Psychotherapy Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tpsr20

Efficacy of group psychotherapy for anxiety disorders: A systematic review and meta-analysis

Sarah Barkowski, Dominique Schwartze, Bernhard Strauss, Gary M. Burlingame & Jenny Rosendahl

To cite this article: Sarah Barkowski, Dominique Schwartze, Bernhard Strauss, Gary M. Burlingame & Jenny Rosendahl (2020) Efficacy of group psychotherapy for anxiety disorders: A systematic review and meta-analysis, Psychotherapy Research, 30:8, 965-982, DOI: 10.1080/10503307.2020.1729440

To link to this article: https://doi.org/10.1080/10503307.2020.1729440

Psychotherapy Research, 2020
Vol. 30, No. 8, 965–982, https://doi.org/10.1080/10503307.2020.1729440

EMPIRICAL PAPER

Efficacy of group psychotherapy for anxiety disorders: A systematic review and meta-analysis

SARAH BARKOWSKI[1], DOMINIQUE SCHWARTZE [1], BERNHARD STRAUSS[1], GARY M. BURLINGAME [2], & JENNY ROSENDAHL[1]

[1]*Institute of Psychosocial Medicine and Psychotherapy, Friedrich-Schiller University, Jena University Hospital, Jena, Germany & [2]Department of Psychology, Brigham Young University, Provo, UT, USA*

(Received 4 September 2019; revised 30 January 2020; accepted 30 January 2020)

ABSTRACT

Objective: This meta-analysis evaluates the efficacy of group psychotherapy in the treatment of anxiety disorders.
Method: A comprehensive literature search using PubMed, PsychInfo, Web of Science, CENTRAL, and manual searches was conducted to locate randomized controlled trials. We found 57 eligible studies ($k = 76$ comparisons) including 3656 participants receiving group psychotherapy or an alternative treatment for generalized anxiety disorder, social anxiety disorder, and panic disorder.
Results: Effect size estimates show that group psychotherapy reduces specific symptoms of anxiety disorders more effectively than no-treatment control groups ($g = 0.92$, [0.81; 1.03], $k = 43$) and treatments providing common unspecific treatment factors ($g = 0.29$ [0.10; 0.48], $k = 12$). No significant differences were found compared to individual psychotherapy ($g = 0.24$ [−0.09; 0.57], $k = 7$) or pharmacotherapy ($g = -0.05$ [−0.33; 0.23], $k = 6$). The effects were unrelated to factors of the group treatment. Within head-to-head studies, a significant moderating effect emerged for researcher allegiance.
Conclusions: Our results support the efficacy of group psychotherapy for anxiety disorders. They indicate that mixed-diagnoses groups are equally effective as diagnosis-specific groups, although further evidence is required. Future primary studies should address differential effectiveness, include a wider range of therapeutic approaches as well as active comparison groups.

Keywords: group psychotherapy; anxiety; cognitive behavior therapy; outcome research; meta-analysis; researcher allegiance

Clinical or methodological significance of this article: This meta-analysis is the first to systematically examine direct comparisons of group psychotherapy for anxiety disorders to both alternative treatments and control groups. It thereby adds to information from previous meta-analyses that were either restricted to a specific diagnosis or did not regard group treatment in particular. The results provide evidence on the equivalence of group psychotherapy to other treatment options and can thus support its clinical use and recommendation within practice guidelines. By collecting a broad body of evidence, moderator analyses on group treatment factors that may influence the outcome could be examined.

1. Introduction

Anxiety disorders are among the most prevalent classes of mental disorders with a one-year prevalence between 8.3% and 11.6% in community samples (Baxter, Scott, Vos, & Whiteford, 2013; Kessler et al., 2009). Accordingly, the research interest has substantially increased in the last decades, and there is a growing awareness of the individual and societal burden of anxiety disorders, especially if left untreated (Whiteford et al., 2013). The high chronicity, early age at onset and a high comorbidity with other mental disorders, such as substance abuse and depression, have been particularly associated with substantial direct and indirect costs (Kessler & Greenberg, 2002; Lai, Cleary, Sitharthan, & Hunt, 2015). For the individual, anxiety disorders impede everyday role functioning (Kessler et al., 2009) and work performance (Ivandic et al., 2017; Kessler & Greenberg, 2002),

Correspondence concerning this article should be addressed to Jenny Rosendahl, Institute of Psychosocial Medicine and Psychotherapy, Friedrich-Schiller University, Jena University Hospital, Jena, Germany. Email: jenny.rosendahl@med.uni-jena.de

and lead to increased use of mental health care services (Kessler & Greenberg, 2002).

1.1. Empirically Supported Treatments for Anxiety Disorders

According to the NICE clinical guidelines, psychological interventions are the treatment of choice for anxiety disorders and should be preferred over pharmacological interventions as less intrusive (NICE, 2014). A number of well-evaluated diagnosis-specific psychotherapeutic treatment manuals are available. Cognitive–behavioral therapy (CBT) programs are dominant followed by interpersonal and psychodynamic psychotherapy (i.e. Leichsenring et al., 2013; Lipsitz et al., 2008). There are several classes of pharmacotherapies that have been shown to be effective for individuals who do not respond well to psychotherapy (Bandelow et al., 2015). Besides individual psychotherapy, the past three decades have seen an increase in randomized controlled trials (RCTs) testing manualized group treatments developed for specific mental disorders (Burlingame, Fuhriman, & Mosier, 2003; Burlingame, Strauss, & Joyce, 2013). The majority are based upon CBT approaches with two exceptions; psychodynamic (Knijnik, Kapczinski, Chachamovich, Margis, & Eizirik, 2004) and interpersonal group treatments (Huang & Liu, 2011) for social anxiety disorder (SAD). An increasing number of group protocols is designed for patients with different anxiety disorders treated in the same treatment group (Norton, 2012; Wolgensinger, 2015). These mixed diagnostic groups assume sufficient similarity in the underlying etiology of different anxiety disorders that they can be treated with a common intervention. For example, the DSM-V asserts that anxiety disorders "share features of excessive fear and anxiety and related behavioral disturbances" (APA, 2013) and anxiety treatments, regardless of the specific disorder, usually use common components, such as exposure and cognitive restructuring to manage avoidance behavior (Norton, 2008; Schmidt et al., 2012).

Taken together, group treatments are becoming a common evidence-based treatment for anxiety disorders. Especially, in-patient settings regularly offer groups to treat their patients (Barghaan, Schulz, Koch, & Watzke, 2009; Weber & Strauss, 2015). However, group interventions do not play a significant role in practice guidelines and are recommended in cases where individual therapy is unavailable (Bandelow et al., 2015; NICE, 2013, 2014). This may be due, in part, to a lack of systematic evidence.

1.2. Group Psychotherapy as a Treatment Option

Some parameters of the group treatment differ from an individual therapy format. Administering psychotherapy in a group supposedly has an economic advantage since several patients are treated together, resulting in a reduction of therapist time per patient. This is important since only a small percentage of primary care patients presenting with symptoms of anxiety disorders actually receive adequate psychosocial care (Stein et al., 2011; Weisberg, Beard, Moitra, Dyck, & Keller, 2014). Offering evidence-based group treatment might address this treatment shortage, especially if it is related to financial issues or the availability of treatment in less developed regions.

Moreover, the treatment is offered in an interpersonal environment that responds to the fact that psychosocial well-being often depends on feeling connected to the world and being respected and valued by others (Yalom & Leszcz, 2005). This treatment feature seems particular relevant to individuals suffering from anxiety disorders and is also reflected in the specific treatment factors that have been proposed as exclusive to groups, such as vicarious and interpersonal learning, experiencing universality, altruism and a sense of belonging and relatedness (see Fuhriman & Burlingame, 1990; Yalom & Leszcz, 2005). Empirical evaluation of group therapeutic factors is scarce, but there are a few studies that show an increase in the rating of group therapeutic factors across treatment duration as well as a relationship to treatment outcome for anxiety disorder patients (Behenck, Wesner, Finkler, & Heldt, 2017; Choi & Park, 2006; Taube-Schiff, Suvak, Antony, Bieling, & McCabe, 2007).

There are also disadvantages when patients with anxiety disorders are treated in groups. The group in itself might constitute an impediment to seeking treatment, as individuals, suffering from an anxiety disorder, might perceive the setting as challenging and therefore seek to avoid it. For example, Strauss, Spangenberg, Brähler, and Bormann (2015) found the attitude towards groups to be negatively related to anxiety. To address this disadvantage, group psychotherapists recommend preparatory sessions where concerns and motivational issues are addressed (e.g. Strauss & Mattke, 2012). A second disadvantage is the group setting provides, by definition, less opportunity to address individual topics and develop individualized etiological models. This can be a detriment to patient's perceiving themselves as cared for and appreciated as an individual (Shechtman & Kiezel, 2016). In light of these advantages and disadvantages, a careful analysis of the differential effectiveness of group and individual therapy can lead to evidence-based practice guidelines.

1.3. Literature Review

There are several meta-analyses of RCTs examining the efficacy of psychotherapeutic approaches including group psychotherapy in patients with anxiety disorders (for an overview see Supplementary Table S1). The most comprehensive review including studies on panic disorder (PD) with or without agoraphobia, generalized anxiety disorder (GAD), SAD (based on DSM criteria) was based on 234 randomized-controlled trials (94 psychotherapy, 110 pharmacotherapy, 28 combinations) published between 1983 and 2013 (Bandelow et al., 2015). Pre–post comparisons on disorder-specific symptoms revealed significant and large effects for individual CBT/exposure ($d = 1.30$, $k = 93$) and group CBT/exposure ($d = 1.22$, $k = 18$). Pre–post effect sizes for pharmacotherapy were somewhat larger ($d = 2.02$, $k = 206$). The authors further provided indirect effects of individual and group psychotherapy in comparison to either waitlist or to pill placebo. The effects compared to waitlist were similar for individual CBT/exposure ($d = 1.23$, $I^2 = 70.0\%$, $k = 25$) and group CBT/exposure ($d = 1.33$, $I^2 = 0\%$, $k = 7$), though a statistical test for the difference was not provided.

Compared to pill placebo, individual CBT revealed a significant effect ($d = 0.57$, $I^2 = 64.8\%$, $k = 9$) at a moderate level, whereas the pooled effect of group CBT was smaller and not significant ($d = 0.12$, $k = 5$). It is unfortunate that the most comprehensive review on the efficacy of psychotherapeutic treatment provided no direct comparison of group therapy to other therapies (i.e. individual psychotherapy or pharmacotherapy). Moreover, heterogeneity was substantial for all comparisons except for group CBT. Hence, results should be interpreted with caution.

Another review (Gould, Coulson, & Howard, 2012) summarized the evidence of CBT for anxiety disorders, focusing particularly on the treatment of older adults (55+). It is based on 12 trials published between 1996 and 2010 including PD, GAD, agoraphobia, phobia, and anxiety disorders not otherwise specified, but also on posttraumatic stress disorder (PTSD) and obsessive-compulsive disorder. CBT and group CBT were each compared against different active and non-active control groups such as waitlist, treatment as usual, pharmacotherapy, or supportive counselling/therapy. Meta-regression analyses revealed non-significant differences between individual and group CBT for both anxiety ($\beta = -0.20$; $p = .36$) and depression ($\beta = 0.04$; $p = .89$; 0 = individual, 1 = group). However, there was no differentiation between control groups and heterogeneity was not reported making interpretation difficult. Furthermore, the inclusion of older adults restricts generalizability. In a recent review on the placebo-controlled efficacy of CBT for anxiety disorders, Carpenter and colleagues (2018) compared effect sizes of individual and group therapy studies for patients of SAD and PTSD. They found a significant difference in favor of the individual CBT studies (Hedges' $g = 0.54$, 95% CI [0.40;0.68]; group CBT: Hedges' $g = 0.16$, 95% CI [0.01;0.31]). However, this finding is based upon indirect comparisons and limited to SAD and PTSD.

Beyond these comprehensive reviews, several meta-analyses of RCTs have been published on the efficacy of psychotherapy focusing on single anxiety disorders, i.e. GAD (Cuijpers et al., 2014; Hunot, Churchill, Silva de Lima, & Teixeira, 2007) and SAD (Acarturk, Cuijpers, van Straten, & de Graaf, 2009; Barkowski et al., 2016; Mayo-Wilson et al., 2014; Powers, Sigmarsson, & Emmelkamp, 2008; Wersebe, Sijbrandij, & Cuijpers, 2013). For PD with/without agoraphobia, two meta-analyses (Sánchez-Meca, Rosa-Alcazar, Marin-Martinez, & Gomez-Conesa, 2010; Trull, Nietzel, & Main, 1988) and one network meta-analysis (Pompoli et al., 2016) have been published. However, the two former meta-analyses included non-randomized trials and the network meta-analysis did not separate results for group therapy. A sensitivity analysis on individual therapy studies was performed which yielded results comparable to the full sample. One meta-analysis (Schwartze et al., 2017) focused on group therapy for PD and results produced a large effect size on the primary outcome compared to no-treatment control groups ($k = 9$; $g = 1.08$; 95% CI [0.82, 1.34]; $p < .001$). A second comparison with active treatment (e.g., individual therapy, pharmacotherapy and relaxation therapy) found non-significant differences on the primary outcome ($k = 6$; $g = 0.18$; 95% CI [−0.14, 0.49]; $p = .264$).

The interpretation of the efficacy of group psychotherapy is limited since there is a lack of direct comparisons of group psychotherapy against active treatment conditions such as individual psychotherapy or pharmacotherapy. Consequently, these effects might be biased by a host of confounds (e.g., characteristics of the patients, treatment, and therapists) that vary between studies. There are only two reviews that directly compare group therapy to other specific treatment approaches in the same study: (i) Mayo-Wilson et al. (2014) published a network meta-analysis including 101 RCTs with 15 comparisons examining individual psychotherapy and 28 comparisons group therapy, respectively, compared to WL, pill placebo, and psychological placebo. Estimated direct effects of group CBT revealed no significant differences in comparison to individual CBT ($d = -0.27$), and to pharmacotherapy

(MAOIs; $d = -0.09$), respectively. (ii) Barkowski et al. (2016) also provided results on direct comparisons, finding no differences between group psychotherapy and individual therapy ($g = 0.23$; $k = 4$) or pharmacotherapy ($g = -0.15$, $k = 4$), respectively. However, study inclusion in both reviews was limited to SAD exclusively.

Taken together, results on comparisons of group psychotherapy to control groups in the treatment of anxiety disorders are promising and effect sizes are similar to individual psychotherapy. However, the current evidence is limited by small sample sizes, unexplained heterogeneity and a lack of reporting on group therapy moderators. Data on the relative efficacy compared to other treatment approaches, are further restricted by the lack of direct within study comparisons. Finally, no meta-analysis that aggregates data across anxiety disorders focusses on questions of the group treatment.

1.4. Objective

Given the clinical and economic importance of groups in treating anxiety disorders and the limitations of past reviews, the purpose of the present study was to focus on group psychotherapy studies and include exclusively direct within study comparisons of group therapy to control groups and alternative treatments (e.g., individual psychotherapy and pharmacotherapy). We regard all anxiety disorders (SAD, PD with/without agoraphobia, GAD) and analyze them as a group following past meta-analyses that assume common features across anxiety disorders (Bandelow et al., 2015). This approach creates a broader pool of evidence, includes mixed diagnoses groups matching clinical practice and creates sufficient data for subgroup and moderator analyses testing factors that may influence treatment outcome. We consider the moderator analyses as explorative and included group structure (e.g., dose, size), leadership (single vs. co-led) and patient characteristics as well as study design and quality.

2. Method

Objectives, inclusion criteria, and methods were prespecified in a review protocol (Rosendahl, Barkowski, Schwartze, Tefikow, & Strauss, 2013).

2.1. Study Selection and Coding

2.1.1 Eligibility criteria. All RCTs from 1990 onwards were considered for inclusion, regardless of publication status and language (if English abstract was available). Studies were eligible if they fulfilled the following criteria:

(a) Population of adult patients with a diagnosis of anxiety disorder (i.e. SAD, PD, GAD) according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R, DSM-IV, DSM-IV-TR, DSM-V; American Psychiatric Association, 2013) or the International Classification of Diseases (ICD-10; World Health Organization, 2004).
(b) Provision of group psychotherapy based upon a psychotherapeutic formal change theory (e.g. cognitive behavioral, psychodynamic, interpersonal), homogeneously related to the disorder, performed by a professional therapist, administered in groups of at least three patients with a minimum of five sessions. Only interventions aiming at change of the primary psychopathology of anxiety disorders were considered.
(c) Eligible control groups were "no-treatment/wait-list control," "common factor control groups" (defined as "experimental condition used in an attempt to control for psychotherapy's common factors by providing a treatment devoid of specific ingredients," Baskin, Tierney, Minami, & Wampold, 2003, p. 975; since "treatment as usual" generally contains specific ingredients, it was not included here), individual psychotherapy, and pharmacotherapy. We placed minimal attention control groups with the no-treatment controls.
(d) Included trials reported on at least one of the following outcomes: (i) specific psychopathology, i.e., severity of disorder-specific symptoms, (ii) depression, and (iii) anxiety. Disorder-specific symptom severity was considered the primary outcome. Domains (ii) and (iii) summarize global ratings on general psychopathology; i.e. measures in the domain "anxiety" rate the general level of anxiety (e.g. BAI, HAMA) and not symptoms specific to the respective diagnosis. They are considered secondary, as they may not necessarily have been addressed by the treatment. Allocation of outcome measure to domain can be found in Supplementary Table S2.
(e) Patient assignment to treatment condition was random. Cluster randomization was eligible, as well as consecutive assignment if the first assignment was randomized.

2.1.2 Information sources and search. Within the scope of a comprehensive literature search, we searched the databases CENTRAL, Medline

(Ovid), PsychINFO and Web of Science according to a pre-specified search strategy designed to retrieve RCTs on the treatment of anxiety disorders in a group format. The final search strategy for Medline is displayed in the Supplementary Material S3. The database search was last updated on August 20th, 2015.

The database search was complemented by manual searches of the reference sections of recent reviews and meta-analyses published on the efficacy of psychological interventions for anxiety disorders and on the efficacy of group psychotherapy, respectively as well as by searching references of included primary studies. In order to identify unpublished studies, we used the ProQuest Dissertations and Theses Full Text Database and contacted the authors of relevant studies asking for information about unpublished data.

2.1.3 Study selection and data collection process. Titles and abstracts of database records were screened independently by two authors (SB, DS). Studies that were eligible according to the information in title and abstract were read in full text before the final decision on inclusion. Discussions in team meetings resolved unclear cases (with BS and JR). Two authors (DS, SB) extracted descriptive and statistical data from primary studies according to a pilot-tested coding book requiring the following information: (i) characteristics of participants (age, gender, type and method of diagnosis, comorbidity, medication use); (ii) characteristics of the intervention (theoretical orientation, duration and frequency of the group treatment, number of patients in a group); (iii) characteristics of the control group (type, duration and frequency of the control treatment); (iv) information on the outcome measure (type of measurement, length of follow up, study drop-out); (v) effect size related parameters. All outcome data from published measures were included into the analyses, regardless of self- or observer-report.

When available, we extracted outcome data for the pre-treatment time point and all assessment points after completion of the treatment, leading to post-treatment (first assessment after the last session and within one week of program termination), short-term follow-up ($\leq$ 6 months after the intervention), mid-term follow-up (6–18 months after the intervention) and long-term follow-up ($\geq$ 18 months after the intervention) effect size estimates.

In a pilot phase, five studies were double-coded independently by two authors (DS, SB). Inter-rater reliability for this coding was good to excellent, yielding a mean Cohen's Kappa of $\kappa = .88$ for categorical variables with a range between $\kappa = .66$ and $\kappa = 1.0$ and a mean single-score intra-class correlation coefficient of ICC = .98 for metric variables with a range between ICC = .67 and ICC = 1.0. Reasons for discrepancies were discussed and where necessary, the instructions in the coding manual were adapted.

To insure the validity of effect size parameters, outcomes were double-coded by JR. Raw population data (M, SD, N) were preferred over test results (p-value, F-value, etc.) and where possible pre-test data were used in the effect size calculation.

In case of missing information, we contacted the study authors. Out of six contacted authors, two responded and provided missing data for effect size calculation. If information on effect sizes could not be retrieved, we approximated data using different estimation methods, i.e., setting an effect size to zero if non-significant results were mentioned without reporting statistical parameters, or using conservative p-value estimations if exact p-values were not given (e.g. using $p = 0.05$ if $p < .05$ was given). If none of the coders was familiar with the publication language, a translator was consulted and the rating was done cooperatively. Since no translator for Farsi was available, three studies were excluded.

2.1.4 Risk of bias in individual studies. We used the Cochrane risk of bias tool to evaluate the validity of the included studies (Higgins, Altman, & Sterne, 2011a). The following sources of bias were coded: risk of selection bias (adequate randomization procedure and concealment of allocation), reporting bias (complete outcome reporting), detection bias (blinding of outcome assessors for observer-rated outcomes) and attrition bias (handling incomplete outcome data). The risk of performance bias (blinding of participants and personnel) was not assessed since by definition both personnel and patients are aware of the type of treatment. Studies were additionally rated on efforts to ensure treatment implementation according to protocol, either by providing specific training and regular supervision or by conducting adherence checks to a manual. The evaluation of this implementation bias has been proposed by the PPRISMA working group for quality rating in psychotherapy studies (Barth et al., 2011, personal communication) and was investigated within previous meta-analyses on psychological treatment (Tefikow et al., 2013). Some quality ratings were only applied to subgroups. The credibility of common factor control groups as a non-specific treatment factor was estimated either from credibility ratings or by using criteria, such as a positive rationale, discussion of the target problem and an equivalent treatment structure as a proxy (compare Baskin et al., 2003). Studies that compared two state-of-the-art treatments (i.e. comparisons to individual psychotherapy and pharmacotherapy, subsequently referred to as head-to-head) were rated on the researcher's allegiance to one of the approaches. Having developed a treatment or

stating hypotheses specifically favoring one of the treatments was regarded as indicating a potential risk of allegiance bias (compare Munder, Gerger, Trelle, & Barth, 2011). All studies with two treatment groups were assessed as to the equivalence of the therapist's training. According to Cochranés criteria of judging risk of bias, judgments were categorized as "low risk" or "high risk" of bias, respectively (Higgins et al., 2011a). If insufficient detail was reported to make an informed decision, studies were rated as "unclear risk" of bias and were added to the subgroup of studies with high risk of bias for analyses. Two separate independent coders (AL, JR) conducted the risk of bias assessment in duplicate.

2.2. Statistical Analyses

2.2.1 Summary measures. Between-group effect sizes Hedges' g were calculated for continuous outcome data using Comprehensive Meta-Analysis software (Borenstein, Hedges, Higgins, & Rothstein, 2013). We obtained effect sizes by subtracting the posttest mean of the comparison group from the posttest mean of the group treatment and dividing the result by the pooled standard deviation of the two groups. Whenever pre- and posttest data were available, the pooled standard deviation was taken from the standard deviations of the mean pre–post difference scores of the two groups. In cases where only posttest data were available, the standard deviations of the posttest means for the two groups were used. A small sample bias correction was applied to all effect sizes (Hedges & Olkin, 1985). If available, intention-to treat data were chosen over completer data. Positive effect sizes indicate a superiority of the group treatment, while negative effect sizes indicate effects in favor of the control treatment. The magnitude of Hedges' g was interpreted within the same framework as Cohen's d, regarding 0.20, 0.50, and 0.80 as small, medium, and large effect sizes, respectively (Cohen, 1992).

Since many studies reported multiple outcomes for one outcome domain, effect sizes were subsequently aggregated within domains for each unit of analysis using the agg-function provided by the "MAd" package version 0.8-1 (Del Re & Hoyt, 2014), implemented in the R statistical computing language and environment version 3.1.1 (R Core Team, 2014). Correlations between outcomes (Borenstein, Hedges, Higgins, & Rothstein, 2009; Hedges & Olkin, 1985) were accounted for and set at $r = .50$ (compare Wampold et al., 1997).

Outliers were identified by executing leave-one-out analyses under the exclusion of one study at a time. If the exclusion of a study resulted in the reduction of overall heterogeneity by more than 10%, the study was excluded for further analyses.

2.2.2 Synthesis of results. Meta-analyses were conducted in the R statistical computing language and environment, version 3.4.3 (R Core Team, 2017) for all subgroups and outcome categories, using the package "metafor," version 2.0-1 (Viechtbauer, 2010). Random-effects models (Hedges & Vevea, 1998) were applied with heterogeneity estimated using the DerSimonian-Laird method (DerSimonian & Laird, 1986). Between-study heterogeneity was assessed with χ^2 heterogeneity tests (Cochran's Q) and I^2 statistic (Higgins, Thompson, Deeks, & Altman, 2003). Values of 25%, 50% and 75% are commonly interpreted as low, moderate and high amounts of heterogeneity, respectively (Borenstein et al., 2009; Higgins et al., 2003).

We performed meta-analyses separately for comparisons to no-treatment control groups and for comparisons to active control groups. The latter were further stratified by type of comparison (common factor control group, individual psychotherapy, and pharmacotherapy; cf. Rosendahl et al., 2013). Power was calculated for this subgroup (Borenstein et al., 2009), since effect sizes as well as sample sizes were expected to be small.

Whenever a study had two control groups or two group treatments, these were entered to the analyses in separate comparisons. To overcome a unit-of-analysis error for multiple correlated comparisons within a study, the sample size of the shared experimental group was divided evenly by number of comparisons (Higgins, Deeks, & Altman, 2011b).

2.2.3 Risk of bias across studies. Publication bias was assessed via examination of the funnel plot by using the Egger test (Egger, Smith, Schneider, & Minder, 1997). If the Egger test suggested funnel plot asymmetry, a trim and fill analysis was used to obtain an adjusted estimate of the treatment effect by imputing potentially missing studies (Duval & Tweedie, 2000).

2.2.4 Moderator analyses. Moderator analyses were conducted for study characteristics (publication year), patient characteristics (type of anxiety disorder, recruitment strategy), characteristics of the group treatment (type of treatment, type of leadership, number of therapists per group, number of patients per group, treatment dose, length of session) and the risk of bias within studies. For comparisons to active control groups, the type of comparison group and the equivalence of overall treatment time and number of treatment sessions, as well as the specific risk of bias variables were considered additionally. We also intended to investigate the effect of an individual preparatory session before the group treatment as well as further group process-related factors, but

abandoned this objective due to insufficient reporting in primary studies.

We used the metafor package in R (Viechtbauer, 2010) to compute mixed-effects meta-regression analyses. Moderator variables were entered into the regression analyses as potential predictors of the effect size and investigated regarding their potential to explain heterogeneity. Categorical moderators were dummy-coded. Beta coefficients and a test of moderator significance on the Q-statistic are provided. Additionally, the amount of variance explained by the moderator is given as R^2. Dichotomous moderators were required to have at least $k = 5$ comparisons in each subgroup to ensure sufficient power for analyses.

2.2.5 Sensitivity analyses. In order to test the robustness of findings we conducted sensitivity analyses on the primary outcome measure with regard to (i) a more conservative effect size estimation and (ii) characteristics of the study design.

(a) With respect to a robust effect size estimation we restricted analyses to (i) effect sizes from standard self-report measures because they are more conservative (Cuijpers, Li, Hofmann, & Andersson, 2010a) and (ii) excluded all approximated effect sizes (i.e. estimated from insufficient data).
(b) More conservative study designs were examined by performing analyses on subsamples of studies, excluding (i) studies that reported an inadequate randomization procedure, (ii) studies that did not ensure adherence to the manual, (iii) studies that allowed for concurrent medication, (iv) non-CBT treatment approaches, (v) treatments provided without presence of a professional therapist.

2.2.6 Treatment acceptance. As a proxy to treatment acceptance, we conducted meta-analyses on the raw proportion of dropout across studies for each experimental group separately, deriving a standardized percentage of dropout. If given, we used the number of patients finishing treatment to calculate the dropout rate. Otherwise, the number of patients providing post-assessment was taken as an approximation. To compare treatment acceptance between conditions, we used meta-analyses on the relative risk (RR) of dropout, considering only direct within-study comparisons. RR higher one indicates higher dropout in the group treatment condition and below one a higher dropout in the comparison condition.

3. Results

3.1. Study Pool

We screened a total of 3831 records and finally included $N = 57$ RCTs reporting $k = 76$ comparisons in the meta-analysis. References of included studies can be found in Supplementary Material S4. Figure 1 contains the flow chart of the study selection process. 3.1.1 Study characteristics. Characteristics of included studies are depicted in Supplementary Table S5. Thirty-four studies (59.6%) reported on SAD patients, 13 (22.8%) on PD patients, five (8.8%) on GAD patients and five (8.8%) on mixed anxiety disorder diagnoses. Sixty-seven group psychotherapeutic interventions were reported, $n = 57$ of which followed a full CBT approach, $n = 6$ provided for exposure treatment alone and $n = 4$ for a different treatment approach ($n = 1$ cognitive therapy, $n = 1$ psychodynamic psychotherapy, $n = 1$ interpersonal psychotherapy and $n = 1$ social skills training). These interventions were directly compared to no-treatment controls ($k = 48$), common factor controls ($k = 12$), individual treatments ($k = 8$), and pharmacotherapy ($k = 8$). In total, 1922 patients received a group psychotherapeutic treatment and 1734 were allocated to a control group (for results by diagnosis on sample and treatment characteristics see Appendix, Table A1).

Gender was evenly distributed with a median of 57.5% female patients (interquartile range: $IQR = 19.1$; $N = 56$). The median age across studies was 35 years ($IQR = 6$, $N = 55$) and the typical group had six members ($IQR = 2$, $N = 52$) with 12 sessions of 120 min.

3.1.2 Risk of bias within studies. Figures A1 and A2 in the Appendix present the distribution of risk of bias across studies (assessment of single studies can be found in Supplementary Table S6). Nine studies were rated as low risk of selection bias (low risk of bias from random assignment and allocation concealment). Insufficient reporting was especially an issue in the rating of random assignment, allocation concealment and selected outcome reporting with over 60% of the studies rated as unclear.

3.2. Statistical Analyses

3.2.1 Group psychotherapy compared to no-treatment control groups.

Main results. Large and significant positive standardized effects were derived for all comparisons to no-treatment control groups on the primary treatment outcome, regardless of diagnosis. Results per subgroup are displayed in Table I. The aggregated mean effect for all studies was large, $g = 1.02$ (95%

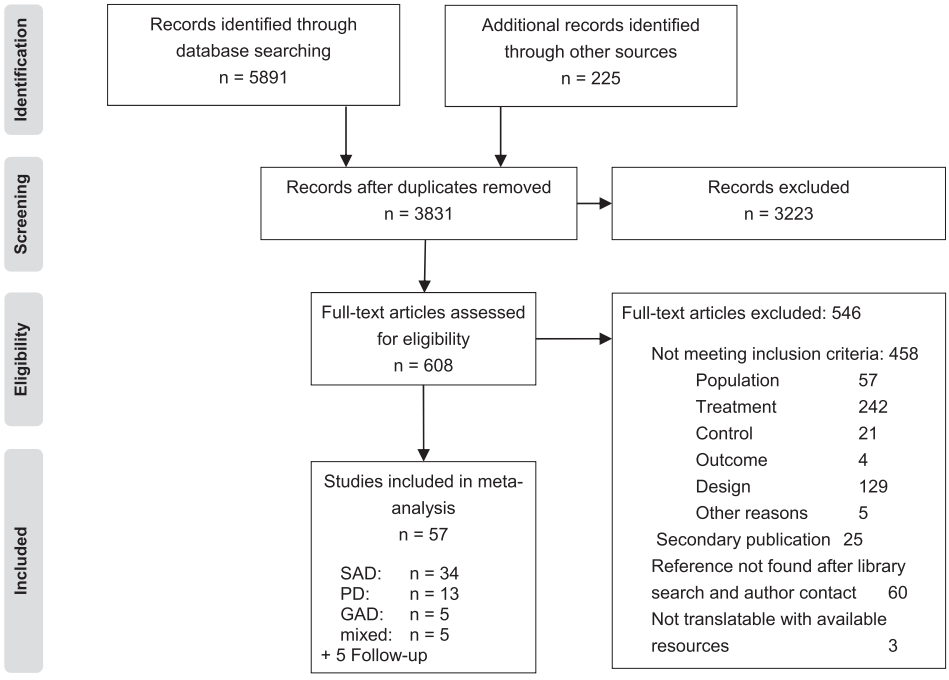

Figure 1. Flow diagram of study selection. GAD = generalized anxiety disorder; PD = panic disorder; SAD = social anxiety disorder.

CI [0.87; 1.18]; $p < .001$; $k = 45$), with significant between-study heterogeneity; $I^2 = 65.6\%$ ($Q = 127.92$; $df = 44$; $p < .001$). Leave-one-out analyses revealed two studies that reduced heterogeneity by over 10% if omitted (D'El Rey, Lacava, Cejkinski, & Mello, 2008: new $I^2 = 54.3\%$; Pishyar, Harris, & Menzies, 2008: new $I^2 = 54.9\%$). Both effect sizes were positive and unusually high for psychotherapy outcome research ($g = 4.49$ [3.29; 5.68]/$g = 3.83$ [2.84; 4.82]). Elimination of the two outliers reduced heterogeneity to a low level ($I^2 = 31.3\%$; $Q = 61.15$; $df = 42$; $p = .028$, $k = 43$) and produced a slightly smaller effect size of $g = 0.92$ (95% CI [0.81; 1.03]; $p < .001$; compare Figure 2). Egger's regression test for the whole sample was significant ($z = 4.38$; $p < .001$), but trim-and-fill analyses did not suggest the existence of omitted studies (for consultation of the funnel plot, see Appendix, Figure A3). Egger's regression test was non-significant without outliers ($z = 1.77$; $p = .076$) and trim-and-fill analyses did not suggest further studies.

Moderator analyses. Potential effect modifiers were analyzed in univariate analyses (compare Appendix Table A2). The type of anxiety diagnosis did not significantly moderate the effect size variance ($p = .196$; $R^2 = 12.6\%$) although three studies with mixed anxiety diagnoses produced a high effect size. None of the study quality indicators were significant moderators. There were no studies with a low reporting bias and too little variance within the blinding of

Table I. Results from meta-analyses on the primary treatment outcome.

	Comparisons to no-treatment control groups					Comparisons to active treatment groups				
				Heterogeneity					Heterogeneity	
Subgroups	*k*	*g*	CI(*g*)	*p*(Q)	I^2	*k*	*g*	CI(*g*)	*p*(Q)	I^2
Overall	45	1.02	0.87; 1.18	<.001	65.6	26	0.15	−0.04; 0.33	<.001	75.1
w/o outliers	43	0.92	0.81; 1.03	.028	31.3	25	0.20	0.05; 0.35	<.001	61.7
SAD	28	1.03	0.79; 1.26	<.001	72.6	16	0.08	−0.17; 0.33	<.001	78.0
w/o outliers	26	0.83	0.70; 0.96	.268	13.5	15	0.17	−0.02; 0.35	.003	57.0
PD	10	1.04	0.76; 1.31	.017	55.5	7	0.20	−0.09; 0.49	.016	61.6
GAD	4	0.94	0.64; 1.23	.197	35.9	2	0.05	−0.40; 0.50	.231	30.3
Mixed	3	1.23	0.94; 1.62	.841	0.0	1	0.93	0.51; 1.36	–	–

Note. *k* = Number of comparisons included in the analysis, *g* = Hedges' *g*, CI(*g*) = 95% confidence interval for *g*, *p*(Q) = level of significance for the Q-statistic, I^2 = proportion of heterogeneity not explained by within-study variance.

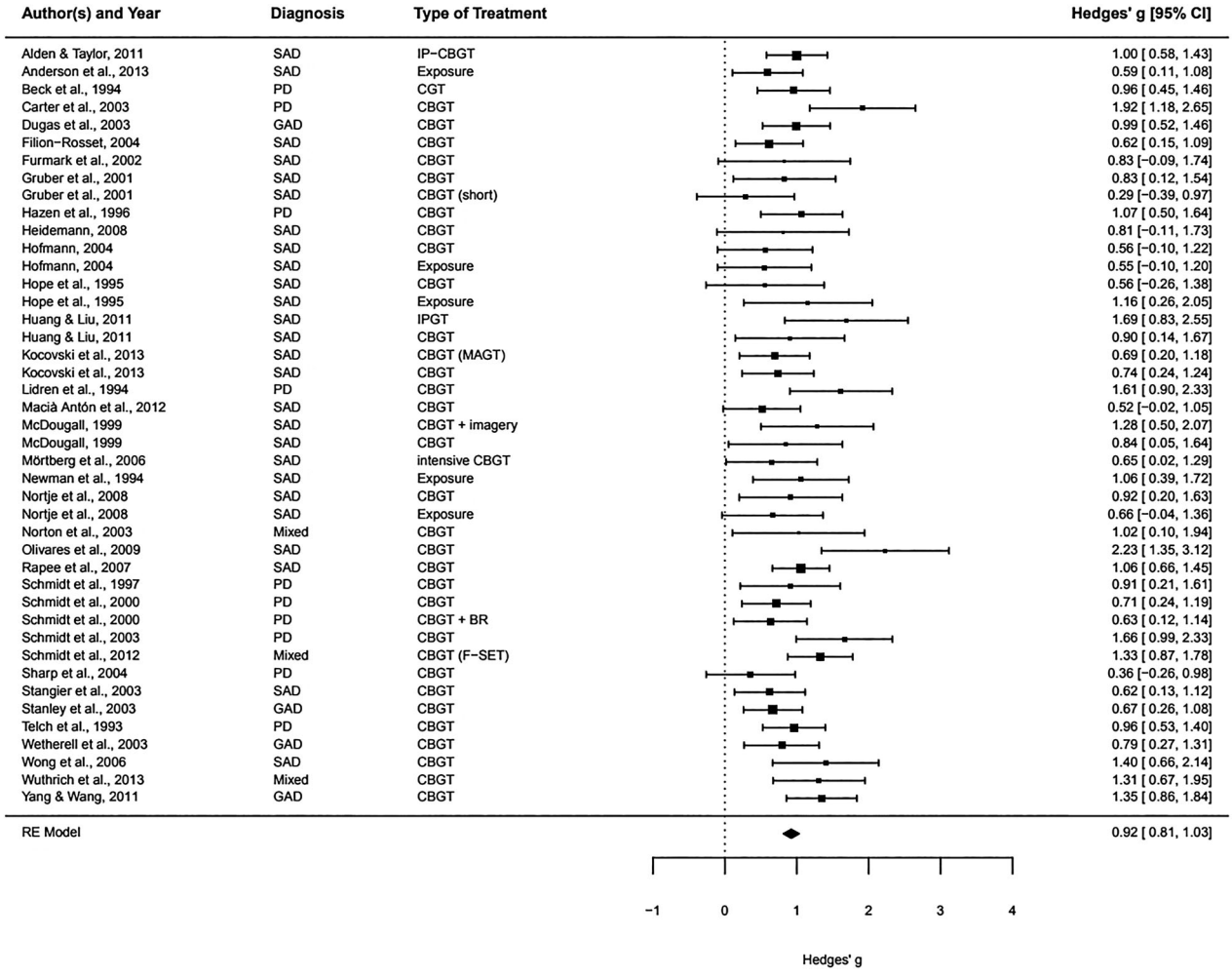

Figure 2. Meta-analytic results on the efficacy of group psychotherapy compared to no-treatment control groups on primary outcome symptom severity. GAD = generalized anxiety disorder; PD = panic disorder; SAD = social anxiety disorder; BR = breathing retraining; CBGT = cognitive-behavioral group therapy; CGT = cognitive group therapy; IP-CBGT = interpersonal cognitive-behavioral therapy; IPGT = interpersonal group therapy.

outcome assessment, therefore these quality indicators could not be analyzed. A higher effect size was found for studies when recruitment strategy was specified as "other" which involves whole population screening instead of referral or recruitment strategies and was typically associated with university students. However, the moderator did not obtain significance (compare Table A2). As indicated by the R^2-values, most heterogeneity was explained by treatment dose ($R^2 = 20.8$), with effect sizes rising as dose decreases (compare Figure 3). However, the moderator did not obtain significance either.

Sensitivity analyses. The results remained significant for data restricted to more robust effect estimators from standard self-report measures and non-approximated effect sizes and there was no indication that results differed on the conservative study inclusion criteria (compare Appendix Table A3). A significant controlled treatment effect emerged for

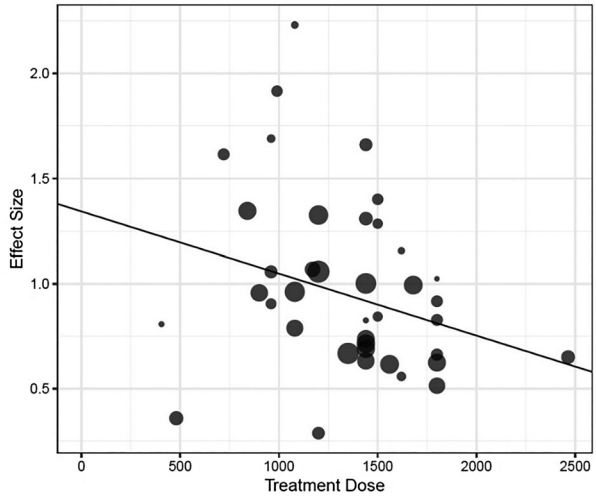

Figure 3. Scatterplot on treatment dose (in minutes) as a moderator of effect size for comparisons to no-treatment control groups.

all subgroups with effect sizes ranging between $g = 0.86$ and $g = 1.02$. Heterogeneity was medium to high and significant for all subgroups except for manualized treatments, professional therapist-led treatments and anxiety specific treatments. A small sample of eight studies prohibited the use of concurrent psychotropic medication and heterogeneity for this small sample was especially high ($I^2 = 66.2\%$), suggesting that the treatment effect might be moderated.

Results on secondary outcomes. Group psychotherapy showed positive effects on overall depression ($g = 0.77$ [0.59;0.96]; $p < .001$; $k = 23$) and anxiety ($g = 0.97$ [0.77;1.18]; $p < .001$; $k = 19$) with significant between-study heterogeneity at a moderate level for both outcome domains (depression: $I^2 = 42.4\%$; $Q = 38.20$; $df = 22$; $p = .017$; anxiety: $I^2 = 58.5\%$; $Q = 43.35$; $df = 18$; $p < .001$). Exclusion of one outlier (Pishyar et al., 2008; $g = 3.24$ [2.20;4.28]) eliminated all between-study heterogeneity for the depression outcomes ($g = 0.71$ [0.57;0.84], $p < .001$; $I^2 = 0.0\%$; $Q = 15.90$; $df = 21$; $p = .755$; $k = 22$) but there was no single study responsible for the heterogeneity in anxiety.

Results from follow-up analyses. A reduced number of studies reported short-term follow-up results and effects were maintained for all outcome domains with a specific symptomatology controlled effect size of $g = 0.96$ (95% CI [0.58; 1.33], $p < .001$, $k = 8$), depression of $g = 0.77$ (95% CI [0.46; 1.08], $p < .001$, $k = 6$), and anxiety of $g = 0.64$ (95% CI [0.23; 1.04], $p = .002$, $k = 3$). Heterogeneity was significant only for specific symptomatology ($I^2 = 63.5\%$, $Q = 19.18$, $df = 7$; $p = .008$), while it remained at $I^2 = 0.0\%$ both for depression ($Q = 3.98$, $df = 5$, $p = .552$) and anxiety ($Q = 0.21$, $df = 2$, $p = .900$). Leave-one-out analyses detected one study that was responsible for all remaining heterogeneity in the results regarding specific symptomatology (Olivares, Rosa-Alkázar, Olivares-Olivares, & Rosa-Alkázar, 2009; $g = 2.55$ [1.62; 3.48]). The magnitude of the effect was reduced for the remaining studies ($g = 0.76$ [0.53; 0.99], $p < .001$, $k = 7$, $I^2 = 0.0\%$, $Q = 5.71$, $df = 6$, $p = .456$). Only one study reported results for mid-term follow-up (Olivares et al., 2009), showing an unusually high effect size of $g = 2.99$ (95% CI [1.97; 4.00]). Therefore, this result should be interpreted with great caution. There were no long-term follow-up results reported.

3.2.2 Group psychotherapy compared to active treatment groups.

Main results. Group psychotherapy was compared to active treatment conditions yielding non-significant positive effect sizes on the primary outcome for all diagnostic subgroups (see Table I) except for mixed anxiety diagnoses, which consisted of only one comparison with a significant positive effect size of $g = 0.93$ (95% CI [0.51; 1.36], $p < .001$). Diagnostic subgroups were combined, resulting in a non-significant effect size of $g = 0.15$ (95% CI [−0.04; 0.33], $p = .123$, $k = 26$) with moderate to high between-study heterogeneity of $I^2 = 75.1\%$ ($Q = 100.37$, $df = 25$, $p < .001$). Power was calculated to be at 1-$\beta = .86$ and 1-$\beta = .66$ for an effect size of $g = .20$ and $g = .15$, respectively. Leave-one-out analyses identified one comparison as an outlier (Heimberg et al., 1998; comparison to pharmacotherapy) with a reduction of heterogeneity of over 10% after elimination, resulting in $I^2 = 61.7\%$ ($Q = 62.72$, $df = 24$, $p < .001$). The new effect size estimate gained significance with $g = 0.20$ (95% CI [0.05; 0.35], $p = .009$, $k = 25$; compare Figure 4). Power was estimated to be at 1-$\beta = .90$ and 1-$\beta = .72$ for an effect size of $g = 0.20$ and $g = 0.15$, respectively. Egger's regression test for funnel plot asymmetry for the whole study pool did not gain significance ($z = 1.24$, $p = .215$), however, visual examination of the funnel plot suggested omission of studies in the lower left corner and trim-and-fill analyses added seven potential studies ($g = -0.03$ [−0.23; 0.18], $p = .750$, $k = 33$, $Q = 167.85$, $df = 32$, $p < .001$, $I^2 = 80.9$; for consultation of the funnel plot, see Appendix, Figure A4). Egger's regression test for the reduced study pool did not gain significance either ($z = 1.65$, $p = .099$) but trim-and-fill analyses added six studies to the model ($g = 0.06$ [−0.11; 0.23], $p = .466$, $k = 31$, $Q = 111.22$, $df = 30$, $p < .001$, $I^2 = 73.0$), resulting in a non-significant overall effect size estimate.

Moderator analyses. Type of diagnosis did not moderate treatment effects ($p = .184$, $R^2 = 13.8\%$) although one diagnosis-mixed study produced an effect size that differed significantly from the intercept (compare Appendix Table A2 for moderator analyses). Type of comparison group did not moderate treatment effects ($p = .208$, $R^2 = 14.4\%$) but a subgroup analyses produced a significant effect in favor of group psychotherapy when compared to common factor control groups ($g = 0.29$ [0.10; 0.48], $p = .003$, $k = 12$, $I^2 = 56.7\%$, $Q = 25.39$, $df = 11$, $p = .008$). The two remaining subgroups were non-significant (compare Appendix Table A4), which may be related to the lower power to detect a small effect of $g = .20$ for both subgroups (individual PT: 1-$\beta = .45$; pharmacotherapy: 1-$\beta = .42$).

Due to insufficient sample size, a number of categorical moderator variables could not be entered into analyses (type of treatment, recruitment strategy, reporting bias, non-specific treatment factors (credibility), equivalent treatment time and treatment sessions) restricting our ability to explore differential treatment effects. Only one moderator was significant; allegiance to a particular treatment explained

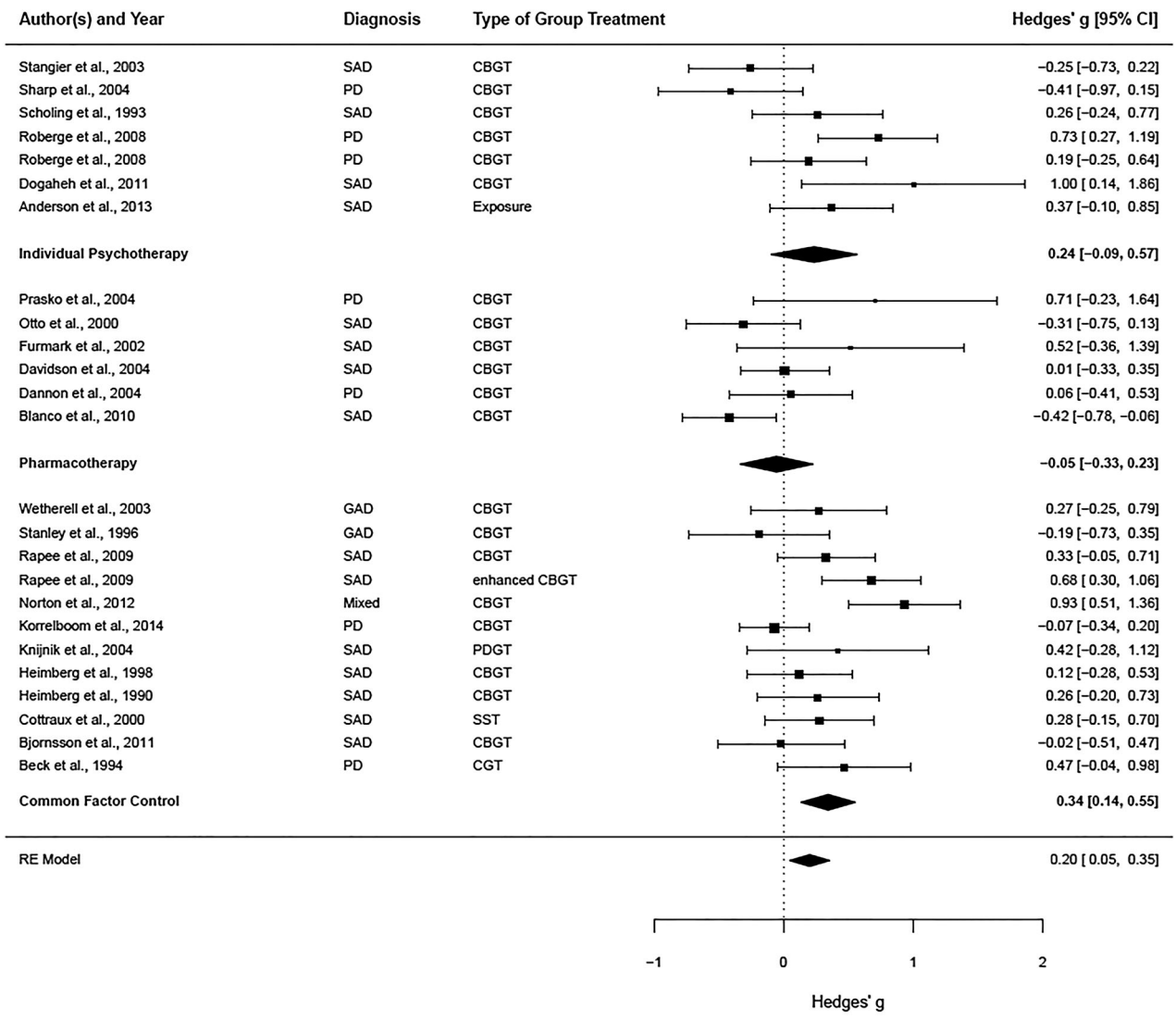

Figure 4. Meta-analytic results on the efficacy of group psychotherapy compared against active treatment conditions on primary outcome symptom severity. Note. Positive effect sizes indicate a higher effect for the group treatment compared to the comparison. GAD = generalized anxiety disorder; PD = panic disorder; SAD = social anxiety disorder; CBGT = cognitive-behavioral group therapy; CGT = cognitive group therapy; PDGT = psychodynamic group therapy; SST = social skills training.

effect size variability in head-to-head-studies ($k = 13$, $p = .034$, $R^2 = 31.1\%$); studies without allegiance had higher effect sizes on average. Among those with a treatment allegiance, two studies favored group treatment over pharmacotherapy and one study favored individual treatment over group treatment. Qualitative examination of the allegiance found effect sizes of $g = -0.31$ ($SE = 0.22$; Otto et al., 2000) and $g = -0.42$ ($SE = 0.18$; Blanco et al., 2010) for studies favoring group treatment over pharmacotherapy and an effect size of $g = 0.37$ ($SE = 0.24$; Anderson et al., 2013) for the study favoring individual treatment over group treatment. The effect sizes of studies rated as unclear were at $g = 0.06$ ($SE = 0.24$; Dannon, Gon-Usishkin, Gelbert, Lowengrub, & Grunhaus, 2004) and $g = -0.41$ ($SE = 0.29$; Sharp, Power, & Swanson, 2004).

Sensitivity analyses. (a) Analyses restricted to standard self-report measures and non-approximated effect sizes provided equivalent results to the whole sample (compare Appendix Table A3).

(b) Similarly, applying more conservative study inclusion criteria brought results on the level of small positive effects with a range between $g = 0.02$ and $g = 0.27$. For some subgroups these effects were non-significant (i.e. studies proscribing concurrent medication, studies requiring the presence of a professional therapist; compare Appendix Table A3).

Results on secondary outcomes. Effects on depression ($g = -0.06$ [−0.28; 0.15], $p = .559$, $k =$

15) and anxiety ($g = -0.42$ [−0.98; 0.14], $p = .139$, $k = 14$) did not differ significantly for group psychotherapy and active control groups. However, medium heterogeneity between studies was present for depression ($I^2 = 54.3\%$, $Q = 30.61$, $df = 14$, $p = .006$) and high heterogeneity for anxiety ($I^2 = 91.0\%$, $Q = 145.11$, $df = 13$, $p < .001$). Leave-one-out analyses identified results from the common factor control group comparison of the Heimberg et al. (1998) study as outlier for depression and both comparisons from the same study as outliers for anxiety. Excluding the respective comparisons led to an $I^2 = 0.0\%$ for depression ($Q = 9.73$, $df = 13$, $p = .716$) and an $I^2 = 65.5\%$ for anxiety ($Q = 31.84$, $df = 11$, $p < .001$). The new effect size estimates remained non-significant (depression: $g = 0.01$ [−0,14; 0.15], $p = .919$, $k = 14$; anxiety: $g = 0.12$ [−0.17; 0.40], $p = .435$, $k = 12$). Stratification for comparison group did not change the basic conclusions (compare Appendix Table A4).

3.2.3 Results on treatment acceptance. The dropout rate for group psychotherapy across all included studies was 16.2% (95% CI [13.2%; 19.2%], $p < .001$, $k = 63$) with a high heterogeneity between studies ($Q = 305.46$; $df = 62$; $p < .001$, $I^2 = 79.7\%$). Neither type of comparison group ($p = .248$), nor diagnostic subgroup ($p = .393$) explained heterogeneity.

We assessed within-study differences of dropout separately for the types of comparator. A significant between-group difference of dropout rates was found for no-treatment controlled comparisons, with fewer dropouts in the no-treatment groups (7.9% [5.2%; 10.6%]) compared to group psychotherapy (15.2% [11.4%; 19.1%]; $RR = 1.28$, 95% CI [1.03; 1.58]; $p = .025$; $k = 37$). Also, a trend emerged for higher dropout rates in group psychotherapy (25.1% [15.8%; 34.4%]) compared to individual psychotherapy (15.3% [10.8%; 19.9%]; $RR = 1.58$ [1.00; 2.49]; $p = .050$; $k = 7$). There was no significant difference between dropout rates of group psychotherapy (21.2% [14.0%; 28.4%]) and common factor control groups (18.7% [10.4%; 27.1%]; $RR = 0.91$ [0.67; 1.22]; $p = .520$, $k = 10$) or group psychotherapy (18.7% [10.0%, 27.4%]) and pharmacotherapy (25.5% [15.9%; 35.0%]; $RR = 0.76$ [0.55; 1.03]; $p = .081$, $k = 8$).

4. Discussion

The present review's aim was to systematically examine the efficacy of group psychotherapy for anxiety disorders with direct comparisons to control groups and alternative treatments. By including 57 studies, we were able to examine moderators better than past disorder-specific reviews. In comparison to no-treatment control groups we found large positive effects for group treatment regardless of patients' diagnosis on the primary outcome of symptoms. Heterogeneity was low suggesting that efficacy of the treatment does not vary considerably between studies. None of the moderators had a significant effect on treatment outcome. Results were similar for the secondary outcomes depression and anxiety and were maintained at follow-up for all outcomes. Therefore, the primary conclusion is that group psychotherapy is an efficacious treatment for anxiety disorders when it comes to primary as well as secondary outcomes.

Comparisons against active control groups and alternative treatments showed small effect sizes for the primary outcome favoring the group treatment with significance depending on study inclusion (with/out outliers and studies omitted due to publication bias). Stratifying the sample by comparison group revealed a superiority of group psychotherapy over common factor control groups and non-significant differences when compared to alternative treatments; i.e., individual psychotherapy and pharmacotherapy. A moderating effect for researcher allegiance was found in head-to-head studies. There were no significant differences between group and active controls on the secondary outcomes anxiety and depression. Thus, group is superior to no-treatment and produces equivalent results when compared to other active treatments. It reduces specific symptomatology better than placebo.

There are some restrictions to the quality of the evidence summarized in this review. Although, given our reliance on RCTs and direct comparisons, this sample of studies strengthens the evidence for group psychotherapy compared to previously available reviews, risk of selection bias could not be dismissed for a few of the studies. More specifically, two studies were judged as high risk of selection bias due to inadequate randomization procedures and 46 studies provided insufficient information to assess randomization or allocation concealment procedures. This leaves nine studies without risk of selection bias. Intent-to-treat analyses were performed in 22 studies and moderator analyses did not suggest an impact on treatment outcome. These limitations should be considered for interpretation. Sensitivity analyses and analyses of publication bias improved the robustness of our findings.

Our results on the overall efficacy of group treatments are comparable to two comprehensive reviews (Bandelow et al., 2015; Gould et al., 2012). In our review, we included 18 studies that have been considered in the meta-analysis of Bandelow et al. (2015) and three studies that have been

analyzed by Gould et al. (2012). In comparison to no-treatment control groups we found large between-group effect sizes similar to within-group (pre–post) effects reported by Bandelow et al. (2015). Moreover, our head-to-head comparisons of individual and group treatment revealed non-significant differences as found by Burlingame et al. (2016) and Gould et al. (2012), but contradicting findings from the study by Carpenter et al. (2018). Additionally, our study pool shows substantial overlap with disorder-specific reviews on the efficacy of group therapy (SAD: Acarturk et al., 2009; Barkowski et al., 2016; Mayo-Wilson et al., 2014; Powers et al., 2008; Wersebe et al., 2013; GAD: Cuijpers et al., 2014; Hunot et al., 2007; PD: Schwartze et al., 2017). Hence, our results are in line with the effects reported in previous reviews on similar research questions.

A strength of our meta-analysis is the examination of moderators of treatment effects that has failed in previous disorder-specific reviews due to a small number of studies. Although no fixed rules exist, a rule of thumb considers 10 studies per subgroup as sufficient to identify significant moderator effects with sufficient power. We were not able to meet this criterion but reported on moderators with at least five studies per subgroup. However, statistical power of meta-regression analyses depends not only on numbers of studies per subgroup but also on heterogeneity.

Recent qualitative reviews (Burlingame et al., 2013) and meta-analyses (Burlingame et al., 2016; Kösters, Burlingame, Nachtigall, & Strauss, 2006) as well as primary studies suggest various potentially influencing factors on the effects of group therapy. Among these are patient characteristics (e.g., attachment style, sense of cohesion), structural characteristics of the group therapy (e.g., individual preparatory sessions, booster sessions, group composition), and leader characteristics (e.g., co-leadership, therapeutic styles). Due to insufficient data, it was not possible to examine most of these factors, and we mainly investigated factors of the general structure of the group treatment. None of these (type of intervention, type of leadership, profession of leader, group size, length of session) were able to explain between-study heterogeneity.

We replicated findings in individual psychotherapy studies where there was no difference in outcome between study pools recruited by referral vs. recruitment (Cuijpers et al., 2013; Cuijpers, van Straten, Bohlmeijer, Hollon, & Andersson, 2010b). However, in our analyses, studies that did not follow either of these strategies yielded an effect size that differed significantly from the intercept.

There was a trend towards a moderating effect of treatment dose, with effect sizes rising with decreasing dose of treatment. This finding does not match the generally postulated negatively accelerating dose-effect relationship that has been found in studies on individual psychotherapy (Howard, Kopta, Krause, & Orlinsky, 1986). However, there are other patterns discussed within studies on dose-effect relationship, such as faster improvement rates for shorter treatments (Evans, Beck, & Burdett, 2017; Stulz, Lutz, Kopta, Minami, & Saunders, 2013). However, most studies focusing on the effect of treatment duration differ from our data, since they follow a design closer to everyday practice and leave the decision on therapy cessation open to therapist and patient. It is therefore difficult to compare these results to ours. We are aware of one meta-analysis of randomized-controlled trials that examined the effect of treatment dose across studies and found an influence of the intensity of treatment, with shorter but more frequent treatments yielding higher effect sizes (Cuijpers et al., 2013). In our analyses, the negative dose-effect relationship reached significance for a subgroup of treatments with exactly one session per week ($k = 31$, $\beta = -0.0007$, $R^2 = 60.75$, $p = .002$) which does not align with the intensity of treatment explanation. We coded treatment dose as total therapy time, which depends not only on the number of sessions but also on the length of each session, a variable that differed to some extent between included primary studies (compare Table A1) and could be related to our divergent finding. However, at the time, we cannot give a convincing rationale for the result. We could have come across a chance finding and more data are necessary to make an informed assessment.

In accordance with previous meta-analyses (for a review compare Munder, Brütsch, Leonhart, Gerger, & Barth, 2013), we found a moderating effect of researcher allegiance on study outcome for the subsample of head-to-head-studies. The standardized mean effect was lower for studies with a risk of allegiance bias, although we did not distinguish the direction of this allegiance. Other meta-analyses that examined group psychotherapy found an effect of allegiance, favoring the treatment that was supported by study authors (Burlingame et al., 2016; McRoberts, Burlingame, & Hoag, 1998).

Acceptance of group treatment differed considerably across studies as estimated by dropout rates. This suggests that there are certain characteristics of the group treatment or other characteristic of the study design (such as patient population) that may have an influence on how the treatment is tolerated. One factor that has been discussed to influence acceptance of the group format is the provision of a preparatory individual session. However, reporting of its application was scarce within the study pool

and impeded a quantitative analysis as possible moderator. A lower treatment acceptance in the group condition was found in comparisons to no-treatment control conditions. For the other pairwise comparisons, the difference was non-significant. A trend emerged towards fewer drop-outs in the individual therapy condition when compared to group therapy. This might indicate that patients tolerate individual psychotherapy better. However, it contradicts findings from a previous meta-analysis that did not find format differences (Swift & Greenberg, 2012), so caution is in order.

4.1. Limitations

One of our initial aims – to strengthen the evidence by focusing on direct comparisons – was achieved only in part. While all our analyses are based on direct comparisons, the study pools on comparisons to individual psychotherapy and pharmacotherapy were still limited, with seven and six comparisons, respectively. Thus, previously examined study pools were augmented by three and two studies, respectively, and head-to-head comparisons for GAD and mixed populations are still absent. So far, larger study pools with direct head-to-head comparisons were obtained by aggregating across disorders (e.g. Burlingame et al., 2016).

Furthermore, although we included a large sample of studies in our review and moderator analyses are based on larger samples than in most of the existing reviews, statistical power of moderator analyses might still pose a problem (López-López, Marín-Martínez, Sánchez-Meca, Van den Noortgate, & Viechtbauer, 2014). Hence, non-significant findings should not be considered as evidence for a non-existent effect (Borenstein et al., 2009; Hempel et al., 2013). In addition to low statistical power, associations found in moderator analyses are observational since they do not imply causal relationship and may be confounded by other factors due to co-linearity (Hedges & Pigott, 2004). Taken together, results of moderator analyses should be interpreted with caution.

Another limitation is dependency among observations and has been noted already in previous research (e.g., Burlingame et al., 2016; Schwartze et al., 2017). When patients are treated in groups, they share a common environment (e.g., having the same therapist, interacting with the same group members). The resulting dependency of data, when not properly accounted for in statistical tests, leads to a predictable inflation of type I error (Baldwin, Murray, & Shadish, 2005). Hence, ignoring data dependency in the primary studies might have resulted in false-positive effects.

As found in many systematic reviews, reporting was not sufficient across studies. For that reason, some of our pre-defined moderator analyses failed because of small samples. Missing information also impeded the risk of bias judgment, with a large percentage of studies assessed as unclear risk of bias. Limitations exist in our search and data collection strategies. Some studies could not be retrieved despite comprehensive library search and efforts to contact the authors. Although inclusion was not restricted to English language, it was limited to languages authors were familiar with. Therefore, a limited number of studies was not included and data were further limited by some study authors not responding to our requests.

It is likely that further eligible studies have been published since the completion of our systematic literature search. These would affect conclusions especially if treatment design or findings differed systematically from those in the present study pool. Recent publications on treatment for SAD (Goldin et al., 2016; Kocovski, Fleming, Hawley, Ringo Ho, & Antony, 2015; Ștefan, Căpraru, & Szilágyi, 2018) show a trend towards the application of mindfulness-based group programs. In direct comparison with CBT groups, these proved to be equally efficacious in reducing disorder-related symptoms (Goldin et al., 2016; Kocovski et al., 2015). In accordance with our results, they also produced high effect sizes around $d = 1.0$ in comparison to wait list control groups in the reduction of anxiety symptoms (Goldin et al., 2016; Ștefan et al., 2018). One study investigating mixed psychiatric disorders including PD (Sundquist, Palmer, Johansson, & Sundquist, 2017), but not exclusively anxiety disorders, compared individually based CBT (considered as TAU) with mindfulness-based group therapy in a Swedish sample. Reduction of psychiatric symptoms did not differ between the two treatments.

The external validity of our results is limited since CBT was the dominant treatment and we had few studies with mixed anxiety diagnoses. Treatment setting was predominantly outpatient and primarily manualized treatments have been evaluated which may not reflect clinical practice.

Finally, publication bias was suggested by funnel plot asymmetry for no-treatment controlled studies; however, statistical tests did not yield any evidence for biased effect estimations.

4.2. Conclusions

Based on the results of our meta-analysis we could conclude that group treatment is highly efficacious

in comparison to no-treatment. Effect sizes are comparable to those found for individual psychotherapy and pharmacotherapy. Group treatment, mostly following a CBT approach, is more efficacious than common factor control in reducing specific symptomatology, indicating the value of specific therapeutic techniques beyond common factors such as therapeutic alliance, attention, instillation of hope and provision of a treatment rational. Since direct comparisons against individual psychotherapy and pharmacotherapy revealed only small, non-significant differences, group therapy could be considered as a treatment alternative for anxiety disorders. This latter finding has relevance to the role group psychotherapy might play within practice guidelines and eventually within health care systems. The evidence herein that supports format equivalence replicates the findings of a meta-analysis that used a much larger sample size (Burlingame et al., 2016). Collectively, these findings support a practitioner freely choosing between formats, depending on patient needs, schedule and availability of treatment resources. However, our data are limited to clinical efficacy and do not allow conclusions on cost effectiveness and feasibility of group psychotherapy within current health care systems. While group offers a greater therapist/patient ratio, group therapy adds other resource demands including space, therapist pre- and post-group preparation time (e.g., case notes) as well as the organizational infrastructure (e.g., group coordinators, multiple therapist referrals, etc.) it takes to have a vibrant group therapy program. Still, the cost-effectiveness of group treatment and its ability to treat more patients at the same time directly address access challenges that are affecting most clinical practice settings. Our results, while preliminary, suggest equal efficacy for mixed diagnoses groups. This finding has direct clinical implications for outpatient treatment settings. More specifically, mixed diagnoses groups are more easily created in a timely manner since a wider range of patients are eligible.

There is the unresolved challenge of statistical heterogeneity indicating that not all patients profit equally from group therapy. Although we analyzed various moderator variables by subgroup analyses and meta-regression analyses, heterogeneity of study results was not explained completely. A trend appeared for a negative dose-efficacy relationship with treatments with a lower dosage yielding higher effect sizes. Since this has not been replicated yet, it could be a chance finding and should be further examined in future meta-analyses. Most clinically relevant moderators could not be tested adequately and future studies should focus on investigating the impact of patient characteristics, treatment factors, and patient-treatment interactions on the efficacy of group therapy. Transparent and complete reporting of information regarding these variables is an essential prerequisite for future studies to allow meta-analysts considering the entire pool of studies for subgroup analyses. Moreover, adequately addressing dependency of observations in statistical tests should be a claim for future studies.

Acknowledgements

We kindly thank Rahel Klatte and Anna Lindner for their support in retrieving primary studies and in the risk of bias coding.

Funding

This research has been funded by the Bundesministerium für Bildung und Forschung [German Ministry of Education and Research; grant number 01KG1216].

Supplemental data

Supplemental data for this article (Appendix and Supplementary Material) can be accessed here. [https://doi.org/10.1080/10503307.2020.1729440].

ORCID

Dominique Schwartze http://orcid.org/0000-0001-5559-4705
Gary M. Burlingame http://orcid.org/0000-0002-8275-4118

References

References marked with an asterisk were included in the meta-analysis.

Acarturk, C., Cuijpers, P., van Straten, A., & de Graaf, R. (2009). Psychological treatment of social anxiety disorder: A meta-analysis. *Psychological Medicine*, *39*(2), 241–254. doi:10.1017/s0033291708003590

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing. doi:10.1176/appi.books.9780890425596.dsm05

*Anderson, P. L., Price, M., Edwards, S. M., Obasaju, M. A., Schmertz, S. K., Zimand, E., & Calamaras, M. R. (2013). Virtual reality exposure therapy for social anxiety disorder: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *81*(5), 751–760. doi: 10.1037/a0033559

Baldwin, S. A., Murray, D. M., & Shadish, W. R. (2005). Empirically supported treatments or type I errors? Problems with the analysis of data from group-administered treatments.

Journal of Consulting and Clinical Psychology, *73*(5), 924–935. doi: 10.1037/0022-006X.73.5.924

Bandelow, B., Reitt, M., Rover, C., Michaelis, S., Gorlich, Y., & Wedekind, D. (2015). Efficacy of treatments for anxiety disorders: A meta-analysis. *International Clinical Psychopharmacology*, *30*(4), 183–192. doi: 10.1097/yic.0000000000000078

Barghaan, D., Schulz, H., Koch, U., & Watzke, B. (2009). Versorgungsstrukturen im stationären setting in Deutschland: Verteilung von Einzel- und Gruppentherapie und deren psychotherapeutischen Ausrichtungen. *Gruppenpsychotherapie und Gruppendynamik*, *45*(2), 83–103. doi:10.13109/grup.2009.45.2.83

Barkowski, S., Schwartze, D., Strauss, B., Burlingame, G. M., Barth, J., & Rosendahl, J. (2016). Efficacy of group psychotherapy for social anxiety disorder: A meta-analysis of randomized-controlled trials. *Journal of Anxiety Disorders*, *39*, 44–64. doi: 10.1016/j.janxdis.2016.02.005

Barth, J., Gerger, H., Munder, T., Rabung, S., Rosendahl, J., & Tefikow, S. (2011). *Reporting meta-analyses of psychotherapeutic effectiveness: PPRISMA*. In 42nd international meeting of the society for psychotherapy research. Bern.

Baskin, T. W., Tierney, S. C., Minami, T., & Wampold, B. E. (2003). Establishing specificity in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology*, *71*(6), 973–979. doi: 10.1037/0022-006X.71.6.973

Baxter, A. J., Scott, K. M., Vos, T., & Whiteford, H. A. (2013). Global prevalence of anxiety disorders: A systematic review and meta-regression. *Psychological Medicine*, *43*(5), 897–910. doi: 10.1017/S003329171200147X

Behenck, A., Wesner, A. C., Finkler, D., & Heldt, E. (2017). Contribution of group therapeutic factors to the outcome of cognitive-behavioral therapy for patients with panic disorder. *Archives of Psychiatric Nursing*, *31*(2), 142–146. doi: 10.1016/j.apnu.2016.09.001

*Blanco, C., Heimberg, R. G., Schneier, F. R., Fresco, D. M., Chen, H., Turk, C. L., … Liebowitz, M. R. (2010). A placebo-controlled trial of phenelzine, cognitive behavioral group therapy, and their combination for social anxiety disorder. *Archives of General Psychiatry*, *67*(3), 286–295. doi:10.1001/archgenpsychiatry.2010.11

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York, NY: Wiley.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2013). *Comprehensive meta-analysis version 2*. Englewood, NJ: Biostat. Inc.

Burlingame, G. M., Fuhriman, A., & Mosier, J. (2003). The differential effectiveness of group psychotherapy: A meta-analytic perspective. *Group Dynamics: Theory, Research, and Practice*, *7*(1), 3–12. doi:10.1037/1089-2699.7.1.3

Burlingame, G. M., Seebeck, J. D., Janis, R. A., Whitcomb, K. E., Barkowski, S., Rosendahl, J., & Strauss, B. (2016). Outcome differences between individual and group formats when identical and nonidentical treatments, patients, and doses are compared: A 25-year meta-analytic perspective. *Psychotherapy (Chic)*, *53*(4), 446–461. doi: 10.1037/pst0000090

Burlingame, G., Strauss, B., & Joyce, A. (2013). Change mechanisms and effectiveness of small group treatments. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 640–689). New York, NY: Wiley & Sons.

Carpenter, J. K., Andrews, L. A., Witcraft, S. M., Powers, M. B., Smits, J. A. J., & Hofmann, S. G. (2018). Cognitive behavioral therapy for anxiety and related disorders: A meta-analysis of randomized placebo-controlled trials. *Depression and Anxiety*, *35*(6), 502–514. doi:10.1002/da.22728

Choi, Y. H., & Park, K. H. (2006). Therapeutic factors of cognitive behavioral group treatment for socical phobia. *Journal of Korean Medical Science*, *21*(2), 333–336. doi:10.3346/jkms.2006.21.2.333

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

Cuijpers, P., Li, J., Hofmann, S. G., & Andersson, G. (2010a). Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: A meta-analysis. *Clinical Psychology Review*, *30*(6), 768–778. doi:10.1016/j.cpr.2010.06.001

Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F. (2013). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: A meta-analysis of direct comparisons. *World Psychiatry*, *12*(2), 137–148. doi:10.1002/wps.20038

Cuijpers, P., Sijbrandij, M., Koole, S., Huibers, M., Berking, M., & Andersson, G. (2014). Psychological treatment of generalized anxiety disorder: A meta-analysis. *Clinical Psychology Review*, *34*(2), 130–140. doi:10.1016/j.cpr.2014.01.002

Cuijpers, P., van Straten, A., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010b). The effects of psychotherapy for adult depression are overestimated: A meta-analysis of study quality and effect size. *Psychological Medicine*, *40*(2), 211–223. doi:10.1017/S0033291709006114

Dannon, P. N., Gon-Usishkin, M., Gelbert, A., Lowengrub, K., & Grunhaus, L. (2004). Cognitive behavioral group therapy in panic disorder patients: The efficacy of CBGT versus drug treatment. *Annals of Clinical Psychiatry*, *16*(1), 41–46. doi:10.1080/10401230490281609

Del Re, A. C., & Hoyt, W. T. (2014). *MAd: Meta-Analysis with Mean Differences. R package version 0.8-2*. Available from http://cran.r-project.org/web/packages/MAd

*D'El Rey, G. J. F., Lacava, J. P. L., Cejkinski, A., & Mello, S. L. (2008). Cognitive-behavioral group treatment in social phobia: 12-week outcome. *Revista de Psiquiatria Clinica*, *35*(2), 79–83. Retrieved from http://www.scielo.br/pdf/rpc/v35n2/a06v35n2.pdf

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*(3), 177–188. doi:10.1016/0197-2456(86)90046-2

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method for testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463. doi:10.1111/j.0006-341X.2000.00455.x

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

Evans, L. J., Beck, A., & Burdett, M. (2017). The effect of length, duration, and intensity of psychological therapy on CORE global distress scores. *Psychology and Psychotherapy*, *90*(3), 389–400. doi:10.1111/papt.12120

Fuhriman, A., & Burlingame, G. M. (1990). Consistency of matter: A comparative analysis of individual and group process variables. *The Counseling Psychologist*, *18*(1), 6–63. doi:10.1177/0011000090181002

Goldin, P. R., Morrison, A., Jazaieri, H., Brozovich, F., Heimberg, R., & Gross, J. J. (2016). Group CBT versus MBSR for social anxiety disorder: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *84*(5), 427–437. doi: 10.1037/ccp0000092

Gould, R. L., Coulson, M. C., & Howard, R. J. (2012). Efficacy of cognitive behavioral therapy for anxiety disorders in older people: A meta-analysis and meta-regression of randomized controlled trials. *Journal of the American Geriatrics Society*, *60*(2), 218–229. doi:10.1111/j.1532-5415.2011.03824.x

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, *9*(4), 426–445. doi: 10.1037/1082-989X.9.4.426

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, *3*(4), 486–504. doi: 10.1037/1082-989X.3.4.486

*Heimberg, R. G., Liebowitz, M. R., Hope, D. A., Schneier, F. R., Holt, C. S., Welkowitz, L. A., … Klein, D. F. (1998). Cognitive behavioral group therapy vs phenelzine therapy for social phobia: 12-week outcome. *Archives of General Psychiatry*, *55* (12), 1133–1141. doi:10.1001/archpsyc.55.12.1133

Hempel, S., Miles, J. N., Booth, M. J., Wang, Z., Morton, S. C., & Shekelle, P. G. (2013). Risk of bias: A simulation study of power to detect study-level moderator effects in meta-analysis. *Systematic Reviews*, *2*, 107. doi:10.1186/2046-4053-2-107

Higgins, J. P. T., Altman, D. G., & Sterne, J. A. C. (2011a). Chapter 8: Assessing risk of bias in included studies. In J. P. T. Higgins, & S. Green (Eds.), *Cochrane Handbook for systematic reviews of interventions version 5.1.0* (updated March 2011). The Cochrane Collaboration. Retrieved from www.cochrane-handbook.org

Higgins, J. P. T., Deeks, J. J., & Altman, D. G. (2011b). Special topics in statistics. In J. P. T. Higgins, & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions. Version 5.1.0* (updated March 2011). The Cochrane Collaboration. Retrieved from www.cochrane-handbook.org

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

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist*, *41*(2), 159–164. doi:10.1037/0003-066X.41.2.159

*Huang, H.-L., & Liu, X.-M. (2011). Effects of group interpersonal psychotherapy and group cognitive behavioral therapy on social anxiety in college students. *Chinese Mental Health Journal*, *25*(5), 324–327.

Hunot, V., Churchill, R., Silva de Lima, M., & Teixeira, V. (2007). Psychological therapies for generalised anxiety disorder. *Cochrane Database of Systematic Reviews*, *1*, Art. No.: CD001848, doi:10.1002/14651858.cd001848

Ivandic, I., Kamenov, K., Rojas, D., Ceron, G., Nowak, D., & Sabariego, C. (2017). Determinants of work performance in workers with depression and anxiety: A cross-sectional study. *International Journal of Environmental Research and Public Health*, *14*(5), doi:10.3390/ijerph14050466

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Chatterji, S., Lee, S., Ormel, J., … Wang, P. S. (2009). The global burden of mental disorders: An update from the WHO World Mental Health (WMH) surveys. *Epidemiologia e psichiatria sociale*, *18* (1), 23–33. doi:10.1017/s1121189 (00001421

Kessler, R. C., & Greenberg, P. E. (2002). The economic burden of anxiety and stress disorders. In D. C. K. L. Davis, J. T. Coyle, & C. Nemeroff (Eds.), *Neuropsychopharmacology: The fifth generation of progress* (pp. 981–992). Philadelphia: Lippencott, Williams and Wilkins.

*Knijnik, D. Z., Kapczinski, F., Chachamovich, E., Margis, R., & Eizirik, C. L. (2004). Psychodynamic group treatment for generalized social phobia. *Revista Brasileira de Psiquiatria*, *26*(2), 77–81. doi:10.1590/S1516-44462004000200003

Kocovski, N. L., Fleming, J. E., Hawley, L. L., Ringo Ho, M.-H., & Antony, M. M. (2015). Mindfulness and acceptance-based group therapy and traditional cognitive behavioral group therapy for social anxiety disorder: Mechanisms of change. *Behaviour Research and Therapy*, *70*, 11–22. doi:10.1016/j.brat.2015.04.005

Kösters, M., Burlingame, G., Nachtigall, C., & Strauss, B. (2006). A meta-analytic review of the effectiveness of inpatient group psychotherapy. *Group Dynamics: Theory, Research, and Practice*, *10*(2), 146–163. doi:10.1037/1089-2699.10.2.146

Lai, H. M. X., Cleary, M., Sitharthan, T., & Hunt, G. E. (2015). Prevalence of comorbid substance use, anxiety and mood disorders in epidemiological surveys, 1990–2014: A systematic review and meta-analysis. *Drug and Alcohol Dependence*, *154*, 1–13. doi:10.1016/j.drugalcdep.2015.05.031

Leichsenring, F., Salzer, S., Beutel, M. E., Herpertz, S., Hiller, W., Hoyer, J., … Leibing, E. (2013). Psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder: A multicenter randomized controlled trial. *American Journal of Psychiatry*, *170* (7), 759–767. doi:10.1176/appi.ajp.2013.12081125

Lipsitz, J. D., Gur, M., Vermes, D., Petkova, E., Cheng, J., Miller, N., … Fyer, A. J. (2008). A randomized trial of interpersonal therapy versus supportive therapy for social anxiety disorder. *Depression & Anxiety*, *25*(6), 542–553. doi:10.1002/da.20364

López-López, J. A., Marín-Martínez, F., Sánchez-Meca, J., Van den Noortgate, W., & Viechtbauer, W. (2014). Estimation of the predictive power of the model in mixed-effects meta-regression: A simulation study. *British Journal of Mathematical and Statistical Psychology*, *67*(1), 30–48. doi: 10.1111/bmsp.12002

Mayo-Wilson, E., Dias, S., Mavranezouli, I., Kew, K., Clark, D. M., Ades, A. E., & Pilling, S. (2014). Psychological and pharmacological interventions for social anxiety disorder in adults: A systematic review and network meta-analysis. *The Lancet. Psychiatry*, *1*(5), 368–376. doi: 10.1016/s2215-0366(14)70329-3

McRoberts, C., Burlingame, G. M., & Hoag, M. J. (1998). Comparative efficacy of individual and group psychotherapy: A meta-analytic perspective. *Group Dynamics: Theory, Research, and Practice*, *2*(2), 101–117. doi:10.1037/1089-2699.2.2.101

Munder, T., Brütsch, O., Leonhart, R., Gerger, H., & Barth, J. (2013). Researcher allegiance in psychotherapy outcome research: An overview of reviews. *Clinical Psychology Review*, *33*(4), 501–511. doi: 10.1016/j.cpr.2013.02.002

Munder, T., Gerger, H., Trelle, S., & Barth, J. (2011). Testing the allegiance bias hypothesis: A meta-analysis. *Psychotherapy Research*, *21*(6), 670–684. doi: 10.1080/10503307.2011.602752

National Institute for Health and Care Excellence (NICE). (2013). *Social anxiety disorder: Recognition, assessment and treatment*. Retrieved from https://www.nice.org.uk/guidance/cg159/chapter/1-Recommendations

National Institute for Health and Care Excellence (NICE). (2014). *Anxiety disorders*. Retrieved from https://www.nice.org.uk/guidance/QS53/chapter/Quality-statement-2-Psychological-interventions

Norton, P. J. (2008). An open trial of a transdiagnostic cognitive-behavioral group therapy for anxiety disorder. *Behavior Therapy*, *39*(3), 242–250. doi: 10.1016/j.beth.2007.08.002

*Norton, P. J. (2012). A randomized clinical trial of transdiagnostic cognitive-behavioral treatments for anxiety disorder by comparison to relaxation training. *Behavior Therapy*, *43*(3), 506–517. doi:10.1016/j.beth.2010.08.011

*Olivares, J., Rosa-Alkázar, A. I., Olivares-Olivares, P. J., & Rosa-Alkázar, Á. (2009). Treatment of young adults with generalized social phobia. *International Journal of Hispanic Psychology*, *1*(2), 1–14.

*Otto, M. W., Pollack, M. H., Gould, R. A., Worthington, J. J., III, McArdle, E. T., & Rosenbaum, J. F. (2000). A comparison of the efficacy of clonazepam and cognitive-behavioral group

therapy for the treatment of social phobia. *Journal of Anxiety Disorders*, *14*(4), 345–358. doi:10.1016/s0887-6185(00)00027-x

*Pishyar, R., Harris, L. M., & Menzies, R. G. (2008). Responsiveness of measures of attentional bias to clinical change in social phobia. *Cognition & Emotion*, *22*(7), 1209–1227. doi: 10.1080/02699930701686008

Pompoli, A., Furukawa, T. A., Imai, H., Tajika, A., Efthimiou, O., & Salanti, G. (2016). Psychological therapies for panic disorder with or without agoraphobia in adults: A network meta-analysis. *Cochrane Database of Systematic Reviews*, *4*), doi:10.1002/14651858.CD011004.pub2

Powers, M. B., Sigmarsson, S. R., & Emmelkamp, P. M. G. (2008). A meta-analytic review of psychological treatments for social anxiety disorder. *International Journal of Cognitive Therapy*, *1*(2), 94–113. doi:10.1680/ijct.2008.1.2.94

R Core Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.R-project.org/

R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.R-project.org/

Rosendahl, J., Barkowski, S., Schwartze, D., Tefikow, S., & Strauss, B. (2013). Systematic review and meta-analyses on the efficacy of small group treatment for mental disorders. PROSPERO: CRD42013004419. Available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013004419

Sánchez-Meca, J., Rosa-Alcazar, A. I., Marin-Martinez, F., & Gomez-Conesa, A. (2010). Psychological treatment of panic disorder with or without agoraphobia: A meta-analysis. *Clinical Psychology Review*, *30*(1), 37–50. doi: 10.1016/j.cpr.2009.08.011

*Schmidt, N. B., Buckner, J. D., Pusser, A., Woolaway-Bickel, K., Preston, J. L., & Norr, A. (2012). Randomized controlled trial of false safety behavior elimination therapy: A unified cognitive behavioral treatment for anxiety psychopathology. *Behavior Therapy*, *43*(3), 518–532. doi:10.1016/j.beth.2012.02.004

Schwartze, D., Barkowski, S., Strauss, B., Burlingame, G. M., Barth, J., & Rosendahl, J. (2017). Efficacy of group psychotherapy for panic disorder: Meta-analysis of randomized, controlled trials. *Group Dynamics: Theory, Research, and Practice*, *21*(2), 77–93. doi:10.1037/gdn0000064

*Sharp, D. M., Power, K. G., & Swanson, V. (2004). A comparison of the efficacy and acceptability of group versus individual cognitive behaviour therapy in the treatment of panic disorder and agoraphobia in primary care. *Clinical Psychology & Psychotherapy*, *11*(2), 73–82. doi:10.1002/cpp.393

Shechtman, Z., & Kiezel, A. (2016). Why do people prefer individual therapy over group therapy? *International Journal of Group Psychotherapy*, *66*(4), 571–591. doi:10.1080/00207284.2016.1180042

Ștefan, C. A., Căpraru, C., & Szilágyi, M. (2018). Investigating effects and mechanisms of a mindfulness-based stress reduction intervention in a sample of college students at risk for social anxiety. *Mindfulness*, *9*(5), 1509–1521. doi:10.1007/s12671-018-0899-y

Stein, M. B., Roy-Byrne, P. P., Craske, M. G., Campbell-Sills, L., Lang, A. J., Golinelli, D., … Sherbourne, C. D. (2011). Quality of and patient satisfaction with primary health care for anxiety disorders. *Journal of Clinical Psychiatry*, *72*(7), 970–976. doi:10.4088/JCP.09m05626blu

Strauss, B., & Mattke, D. (eds.). (2012). *Gruppenpsychotherapie – Lehrbuch für die Praxis*. Stuttgart: Klett Cotta.

Strauss, B., Spangenberg, L., Brähler, E., & Bormann, B. (2015). Attitudes towards (psychotherapy) groups. Results of a survey in a representative sample. *International Journal of Group Psychotherapy*, *65*, 411–430. doi:10.1521/ijgp_2014_64_001

Stulz, N., Lutz, W., Kopta, S. M., Minami, T., & Saunders, S. M. (2013). Dose–effect relationship in routine outpatient psychotherapy: Does treatment duration matter? *Journal of Counseling Psychology*, *60*(4), 593–600. doi: 10.1037/a0033589

Sundquist, J., Palmer, K., Johansson, L. M., & Sundquist, K. (2017). The effect of mindfulness group therapy on a broad range of psychiatric symptoms: A randomised controlled trial in primary health care. *European Psychiatry*, *43*, 19–27. doi:10.1016/j.eurpsy.2017.01.328

Swift, J. K., & Greenberg, R. P. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, *80*(4), 547–559. doi:10.1037/a0028226

Taube-Schiff, M., Suvak, M. K., Antony, M. M., Bieling, P. J., & McCabe, R. E. (2007). Group cohesion in cognitive-behavioral group therapy for social phobia. *Behaviour Research and Therapy*, *45*(4), 687–698. doi: 10.1016/j.brat.2006.06.004

Tefikow, S., Barth, J., Maichrowitz, S., Beelmann, A., Strauss, B., & Rosendahl, J. (2013). Efficacy of hypnosis in adults undergoing surgery or medical procedures: A meta-analysis of randomized controlled trials. *Clinical Psychology Review*, *33*(5), 623–636. doi: 10.1016/j.cpr.2013.03.005

Trull, T. J., Nietzel, M. T., & Main, A. (1988). The use of meta-analysis to assess the clinical significance of behavior therapy for agoraphobia. *Behavior Therapy*, *19*(4), 527–538. doi:10.1016/S0005-7894(88)80021-2

Viechtbauer, W. (2010). metafor: Meta-Analysis Package for R. R package version 1.4-0, Available from http://CRAN.R-project.org/package=metafor

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, 'all must have prizes.'. *Psychological Bulletin*, *122*(3), 203–215. doi: 10.1192/pb.bp.109.025643

Weber, R., & Strauss, B. (2015). Group psychotherapy in Germany. *International Journal of Group Psychotherapy*, *65*(4), 513–525. doi: 10.1521/ijgp.2015.65.4.513

Weisberg, R. B., Beard, C., Moitra, E., Dyck, I., & Keller, M. B. (2014). Adequacy of treatment received by primary care patients with anxiety disorders. *Depression and Anxiety*, *31*(5), 443–450. doi: 10.1002/da.22209

Wersebe, H., Sijbrandij, M., & Cuijpers, P. (2013). Psychological group-treatments of social anxiety disorder: A meta-analysis. *Plos One*, *8*, 11. doi: 10.1371/journal.pone.0079034

Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., … Vos, T. (2013). Global burden of disease attributable to mental and substance use disorders: Findings from the global burden of disease study 2010. *The Lancet*, *382*(9904), 1575–1586. doi:10.1016/S0140-6736(13)61611-6

Wolgensinger, L. (2015). Cognitive behavioral group therapy for anxiety disorders. *Dialogues in Clinical Neuroscience*, *17*(3), 347–351. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4610619/

World Health Organization. (2004). ICD-10*: International statistical classification of diseases and related health problems:* Tenth revision (2nd ed.). World Health Organization. Retrieved from https://apps.who.int/iris/handle/10665/42980

Yalom, I. D., & Leszcz, M. (Collaborator) (2005). *The theory and practice of group psychotherapy* (5th ed.). New York, NY: Basic Books.