitive behavioral therapy for social anxiety and panic disorder in adults. *Behaviour Research and Therapy, 77*, 40–51. https://doi.org/10.1016/j.brat.2015.12.004

*Hauke, C., Gloster, A. T., Gerlach, A. L., Richter, J., Kircher, T., Fehm, L., Stoy, M., Lang, T., Klotsche, J., Einsle, F., Deckert, J., & Wittchen, H.-U. (2014). Standardized treatment manuals: Does adherence matter? Sensoria: A Journal of Mind, *Brain & Culture, 10*(2), 1–13. https://doi.org/10.7790/sa.v0i0.362

Heinonen, E., & Nissen-Lie, H. A. (2020). The professional and personal characteristics of effective psychotherapists: A systematic review. *Psychotherapy Research, 30*(4), 417-432. https://doi.org/10.1080/10503307.2019.1620366

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

Hill, C. E., O'Grady, K. E., & Elkin, I. (1992). Applying the Collaborative Study Psychotherapy Rating Scale to rate therapist adherence in cognitive-behavior therapy, interpersonal therapy, and clinical management. *Journal of Consulting and Clinical Psychology, 60*(1), 73–79. https://doi.org/10.1037/0022-006X.60.1.73

*Hilsenroth, M. J., Ackerman, S. J., Blagys, M. D., Baity, M. R., & Mooney, M. A. (2003). Short-term psychodynamic psychotherapy for depression: An examination of statistical, clinically significant, and technique-specific change. *The Journal of Nervous and Mental Disease, 191*(6), 349–357. http://doi.org/10.1097/01.NMD.0000071582.11781.67

*Hoffart, A., Sexton, H., Nordahl, H. M., & Stiles, T. C. (2005). Connection between patient and therapist and therapist's competence in schema-focused therapy of personality problems. *Psychotherapy Research, 15*(4), 409–419. https://doi.org/10.1080/10503300500091702

Hogue, A., Liddle, H. A., & Rowe, C. (1996). Treatment adherence process research in family therapy: A rationale and some practical guidelines. *Psychotherapy: Theory, Research, Practice, Training*, *33*(2), 332-345. https://doi.org/10.1037/0033-3204.33.2.332

*Holder, N., Holliday, R., Williams, R., Mullen, K., & Surís, A. (2018). A preliminary examination of the role of psychotherapist fidelity on outcomes of cognitive processing therapy during an RCT for military sexual trauma-related PTSD. *Cognitive Behaviour Therapy, 47*(1), 76–89. https://doi.org/10.1080/16506073.2017.1357750

*Huppert, J. D., Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S.W. (2006). The Interaction of Motivation and Therapist Adherence Predicts Outcome in Cognitive Behavioral Therapy for Panic Disorder: Preliminary Findings. *Cognitive and Behavioral Practice, 13*(3), 198–204. https://doi.org/10.1016/j.cbpra.2005.10.001

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

*Katz, M., Hilsenroth, M. J., Gold, J. R., Moore, M., Pitman, S. R., Levy, S. R., & Owen, J. (2019). Adherence, flexibility, and outcome in psychodynamic treatment of depression. *Journal of Counseling Psychology, 66*(1), 94–103. https://doi.org/10.1037/cou0000299

*Kazantzis, N., Clayton, X., Cronin, Timothy. J., Farchione, D., Limburg, K., & Dobson, K.S. (2018). The Cognitive Therapy Scale and Cognitive Therapy Scale-Revised as measures of therapist competence in cognitive behavior therapy for depression: Relations with short and long term outcome. *Cognitive Therapy and Research, 42*(4), 385–397. https://doi.org/10.1007/s10608-018-9919-4

Kazdin, A. E. (1986). Comparative outcome studies of psychotherapy: methodological issues and strategies. *Journal of Consulting and Clinical Psychology*, *54*(1), 95-105. https://doi.org/10.1037/0022-006X.54.1.95

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, *3*, 1-27. https://doi.org/10.1146/annurev.clinpsy.3.022806.091432

Kolden, G. G., Wang, C.-C., Austin, S. B., Chang, Y., & Klein, M. H. (2018). Congruence/genuineness: A meta-analysis. *Psychotherapy, 55*(4), 424–433. https://doi.org/10.1037/pst0000162

*Kramer Schmidt, L., Moyers, T. B., Nielsen, A. S., & Andersen, K. (2019). Is fidelity to motivational interviewing associated with alcohol outcomes in treatment-seeking 60+ year-old citizens? *Journal of Substance Abuse Treatment, 101*, 1–11. https://doi.org/10.1016/j.jsat.2019.03.004

*Kuyken, W., & Tsivrikos, D. (2009). Therapist competence, comorbidity and cognitive-behavioral therapy for depression. *Psychotherapy and Psychosomatics, 78*(1), 42-48. https://doi.org/ 10.1159/000172619

Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, *33*(2), 363-374. https://doi.org/10.2307/2529786

Lemma, A., Roth, A.D., & Pilling, S. (2008). The competences required to deliver effective psychoanalytic/ psychodynamic therapy. University College London. https://www.ucl.ac.uk/drupal/site_pals/sites/pals/files/migrated-files/PPC_Clinicians_Background_Paper.pdf

Lenhard, W. & Lenhard, A. (2016). *Computation of Effect Sizes*. Psychometrica. https://doi.org/10.13140/RG.2.2.17823.92329

Leichsenring, F., Salzer, S., J Hilsenroth, M., Leibing, E., Leweke, F., & Rabung, S. (2011). Treatment integrity: An unresolved issue in psychotherapy research. *Current Psychiatry Reviews*, *7*(4), 313-321. https://doi.org/10.2174/157340011797928259

Light, R. J., & Pillemer, D. B. (1984). *Summing up: the science of reviewing research*. Harvard University Press.

*Liness, S., Beale, S., Lea, S., Byrne, S., Hirsch, C. R., & Clark, D.M. (2019a). The sustained effects of CBT training on therapist competence and patient outcomes. *Cognitive Therapy and Research, 43*(3), 631–641. https://doi.org/10.1007/s10608-018-9987-5

*Liness, S., Beale, S., Lea, S., Byrne, S., Hirsch, C. R., & Clark, D. M. (2019b). Multi-professional IAPT CBT training: Clinical competence and patient outcomes. *Behavioural and Cognitive Psychotherapy, 47*(6), 672–685. https://doi.org/10.1017/S1352465819000201

**Loeb, K. L., Wilson, G. T., Labouvie, E., Pratt, E. M., Hayaki, J., Walsh, B. T., . . . Fairburn, C. G. (2005). Therapeutic alliance and treatment adherence in two interventions for bulimia nervosa: A study of process and outcome. *Journal of Consulting and Clinical Psychology, 73*, 1097–1107. https://doi.org/10.1037/0022-006X.73.6.1097

*Luborsky, L., McLellan, A. T., Woody, G. E., O'Brien, C. P., & Auerbach, A. (1985). Therapist success and its determinants. *Archives of General Psychiatry, 42*(6), 602–611. http://doi.org/10.1001/archpsyc.1985.01790290084010

*Marques, L., Valentine, S. E., Kaysen, D., Mackintosh, M. A., Dixon De Silva, L. E., Ahles, E. M., Youn, S. J., Shtasel, D. L., Simon, N. M., & Wiltsey-Stirman, S. (2019). Provider fidelity and modifications to cognitive processing therapy in a diverse community health clinic: Associations with clinical change. *Journal of Consulting and Clinical Psychology, 87*(4), 357–369. https://doi.org/10.1037/ccp0000384

*Martino, S., Ball, S. A., Nich, C., Frankforter, T. L., & Carroll, K. M. (2008). Community program therapist adherence and competence in motivational enhancement therapy. *Drug and Alcohol Dependence, 96*(1-2), 37–48. https://doi.org/10.1016/j.drugalcdep.2008.01.020

*Marziali, E. A. (1984). Prediction of outcome of brief psychotherapy from therapist interpretive interventions. *Archives of General Psychiatry, 41*(3), 301-304. https://doi.org/10.1001/archpsyc.1984.01790140091011

Mascha, E. J. (2015). Alpha, beta, meta: guidelines for assessing power and type I error in meta-analyses. *Anesthesia & Analgesia, 121*(6), 1430-1433. https://doi.org/10.1213/ANE.0000000000000993

*McCarthy, K. S., Keefe, J. R., & Barber, J. P. (2016). Goldilocks on the couch: Moderate levels of psychodynamic and process-experiential technique predict outcome in psychodynamic therapy. *Psychotherapy Research, 26*(3), 307–317. https://doi.org/10.1080/10503307.2014.973921

McLeod, B. D., Southam-Gerow, M. A., Rodríguez, A., Quinoy, A. M., Arnold, C. C., Kendall, P. C., & Weisz, J. R. (2018). Development and initial psychometrics for a therapist competence instrument for CBT for youth anxiety. *Journal of Clinical Child & Adolescent Psychology,* 47(1), 47-60. https://doi.org/10.1080/15374416.2016.1253018.

McLeod, B. D. Southam-Gerow, M.A., Jensen-Doss, A., Hogue, A., Kendall, P. C. & Weisz, J. R. (2019). Benchmarking treatment adherence and therapist competence in individual cognitive-behavioral treatment for youth anxiety disorders. *Journal of Clinical Child & Adolescent Psychology, 48*, S234 S246. https://doi.org/10.1080/15374416.2017.1381914

Miller, S. J., & Binder, J. L. (2002). The effects of manual-based training on treatment fidelity and outcome: A review of the literature on adult individual psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, *39*(2), 184-198. https://doi.org/10.1037/0033-3204.39.2.184

Milne, D. L., Sheikh, A. I., Pattison, S., & Wilkinson, A. (2011). Evidence-based training for clinical supervisors: A systematic review of 11 controlled studies. *The Clinical Supervisor, 30*(1), 53-71. https://doi.org/10.1080/07325223.2011.564955

*Minonne, G. A. (2009). *Therapist adherence, patient alliance, and depression change in the NIMH Treatment for Depression Collaborative Research Program* [Unpublished doctoral thesis]. The University of Michigan.

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. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

National Collaborating Centre for Mental Health. (2020). *The Improving Access to Psychological Therapies Manual*, version 4, updated March 2020. https://www.england.nhs.uk/publication/the-improving-access-to-psychological-therapies-manual/

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

*Norrie, J., Davidson, K., Tata, P., & Gumley, A. (2013). Influence of therapist competence and quantity of cognitive behavioural therapy on suicidal behaviour and inpatient hospitalisation in a randomised controlled trial in borderline personality disorder: Further analyses of treatment effects in the BOSCOT study. *Psychology and Psychotherapy, 86*(3), 280–293. https://doi.org/10.1111/papt.12004

*Ogrodniczuk, J.S. (1997). *Therapist adherence to treatment manuals and its relation to the therapeutic alliance and therapy outcome: scale development and validation* [Unpublished

doctoral thesis]. Faculty of Graduate Studies and Research: Medical Sciences, University of Alberta.

** O'Malley, S. S., Foley, S. H., Rounsaville, B. J., Watkins, J. T., Sotsky, S. M., Imber, S. D., & Elkin, I. (1988). Therapist competence and patient outcome in interpersonal psychotherapy of depression. *Journal of Consulting and Clinical Psychology, 56*, 496–501. https://doi.org/10.1037/0022-006X.56.4.496

Orlinsky, D. E., & Howard, K. I. (1986). The psychological interior of psychotherapy: Explorations with the Therapy Session Reports. In L. S. Greenberg & W. M. Pinsof (Eds.), *Guilford clinical psychology and psychotherapy series. The psychotherapeutic process: A research handbook* (p. 477–501). Guilford Press.

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

Parry, G., Bennett, D., Roth, A. D., & Kellett, S. (2021). Developing a competence framework for cognitive analytic therapy. *Psychology and Psychotherapy: Theory, Research and Practice*, *94*(1), 151-170. https://doi.org/10.1111/papt.12306

*Patton, J.R. (1997). *A study in reliability and validation: what is the nature of the unique and joint contributions of therapist adherence and therapeutic alliance to treatment outcome?* [Unpublished doctoral thesis]. Graduate Faculty in Psychology, The City University of New York.

Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology*, *75*(6), 829. https://doi.org/10.1037/0022-006X.75.6.829

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice*, 12*,* 365– 383. https://doi.org/10.1093/clipsy.bpi045

*Piper, W. E., Azim, H. F., Joyce, A. S., & McCallum, M. (1991). *Transference interpretations, therapeutic alliance, and outcome in short-term individual psychotherapy*. Archives of General Psychiatry, 48(10), 946-953. https://doi.org/10.1001/archpsyc.1991.01810340078010

**Piper, W. E., Debbane, E. G., Bienvenu, J. P., de Carufel, F., & Garant, J. (1986). Relationships between the object focus of therapist interpretations and outcome in short-term individual psychotherapy. *British Journal of Medical Psychology, 59*, 1–11. https://doi.org/10.1111/j.2044-8341.1986.tb02659.x

Roth, A. D., & Pilling, S. (2008). Using an evidence-based methodology to identify the competences required to deliver effective cognitive and behavioural therapy for depression and anxiety disorders. *Behavioural and Cognitive Psychotherapy*, *36*(2), 129. https://doi.org/10.1017/S1352465808004141

*Sachs, J. S. (1983). Negative factors in brief psychotherapy: An empirical assessment. *Journal of Consulting and Clinical Psychology, 51*(4), 557. https://doi.org/10.1037/0022-006X.51.4.557

*Shaw, B. F., Elkin, I., Yamaguchi, J., Olmsted, M., Vallis, T. M., Dobson, K. S., Lowery, A., Sotsky, S. M., Watkins, J. T., & Imber, S. D. (1999). Therapist competence ratings in relation to clinical outcome in cognitive therapy of depression. *Journal of Consulting and Clinical Psychology, 67*(6), 837–846. https://doi.org/10.1037/0022-006X.67.6.837

*Sinai, D., Gur, M., & Lipsitz, J. D. (2012). Therapist adherence to interpersonal vs. Supportive therapy for social anxiety disorder. *Psychotherapy Research: Journal of the Society for Psychotherapy Research, 22*(4), 381–388. https://doi.org/10.1080/10503307.2012.658452

*Snippe, E., Schroevers, M. J., Tovote, K. A., Sanderman, R., Emmelkamp, P. M. G., & Fleer, J. (2019). Explaining variability in therapist adherence and patient depressive symptom

improvement: The role of therapist interpersonal skills and patient engagement. *Clinical Psychology and Psychotherapy, 26*(1), 84–93. https://doi.org/10.1002/cpp.2332

Southam-Gerow, M. A, McLeod, B. D., Arnold, C. C., Rodríguez, A., Cox, J.R., Reise, S. P., Bonifay, W. E., Weisz, J.R., & Kendall, P. C. (2016). Initial development of a treatment adherence measure for cognitive-behavioral therapy for child anxiety. *Psychological Assessment, 28(1)*, 70-80. https://doi.org/10.1037/pas0000141

*Spektor, D. (2008). *Therapists' adherence to manualized treatments in the context of ruptures* [Unpublished doctoral thesis]. New School for Social Research, The New School.

*Spohr, S. A., Taxman, F. S., Rodriguez, M., & Walters, S. T. (2016). Motivational Interviewing Fidelity in a Community Corrections Setting: Treatment Initiation and Subsequent Drug Use. *Journal of Substance Abuse Treatment, 65*, 20–25. https://doi.org/10.1016/j.jsat.2015.07.012

** Strunk, D. R., Brotman, M. A., & DeRubeis, R. J. (2010). The process of change in cognitive therapy for depression: Predictors of early inter-session symptom gains. *Behaviour Research and Therapy, 48*(7), 599-606. https://doi.org/10.1016/j.brat.2010.03.011

*Strunk, D. R., Brotman, M. A., DeRubeis, R. J., & Hollon, S. D. (2010). Therapist competence in cognitive therapy for depression: Predicting subsequent symptom change. *Journal of Consulting and Clinical Psychology, 78*(3), 429–437. https://doi.org/10.1037/a0019631

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

*Svartberg, M., & Stiles, T. C. (1992). Predicting patient change from therapist competence and patient-therapist complementarity in short-term anxiety-provoking psychotherapy: A pilot study. *Journal of Consulting and Clinical Psychology, 60*(2), 304–307. https://doi.org/10.1037/0022-006X.60.2.304

*Thyrian, J. R., Freyer-Adam, J., Hannöver, W., Röske, K., Mentzel, F., Kufeld, C., Bischof, G., Rumpf, H.-J., John, U., & Hapke, U. (2007). Adherence to the principles of Motivational Interviewing, clients' characteristics and behavior outcome in a smoking cessation and relapse prevention trial in women postpartum. *Addictive Behaviors, 32*(10), 2297–2303. https://doi.org/10.1016/j.addbeh.2007.01.024

*Trepka, C., Rees, A., Shapiro, D. A., Hardy, G. E., & Barkham, M. (2004). Therapist competence and outcome of cognitive therapy for depression. *Cognitive Therapy and Research, 28*(2), 143–157. https://doi.org/10.1023/B:COTR.0000021536.39173.66

Tryon, G. S., Birch, S. E., & Verkuilen, J. (2018). Meta-analyses of the relation of goal consensus and collaboration to psychotherapy outcome. *Psychotherapy, 55*(4), 372–383. https://doi.org/10.1037/pst0000170

*Tschuschke, V., Crameri, A., Koehler, M., Berglar, J., Muth, K., Staczan, P., Von Wyl, A., Schulthess, P., & Koemeda-Lutz, M. (2015). The role of therapists' treatment adherence, professional experience, therapeutic alliance, and clients' severity of psychological problems: Prediction of treatment outcome in eight different psychotherapy approaches. Preliminary results of a naturalistic. *Psychotherapy Research: Journal of the Society for Psychotherapy Research, 25*(4), 420–434. https://doi.org/10.1080/10503307.2014.896055

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, *35*(2), 215-247. https://doi.org/10.3102/1076998609346961

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology*, *61*(4), 620–630. https://doi.org/10.1037/0022-006X.61.4.620

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry, 14*(3), 270-277. https://doi.org/10.1002/wps.20238

Wampold, B. E., & Imel, Z. E. (2015). The great psychotherapy debate: The evidence for what makes psychotherapy work. Routledge.

Wampold, B. E., & Owen, J. (2021). Therapist effects: History, methods, magnitude, and characteristics of effective therapists. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (7th ed., pp. 301-330). Wiley.

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, *78*(2), 200-211. https://doi.org/10.1037/a0018912

*Webb, C. A., Derubeis, R. J., Dimidjian, S., Hollon, S. D., Amsterdam, J. D., & Shelton, R. C. (2012). Predictors of patient cognitive therapy skills and symptom change in two randomized clinical trials: The role of therapist adherence and the therapeutic alliance. *Journal of Consulting and Clinical Psychology, 80*(3), 373–381. https://doi.org/10.1037/a0027663

** Webb, C. A., Gelfand, L. A., DeRubeis, R. J., Amsterdam, J. D., Shelton, R. C., Hollon, S. D., . . . Dimidjian, S. (2009). Mechanisms of change in cognitive therapy for depression: Therapist adherence, symptom change, and the mediating role of patient skills. *Manuscript in preparation*.

*Weck, F., Bohn, C., Ginzburg, D. M., & Stangier, U. (2011). Assessment of adherence and competence in cognitive therapy: Comparing session segments with entire sessions. *Psychotherapy Research: Journal of the Society for Psychotherapy Research, 21*(6), 658–669. https://doi.org/10.1080/10503307.2011.602751

*Weck, F., Richtberg, S., Jakob, M., Neng, J. M. B., & Hofling, V. (2015). Therapist competence and therapeutic alliance are important in the treatment of health anxiety (hypochondriasis). *Psychiatry Research, 228*(1), 53–58. https://doi.org/10.1016/j.psychres.2015.03.042

*Weck, F., Rudari, V., Hilling, C., Hautzinger, M., Heidenreich, T., Schermelleh-Engel, K., & Stangier, U. (2013). Relapses in recurrent depression 1 year after maintenance cognitive-behavioral therapy: The role of therapist adherence, competence, and the therapeutic alliance. *Psychiatry Research, 210*(1), 140–145. https://doi.org/10.1016/j.psychres.2013.05.036

Weck, F., Grikscheit, F., Höfling, V., & Stangier, U. (2014). Assessing treatment integrity in cognitive-behavioral therapy: Comparing session segments with entire sessions. *Behavior Therapy, 45*(4), 541-552. https://doi.org/10.1016/j.beth.2014.03.003

Wells, G., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2000). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottowa Hospital Research Institute. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

*Westra, H. A., Constantino, M. J., Arkowitz, H., & Dozois, D. J. (2011). Therapist differences in cognitive–behavioral psychotherapy for generalized anxiety disorder: A pilot study. *Psychotherapy, 48*(3), 283-292. https://doi.org/10.1037/a0022011

*Wurman, T.V. (2019). *Therapist competence in dynamic interpersonal therapy and its association with treatment outcome* [Unpublished doctoral thesis]. University College London.

Zarafonitis-Müller, S., Kuhr, K., & Bechdolf, A. (2014). The relationship between therapist's competence and adherence to outcome in cognitive-behavioural therapy-results of a metaanalysis. *Fortschritte der Neurologie-psychiatrie*, *82*(9), 502-510. https://doi.org/10.1055/s-0034-1366798

*Zelencich, L. M., Wong, D., Kazantzis, N., McKenzie, D. P., Downing, M., & Ponsford, J. L. (2020). Predictors of Anxiety and Depression Symptom Improvement in CBT Adapted for Traumatic Brain Injury: Pre/Post-Injury and Therapy Process Factors. *Journal of the International Neuropsychological Society, 26*(1), 97–107. https://doi.org/10.1017/S1355617719000791

Table 1

A Summary of Study Characteristics for Studies Using Non-Hierarchical Analyses to Investigate Process-Outcome Associations

Meta-Analysis	Patient N	Therapist N	Treatment Modality	Disorder Treated /Diagnosis	Patient Demographics	Therapist Demographics	% coverage (per patient)	Aspect of measure used	Type of session data rated	Alliance controlled	Country in which study sited	RoB assessment
Non-Hierarchical Adherence-Outcome (k=28, samples n=37)	54.46 (76.06)	10.8 (7.61)	Psychodynamic (18.92%), SEDP (5.41%), CBT (18.92%), CT (13.51%), MI/MET (8.11%), Mixed (5.41%), EFT (8.11%), IPT (5.41%). All remaining were 2.8%: STIP, STSIP, BAP, IET, BRT, ET.	Anxiety Disorders (16.22%), Substance Use/Addiction (16.22%), Depression (29.73%), Mixed (27.03%), Trauma/PTSD (8.11%), BPD (2.7%)	58.91% female, mean age 39.14 (9.55)	63.48% female, mean age 35.98 (4.87)	28.91%	Sub-scale (48.65%), Whole scale (45.95%)	Transcripts (2.7%), Audio recording (54.05%), Video recording (43.24%)	Yes (5.41%), No (94.59%)	USA (59.46%), Germany (16.22%), Canada (24.32%),	Low (64.29%), Moderate (32.14%), High (3.57)
Non-Hierarchical Competence-Outcome (k=27, samples n=32)	127.56 (243.67)	21.81 (43.77)	Psychodynamic (9.38%), CBT (40.63%), CT (12.50%), MI/MET (6.25%), EFT (9.38%). All remaining were 3.13%: CPT, Mixed, Experiential/client-centred, IPT, STAPP, ET, Schema-Focused Therapy	Anxiety Disorders (15.63%), Substance Use/Addiction (12.50%), Depression (28.13%), Mixed (31.25%), Trauma/PTSD (12.50%)	52.93% female, mean age 39.98 (5.9)	72.11% female, mean age 35.81 (5.17)	30.83%	Sub-scale (25.00%), Whole scale (75.00%)	Transcripts (3.13%), Audio recording (50.00%), Video recording (43.75%), NR (3.13%)	Yes (12.5%), No (87.5%)	USA (34.38%), Germany (15.63%), Canada (12.50%), Norway (6.25%), UK (28.13%), Australia (3.13%)	Low (59.26%), Moderate (40.74%), High (0.00%)
Non-Hierarchical Integrity-Outcome (k=3, samples n=4)	134.5 (66.16)	12.00 (14.14)	MI/MET (50.00%), SBNT (25.00%), CBT (25.00%)	Substance Use/Addiction (75.00%),, Anxiety Disorders (25.00%)	36.03% female, mean age 38.38 (4.79)	Mean demographics across studies not calculated as only one study reported	40.63%	Sub-scale (75.00%), Whole scale (25.00%)	Audio recording (25.00%), Video recording (75.00%)	Yes (0.00%), No (100%)	USA (25%), UK (50%), Norway (25%)	Low (100%), Moderate (0.00%), High (0.00%)

Notes. The above table is of effect size samples, not studies. One study/sample may be included across both adherence-outcome and competence-outcome study type as separate meta-analyses were conducted. No participants were included in the same meta-analysis more than once. Some studies did not report particular details and thus the summary characteristics could only be calculated from the studies that did report the relevant information. Risk of bias assessments are the proportions of the individual papers that scored low/moderate/high out of the total number of individual papers in that category (e.g. out of all adherence, non-hierarchical papers). K = total number of individual papers in that category, samples n = the number of individual samples in a particular category (this may be more than the number of individual studies due to some studies having process-outcome effect sizes for numerous sub-samples (e.g. samples receiving different interventions). SEDP = Supportive Expressive Dynamic Psychotherapy, CBT = Cognitive Behaviour Therapy, CT = Cognitive Therapy, IPT = Interpersonal Psychotherapy, MI/MET = Motivational Interviewing/Motivational Enhancement Therapy, EFT = Emotion Focused Therapy, STAGE-12 = Stimulant Abuser Groups to Engage in 12-Step intervention, STDPP = Short Term Psychodynamic Psychotherapy, STIP = Short-Term Interpretative Psychotherapy, STSIP = Short-Term Supportive Individual Psychotherapy, BAP = Brief-Adaptive Psychotherapy, IET = Interpersonal-Experiential Therapy, MBCT = Mindfulness Based Cognitive Therapy. STAPP = Short-Term Anxiety Provoking Psychotherapy, SBNT = Social Behaviour and Network Therapy, UK = United Kingdom, PTSD = Post Traumatic Stress Disorder, ED = Eating Disorders, BPD = Borderline Personality Disorder, USA = United Stated of America.

Table 2

A Summary of Study Characteristics for Studies Using Hierarchical Analyses to Investigate Process-Outcome Associations

Meta-Analysis	Patient N	Therapist N	Treatment Modality	Disorder Treated/Diagnosis	Patient Demographics	Therapist Demographics	% coverage (per patient)	Aspect of measure used	Type of session data rated	Alliance controlled	Country in which study sited	RoB assessment
Hierarchical Adherence-Outcome (k=7, samples n=8)	81.63 (68.77)	22.43 (18.12)	CBT (37.50%) All remaining were 12.50%: Mixed, STAGE-12, STDPP, CPT, MBCT.	Anxiety Disorders (12.50%), Substance Use/Addiction (12.50%), Depression (37.50%), Mixed (12.50%), Trauma/PTSD (12.50%), ED (12.50%)	64.70% female, mean age 33.50 (7.03)	62.30% female, mean age 48.05 (3.32)	36.42%	Sub-scale (25.00%), Whole scale (75.00%)	Audio recording (37.50%), Video recording (50.00%), NR (12.50%)	Yes (25.00%), No (75.00%)	USA (37.50%), Germany (12.50%), The Netherlands (25.00%), Denmark (12.50%), Switzerland (12.50%)	Low (85.71%), Moderate (14.29%), High (0.00%)
Hierarchical Competence-Outcome (k=7, samples n=7)	73.86 (38.79)	11.83 (6.01)	Psychodynamic (28.57%), CBT (28.57%) All remaining were 14.29%: STAGE-12, CPT, MBCT	Substance Use/Addiction (14.29%), Depression (42.86%), Mixed (28.57%), Trauma/PTSD (14.29%)	68.45% female, mean age 40.19 (7.2)	89.63% female, mean age 44.67 (6.49)	63.54%	Sub-scale (28.57%), Whole scale (71.43%)	Audio recording (42.86%), Video recording (42.86%), NR (14.29%)	Yes (42.86%), No (57.14%)	USA (42.86%), The Netherlands (14.29%), Switzerland (14.29%), UK (14.29%), New Zealand (14.29%)	Low (57.14%), Moderate (0.00%), High (42.86%)
Hierarchical Integrity-Outcome (k=3, samples n=3)	135.00 (89.94)	5.33 (1.53)	MI/MET (66.67%), CPT (33.30%)	Substance Use/Addiction (66.67%), Trauma/PTSD (33.33%)	46.87% female, mean age 51.57 (19.18)	Mean demographics across studies not calculated as only one study reported	56.00%	Sub-scale (66.67%), Whole scale (33.33%)	Audio recording (66.67%), Video recording (33.33%)	Yes (0.00%), No (100.00%)	USA (33.33%), Switzerland (33.33%), Multi-National (33.33%)	Low (66.67%), Moderate (33.33%), High (0.00%)

Notes. The above table is of effect size samples, not studies. One study/sample may be included across both adherence-outcome and competence-outcome study type for instance as separate meta-analyses were conducted. No participants were included in the same meta-analysis more than once. Some studies did not report particular details and thus the summary characteristics could only be calculated from the studies that did report the relevant information. Risk of bias assessments are the proportions of the individual papers that scored low/moderate/high out of the total number of individual papers in that category (e.g. out of all adherence, non-hierarchical papers). K = total number of individual papers in that category, samples n = the number of individual samples in a particular category (this may be more than the number of individual studies due to some studies having numerous process-outcome effect sizes for different sub-samples). CBT = Cognitive Behaviour Therapy, MI/MET = Motivational Interviewing/Motivational Enhancement Therapy, STAGE-12 = Stimulant Abuser Groups to Engage in 12-Step intervention, STDPP = Short Term Psychodynamic Psychotherapy, MBCT = Mindfulness Based Cognitive Therapy, UK = United Kingdom, PTSD = Post Traumatic Stress Disorder, USA = United Stated of America.

Table 3

Subgroup Analyses for Non-Hierarchical Adherence-Outcome and Competence-Outcome Effect Sizes

Group	Variable	Subgroup	*k*	Effect Size (*r*)	95% CI	*Q*	I^2 (%)	Diff between subgroups (*p*)
Adherence - outcome								
	Problem Targeted	Anxiety Disorders	6	0.11	-0.19,0.38	14.49**	67.46	0.765
		Depression/ Mood Disorders	11	0.09	-0.06,0.23	13.81	27.57	
		Mixed Problems	10	0.10	-0.13,0.31	22.94	60.77	
		Substance Use/Addiction	6	-0.02	-0.33,0.30	37.52***	86.68	
		Trauma/PTSD	3	-0.11	-0.70,0.58	4.98	59.80	
	Treatment Modality	CBT	7	-0.01	-0.18,0.16	8.70	31.02	0.432
		CT	5	0.23	-0.07,0.50	6.82	41.37	
		EFT	3	-0.11	-0.7,0.58	4.98	59.80	
		IPT	2	0.21	-0.12,0.49	0.07	0.00	
		MI/MET	3	0.05	-0.28,0.37	5.81	65.60	
		Mixed	2	0.03	-1.0,1.0	23.61***	95.76	
		Psychodynamic	7	0.17	-0.27,0.55	29.39	79.58	
		SEDP	2	-0.07	-0.98,0.98	2.77	63.85	
	Alliance Confound	Controlled	2	0.02	-0.64,0.67	0.46	0.00	0.673
		Not Controlled	35	0.07	-0.03,0.17	101.34***	66.45	
Competence - outcome								
	Problem Targeted	Anxiety Disorders	5	0.34	0.00,0.60	9.68*	58.69	**0.002***
		Depression/ Mood Disorders	10	0.25	0.06,0.43	26.91**	66.55	
		Mixed Problems	9	0.21	0.09,0.33	17.32**	53.81	
		Substance Use/Addiction	4	-0.11	-0.36,0.14	10.40*	71.15	
		Trauma/PTSD	4	0.06	-0.35,0.45	6.00	50.02	
	Treatment Modality							
		CBT	13	0.21	0.07,0.34	49.21***	75.61	**<0.001****
		CT	4	0.37	0.15,0.55	2.21	0.00	
		EFT	3	-0.05	-0.49,0.42	1.96	0.00	
		MI/MET	2	0.01	-0.17,0.18	0.10	0.00	
		Psychodynamic	3	0.16	-0.84,0.91	10.58**	81.10	
	Alliance Confound	Not Controlled	28	0.19	0.09,0.29	107.12***	75.58	0.101
		Controlled	4	-0.01	-0.36,0.35	8.37*	64.15	

Notes. CI = Confidence Interval. CBT = cognitive behaviour therapy, CT = cognitive therapy, EFT = emotion focused therapy, IPT = interpersonal psychotherapy, MI/MET = motivational interviewing/motivational enhancement therapy, SEDP = supportive expressive dynamic psychotherapy. *significant at $p < .01$ threshold, **significant at $p < .001$ threshold, between subgroup differences significant at Bonferroni adjusted $p < .017$ threshold for multiple testing (in bold).

Figure 1

Forest Plot for Non-Hierarchical Adherence-Outcome Meta-Analysis

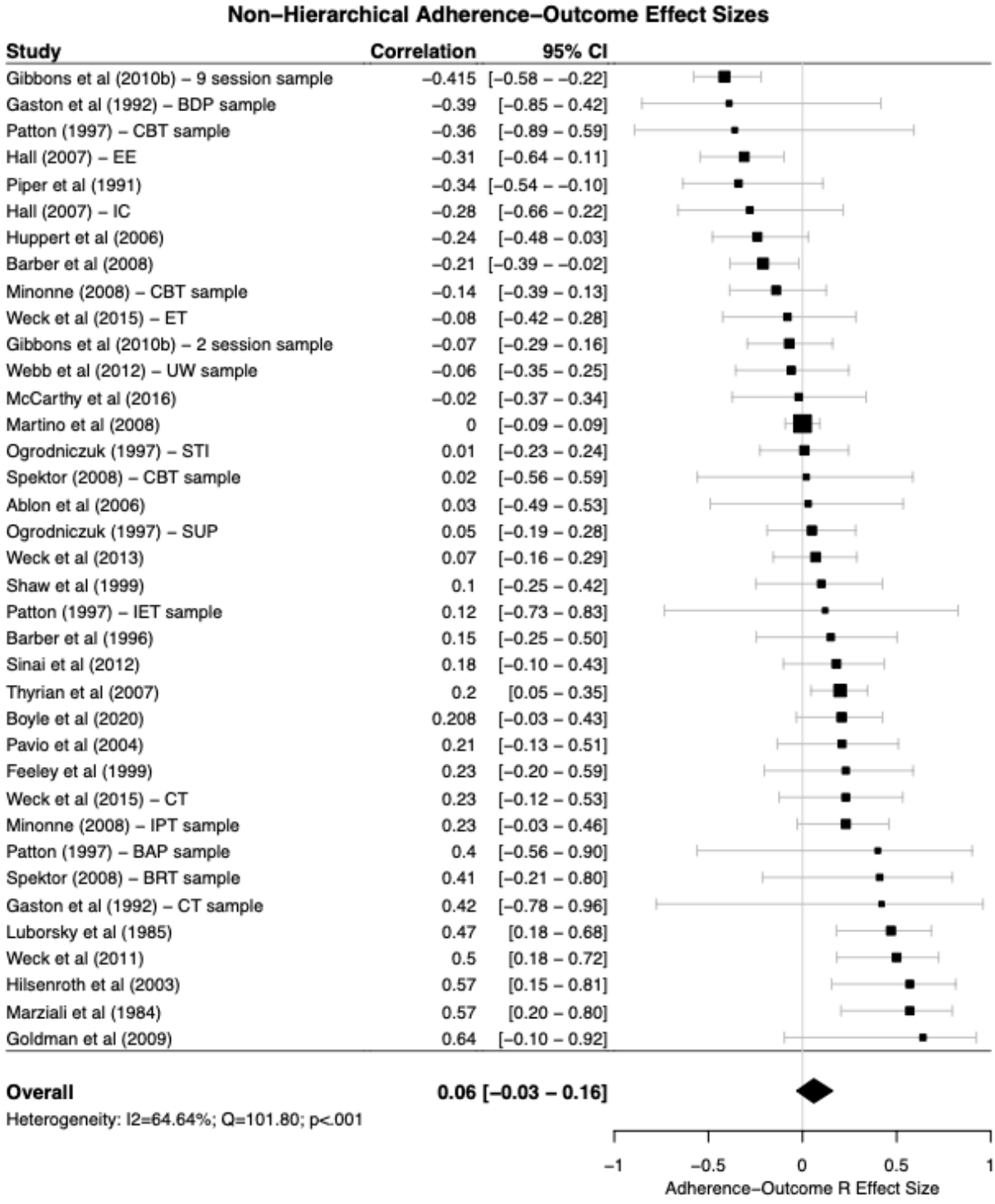

Figure 2

Funnel Plot for Non-Hierarchical Competence-Outcome Meta-Analysis

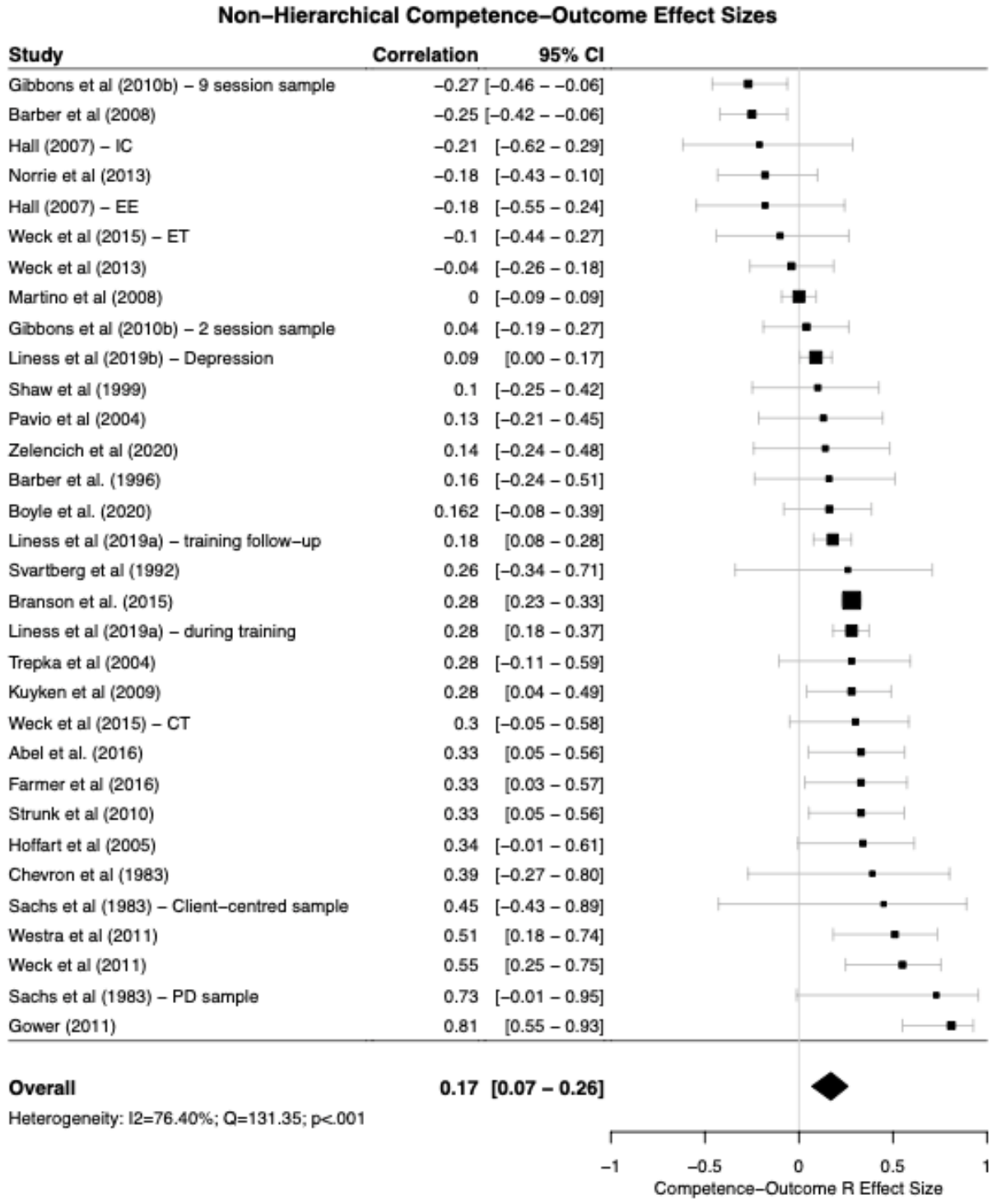

Figure 3

Forest Plot for Non-Hierarchical Integrity-Outcome Meta-Analysis

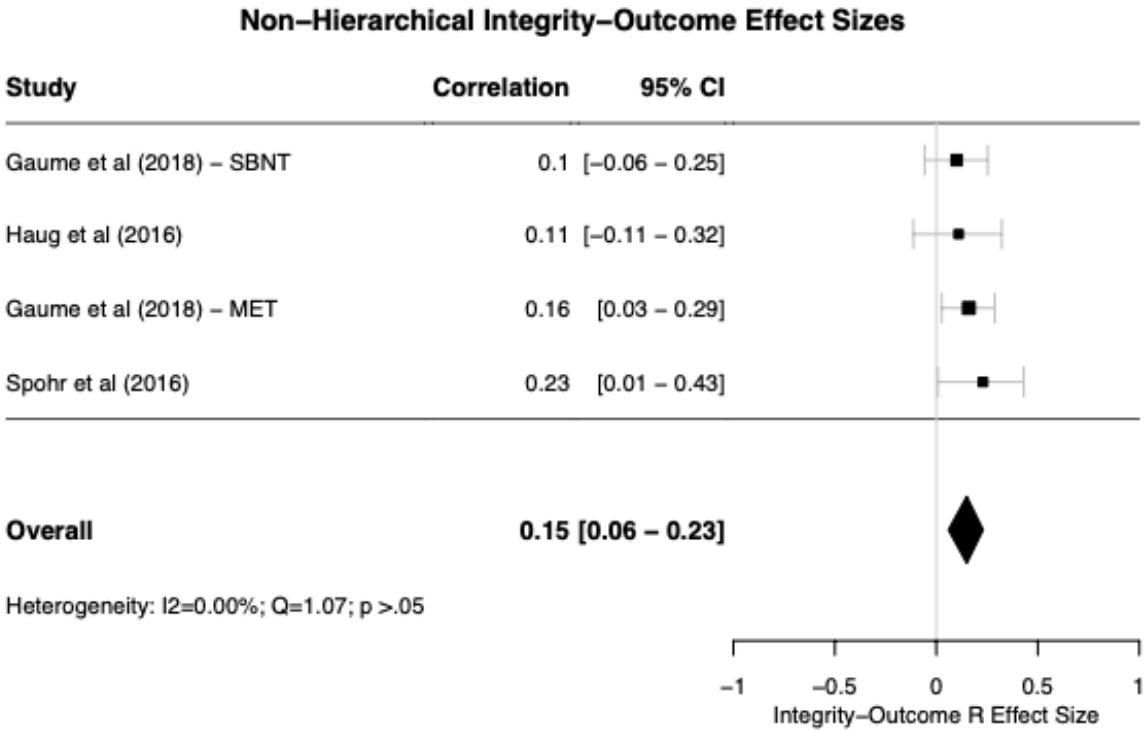

Figure 4

Forest Plots for Hierarchical Adherence-Outcome, Competence-Outcome and Integrity-Outcome Meta-Analyses