DOCUMENT RESUME

ED 393 054 CG 026 894

AUTHOR Burlingame, Gary M.; And Others
TITLE Group Psychotherapy Efficacy: A Meta-Analytic Perspective.
PUB DATE Aug 95
NOTE 52p.; Paper presented at the Annual Meeting of the American Psychological Association (103rd, New York, NY, August 11-15, 1995).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Comparative Analysis; *Counseling Effectiveness; Effect Size; *Group Counseling; Group Dynamics; Group Therapy; *Meta Analysis; *Psychotherapy

ABSTRACT

Analyses of the efficacy of group psychotherapy indicate that group therapy demonstrates, in a majority of reviews, significant improvement over inert comparison groups and proves comparable or superior to other active treatment conditions. Because group therapy is a viable cost-efficient treatment option being used with increasing regularity among diverse populations, and with varying structural formats, a central task demanding attention is careful analysis of the differential effectiveness of group therapy across treatment variations or dimensions. Using a common measuring standard called an effect size, meta analyses can represent the average amount of change one could expect in the average client who receives a given treatment. The present study sought to explore systematically the relationships between improvement rates in group psychotherapy (effect size) and several treatment, therapist, client, and methodological variables using meta-analytic techniques. The review was based on the cumulative results of 12 years of group psychotherapy outcome literature. Results were reported on: component characteristics; post treatment change comparisons--treatment type comparison with wait-list controls, group treatment type comparison, outcome source comparison on post-treatment change, and outcome content comparison on post-treatment change; and pre- post-treatment change comparisons--client population/diagnosis client gender, composition, and attrition. Contains 47 references. (JBJ)

Group Psychotherapy Efficacy: A Meta-Analytic Perspective

Gary M. Burlingame

Addie Fuhriman

Edward Anderson

Brigham Young University

Paper presented at the 103rd Annual Convention of the American Psychological Association, New York City, New York, August, 1995.

Group Psychotherapy Efficacy: A Meta-Analytic Perspective

The efficacy of group psychotherapy has been addressed in numerous evaluations over the last three decades. A recent analysis of such reviews (Fuhriman & Burlingame, 1994) indicates that group therapy demonstrates, in a majority of the reviews, significant improvement over inert comparison groups and proves comparable or superior to other active treatment conditions. These results are timely and especially important in the current era of health care reform in which the efficacy and cost of mental health services are paramount and are a central issue. Because group therapy is a viable cost-efficient treatment option being used with increasing regularity, among diverse populations, and with varying structural formats, a central task demanding our attention is a careful analysis of the differential effectiveness of group therapy across treatment variations or dimensions, such as duration, setting, patient diagnosis, theoretical orientation, and other treatment features.

Although both narrative and meta-analytic studies have formed the basis of the efficacy studies of group psychotherapy over the past few decades, more recent times find meta-analytic techniques being used with greater regularity. Such analyses provide qualitative summary statements about treatments of interest based on a large number of studies that address a particular theme or subject matter (e.g., substance abuse, depression, schizophrenia). The meta-analytic approach quantifies the effectiveness of a particular form of treatment by using a common measuring standard called an effect size—an index that represents the average amount of change one could expect in the average client who receives a given treatment.

Although meta-analytic techniques have been applied to a wide variety of client populations and treatment approaches, very few have focused on group therapy as the primary treatment of interest. More typically, the estimate of the effectiveness of group psychotherapy appears to be derived as a post hoc comparison from studies investigating psychotherapy in general or a specific patient population. As a result, few analyses focus on variables that have the capacity to explain the differential effectiveness of group psychotherapy. Nonetheless, the extant meta-analyses that address group psychotherapy provide the context for the present, extensive meta-analytic study, the purpose of which is to explore factors that will lend understanding to the differential effectiveness of group treatment.

To date, fourteen published meta-analytic studies speak to the effectiveness of group psychotherapy—twelve provide comparative estimates of the effectiveness of group psychotherapy; two give estimates of group effectiveness with specific populations (Table 1). Two-thirds of these studies give a single comparative analysis of the effects of group versus individual therapy, concluding that no reliable difference exists between these two modalities (Baer & Nietzel, 1991; Casey & Berman, 1985; Miller & Berman, 1983; Robinson, Berman, & Neimeyer, 1990; Shapiro & Shapiro, 1982a; Smith, Glass, & Miller, 1980; Tillitski, 1990; Weisz, Weiss, Alicke, & Klotz, 1987). The first examination (based on 475 studies) of differential effectiveness of group versus individual therapy found an effect size of .87 for individual therapy and .83 for group therapy (Smith, et al., 1980). Covering a five-year period (1975-1979), Shapiro and Shapiro (1982b) examined outcome studies that compared

two or more treatments to a control group and found an overall effect size of 1.12 for individual and .89 for group.

Other studies reporting no differences between these two modalities include Robinson's et al. (1990) study of treatments for depression, Tillitski's (1990) comparison of group, individual, and control treatments, and Miller and Berman's (1983) analysis of cognitive behavioral treatments. Tillitski (1990) found that larger effect sizes were reported in more recent studies than those in earlier ones. In addition, he found group treatment to be less effective for children and more effective with adolescents when compared with individual treatment. Other meta-analyses that investigate treatment effectiveness for children and adolescents found no difference between group and individual therapy, specifically, studies from 1952 to 1983 (Casey and Berman, 1985) and studies from 1970-1985 (Weisz, et al., 1987). Finally, Baer and Neitzel (1991) found no difference between group (ES=.79) and individual therapy (ES=.44) when comparing individual cognitive and behavioral treatment of impulsivity in children.

Two meta-analyses reported individual therapy to be more effective with adult populations than group psychotherapy (Table 2). In an investigation of self-statement modification therapies, Dush, Hirt, and Schroeder (1983) found individual therapy (ES=.93) to be more effective than group therapy (ES=.58). As well, Nietzel, Russell, Hummings, and Gretter (1987), in exploring the effectiveness of cognitive behavioral approaches in the treatment of unipolar depression, indicated that individuals treated in groups have significantly more depressive symptoms at the end of treatment than did those in individual treatment. Additionally, others found individual therapy (ES=1.31) to be more effective

than group therapy (ES=.10) in impacting language proficiencies in children (Russell, Greenwald, & Shirk, 1991). However, Fuhriman and Burlingame (1994) point out that in at least two of these meta-analyses (Dush, et al., 1983; Nietzel, et al., 1987), the included studies appear to be using the group as a "convenient" format to deliver predetermined treatment interventions, rather than to utilize the unique properties deemed therapeutic to the group therapy format. As a result, these authors suggest that the findings from these two meta-analyses are limited in addressing the comparative efficacy of the two approaches. In contrast to these findings, Grossman & Hughes' review (1992) found group therapy (ES=.80) to be more effective than individual therapy (ES=.40) in addressing internalizing disorders in children.

Two studies have applied the meta-analytic procedures to individual populations of interest, wherein they provide estimates of group treatment effectiveness with specific clinical populations. Fettes & Peters' study (1992) investigated group treatments for bulimia and reported an effect size estimate of .75 (based on pre-post comparisons). Larger effect sizes were associated with more hours of therapy per week as well as with the addition of other treatment components (e.g., medication, individual treatment, family treatment). In a study of group treatment for older depressed patients, Gorey & Crynes (1991) found an effect size estimate of .68 for the older depressed group (when compared to a wait-list control group), but attributed 87% of this improvement to client variables (e.g., group member living alone, severity of depression).

In summary, of the existing fourteen meta-analyses, only three have group psychotherapy as the principal focus of the investigation (Fettes & Peters, 1992; Gorey &

Crynes, 1991; Tillitski, 1990). Unfortunately, these three studies are limited by sample size (e.g., Tillitski used nine studies) and generalizability (e.g., applicable to eating disorders and older, depressed patients). To date, no large scale meta-analysis has investigated variables that might account for, or illuminate the differential effectiveness of group psychotherapy with an adult clinical population. In other words, what might explain differences in client improvement in group treatment (e.g., theoretical model, setting, diagnosis) has been virtually unexplained by meta-analytic procedures.

Thus, the present study sought to explore systematically the relationships between improvement rates in group psychotherapy (effect size) and several treatment (e.g., orientation, setting, dosage, group size, composition), therapist (experience), client (e.g., diagnosis, chronicity of disorder, gender, age) and methodological (e.g., random assignment, attrition) variables. The intent was to focus on those studies in which group psychotherapy was a primary treatment modality, and to explore variables that might explain differential improvement in group treatment.

Method

The present meta-analytic review is based on the cumulative results of twelve years of group psychotherapy outcome literature. Included studies were identified in a computer search of the *Psychlit* database for the years 1980 through 1992 and by reviewing the reference sections of recent relevant group therapy meta-analyses, narrative reviews, and outcome studies (e.g., Bergin & Garfield, 1994; Dush, Hirt, & Schroeder, 1983; Forsyth, 1990; Fuhriman and Burlingame, 1994; Miller & Berman, 1983; Nietzel, et al., 1987; Robinson, Berman, & Neimeyer, 1990; Shapiro & Shapiro, 1982a, 1982b; Smith, Glass, &

Miller, 1980; Tillitski, 1990; Toseland & Siporin, 1986; Yalom, 1985), resulting in 2025 studies identified.

Following the acquisition and reading of the abstracts of the published studies, studies were included in the meta-analysis if they met the following criteria (Figure 1): a) investigate group therapy and have a control or comparison group within the same study; (b) groups meet regularly with an identified therapist and for a particular purpose; (c) clients exhibit an identifiable problem representative of clients treated within a group format; (d) study is experimental or quasi-experimental; (f) outcome results stated such that effect size estimates can be calculated; (g) reports written in English. In addition, studies which utilized children or adolescent subjects were excluded as therapeutic procedures for these populations often differ from those of adult populations (Dagley, Gazda, Eppinger & Stewart, 1994). Studies involving groups that were primarily psychoeducational in form (e.g., classroom instruction) or those utilizing treatments that did not have a prominent verbal component (e.g., bibliotherapy, exercise) were also excluded because of their lack of client interaction. Marital and family therapy studies were also omitted because the goals of such therapies appear significantly different from the individualized goals being investigated in the majority of the group studies. One hundred and sixteen studies met the criteria for the meta-analysis and formed the final sample.

Selected articles meeting the inclusion criteria were coded on a variety of client, leadership, treatment, and methodological variables which past research and theory indicate may correlate with effect size or have significant correlations with effect size, as demonstrated in previous meta-analyses. Client variables included: diagnosis, chronicity,

entrance into therapy, concurrent therapies, inpatient or outpatient status, client gender, and age. Characteristics of therapist leadership consisted of: theoretical orientation, experience and training, number of leaders, and treatment focus. Methodological variables included: source and content of the outcome measures utilized within a study, indication of random assignment, treatment integrity, sample size, indication and percent of attrition (Bergin & Garfield, 1994; Burlingame, Kircher, & Taylor, 1994; Dush, Hirt, & Schroeder, 1983; Forsyth, 1990; Fuhriman and Burlingame, 1994b; Miller & Berman, 1983; Nietzel, Russell, et al., 1987; Robinson, Berman, & Neimeyer, 1990; Shapiro & Shapiro, 1982a, 1982b; Smith, Glass, & Miller, 1980; Tillitski, 1990; Toseland & Siporin, 1986; Yalom, 1985).

Given the presumption that the process of group therapy is thought to be qualitatively different from that of individual therapy, it follows that to evaluate group therapy solely on the basis of variables traditionally used to evaluate individual therapy may be problematic. To remedy this problem, several attributes thought to correlate with outcome in group therapy were also coded in an effort to test for an association between these factors and effect size. Selected group features included: presence of pregroup training, size, composition (heterogenous, homogenous), therapy dosage, and setting.

In addition, a four-level categorization was used to discriminate the various levels of structure inherent in the therapy offered (Figure 2) (Burlingame and Fuhriman, 1995). Key reviewers indicate that the level of structure found in group treatment not only is related to outcome but also characterizes the type of group treatment being offered (Burlingame & Fuhriman, 1994; Dies, 1994). Thus, the categories were designed to discriminate whether the therapy process was guided by a therapist/manual or by the group members and whether

the interaction was topic or member interaction focused. Fuhriman and Burlingame (1990, 1994) suggest that the elements of interaction and group focus are key characteristics of traditional group psychotherapy and hypothesize that these may be differentially related to the effectiveness of group treatment for specific clientele and disorders.

Following a semester-long coding training program, variable coding was conducted by clinical psychology doctoral and advanced undergraduate psychology students. Coders rated in teams of two—both raters on a team independently rated each article and when differences in their ratings arose they were required to come to consensus on their ratings. Average Kappa values for the independent ratings, prior to consensus coding, were adequate (.81), with an average rate of agreement of 83 percent across all categories.

Meta-analytic Calculation

Outcome statistics for each measure reported in a study were expressed as standardized effect size estimates according to the meta-analysis formula (Cohen, 1977):

$$d = \frac{M1 - M2}{Sp}$$

Utilizing this formula, effect sizes were calculated in two ways. In the first, d represents the estimated standardized effect size; M1 and M2, the means of the groups being compared; and Sp, the pooled within groups standard deviation. This procedure allows for comparison of the relative effectiveness of two groups within the same study. For example, when a wait-list control group was investigated in the same study with a treatment group, M1 represented the mean of the group treatment and M2 the mean of the wait-list control group. Thus, a direct comparison is made between the outcome of the group therapy and control conditions.

When calculated in this manner an effect size of 1.00 indicates that the M1 group achieved an effect one standard deviation above the effect obtained by the M2 group. Thus, the average person in the group therapy condition achieved a better outcome than 84% of the people in the control group. This formula was utilized to calculate effect sizes that directly compare group therapy to control treatments as well as effect sizes that compare the effect of different group treatments utilized within the same study.

The second method for calculating effect sizes treats the posttest mean for a group within a study as M1 and the pretest mean as M2. By calculating effect sizes in this way, an effect size is obtained that represents the absolute standardized treatment effect for that particular group. Thus, an effect size of 1.0 indicates that the group members, on average, improved one standard deviation from their pretest scores. Likewise, an effect size of -1.0 indicates that the clients did not improve on average and, in fact, worsened relative to their pretest scores. An effect size was calculated for each measure used in the study. When outcome measures were described in the method section, but statistics were not reported in the results section, an effect size of zero was assigned to that measure. Zero was also assigned as an effect size when results were reported as nonsignificant. As Casey and Berman (1985) point out, this is a conservative procedure and when it is not done, inflated effect size estimates result.

Although it is common for outcome studies to utilize more than one outcome measure within a given study (Lambert & Hill, 1994; Robinson, Berman, & Neimeyer, 1990), nevertheless, the practice does create problems of independence if a study is allowed to contribute more than one effect size to the overall average effect size estimate. Doing so

also gives extra weight to those studies using the greatest number of outcome measures. Consequently, in this study, the effect sizes of all measures calculated within a given study were averaged so that each study contributes a single effect size to the overall average standard effect size estimate. These average effect sizes were then used to provide an overall average standard effect size for all studies under review.

One concern in the calculation of effect sizes relates to the outcome measures which are utilized within a given study. Lambert and Hill (1994) indicate that the manner in which outcome is measured (e.g., source of measure, type of measure) is related to the effect size obtained from that measure. For example, in source of measure, therapist and self-report instruments generally obtain higher effect sizes than trained observer or physiological measures (Lambert & Hill, 1994). A self-report measure of alcohol intake is likely to give a higher effect size for a substance control treatment than is a physiological measure designed to test blood alcohol. In similar fashion, the type of outcome measure can influence the effect size obtained. Measures which estimate the effect of treatment relative to a specific goal of treatment (e.g., the BDI used to measure decreases in depression) tend to give higher effect sizes than do those that measure global behavior (Lambert & Hill, 1994). Thus, differing types and sources of outcome measures were examined for variation on effect size.

Effect Size Analysis

Hedges and Olkin (1985) assert that studies with large sample sizes tend to estimate the effect size true score more reliably than do studies with small sample sizes, and suggest weighting the effect sizes by the inverse of their variance, thus giving more influence to studies which have larger sample sizes. The appropriateness of this assumption was

empirically tested to determine whether weighted effect sizes should be used by two separate analyses conducted prior to the primary analysis. To test whether effect sizes derived from small N studies were more variable (or less reliable) than studies with large samples, effect sizes were divided into five categories according to the magnitude of the sample size in the study. The variability in effect size within these categories was then tested for homogeneity with an F ratio of the largest variance to the smallest variance, resulting in the variances of each of the five categories being compared to each other. Results indicated that a curvilinear relationship existed between sample size and variability in effect size. Both the largest and smallest samples produced less variable effect sizes than studies with medium sized samples. This finding is contrary to the assumption made by Hedges and Olkin since the variance is expected to decrease as the sample size increases. In order to test whether larger N studies produced different effect sizes than small N studies, the correlation between effect size and sample size was calculated. No significant linear relationship was found ($r(50) = -.2235$. $p = .119$), indicating that large N and small N studies produced comparable magnitude effect sizes. Given these empirical findings, the weighting suggested by Hedges and Olkin (1985) appears to be inappropriate and, thus, effect sizes were not weighted in any analysis.

Comparison with wait-list controls. Fifty of the selected studies compared group treatment to a wait-list control within the same study. Cohen's (1977) formula was utilized to calculate an effect, and in doing so, the posttest mean of the wait-list group was always subtracted from the posttest mean of the therapy group. Once effect sizes were calculated for each measure utilized within a study, several analyses were performed. First, an average effect size by source of outcome (i.e., self-report, independent observer rating, significant

other rating, therapist rating, objective/instrumental measure) was calculated by averaging the effect sizes within each source level. In the same manner, an average effect was calculated by content or focus of the outcome measure (e.g., general symptomatology, personality, target symptoms, somatic complaints, social adjustment). F tests were then conducted comparing the average effect sizes obtained across source categories and across content categories, thus allowing for a comparison of effect sizes by both the source and content of all outcome measures.

Second, for studies which included a wait-list control, the association between type of group (i.e., kind and amount of structure) and effect size was investigated and effect sizes were categorized according to the type of structure (ranging from high therapist direction/low client/group interaction to low therapist direction/high client/group interaction, see Figure 2) utilized in the group. F-test and post hoc comparisons were made to determine which of these structural types were most highly related to effect size.

Finally, to determine if gains made by group treatments exceeded those made by wait-list controls, effect sizes for each measure within a study were averaged to achieve a study effect size; subsequently, these study effect sizes were averaged across studies to achieve an overall average standard effect size. At this point, a t-test comparing the overall average standard effect size to 0.0 was conducted—an effect significantly greater than 0.0 indicates group treatment better than no treatment.

Comparisons between types of group treatments. Thirty-four studies directly compared two or more different group treatments within the same study. Each group treatment utilized within a given study was categorized according to the four structural

categories (Figure 2) and effect sizes comparing treatments within studies across this variable were calculated. Calculation of the difference between the means of two different treatment structures was done in a consistent manner so that a positive effect size indicated more improvement in the first group; negative effect sizes indicate more improvement in the second group.

Once an alternative treatment effect size was calculated for each study, effect sizes were averaged across studies, thus affording comparisons between the same types of group treatment. For example, if five studies compared a low therapist direction/low client/group interaction treatment to a treatment utilizing high therapist direction/low client/group interaction, the effect sizes for these studies were averaged in order to obtain an overall average effect size for that comparison. Next, a t-test was conducted to test the overall average effect size for that comparison against the null hypothesis that it would equal 0.0—an effect size significantly different from 0.0 indicates one treatment obtains better outcomes than the other treatment.

Estimates of absolute effect size. As noted above, effect sizes were computed for individual treatments and wait-list controls by calculating the deviation score of the posttest mean minus the pretest mean. This procedure was performed for each treatment utilized within a study by deriving a pre- post effect size for each measure employed within a study and then averaging effect sizes across measures; in this manner, each study contributed one effect size for each group examined within it.

These pre-to-post treatment allowed a test of the association between the coded moderator variables and the variation in overall treatment improvement. This enables one to

determine whether differential effectiveness exists and answers one question posed by this study: "Is group therapy differentially affected when client, therapist, group, and methodological variables differ? In this analysis, the effect sizes contributing to the overall average standard effect size are sorted according to the levels of a specific moderator variable (e.g., the diagnosis of a patient) and an F test is conducted to determine if the mean effect sizes of the levels are different. If different, post hoc t-tests for the difference between two mean effect sizes can be conducted to determine which means are different, thereby providing information regarding factors which may relate to the effect size, or to the differential in group therapy.

The relationship between continuous variables and effect size can also be assessed by conducting a linear regression analysis—with effect size as the predicted variable and the continuous variable as the predictor. For example, such an analysis could be used to determine the relationship between effect size and treatment dosage for studies included in the meta-analysis by regressing dosage against effect size and determine if such a regression is significant. If significant, the relationship between dose and effect size can be quantified.

When a continuous variable, such as age, was being investigated, the correlation between the variable and effect size was computed and tested for significance. In other words, the correlation between effect size and age was determined and tested for significance. A standard alpha level of .05 was used in all analyses. However, given the absence in the literature of any large scale meta-analysis that uses studies focusing on group treatment as a predominant modality, trends ($p < .10$) are also reported.

Results

Component Characteristics

An examination of various components provides some interesting observations regarding the treatment setting and theoretical focus, the professional alliance of the therapists, and the client diagnoses present in tne 116 studies (Table 3). It is both striking and disappointing that in two-thirds of the studies, they either do not report, or the report is so unclear that treatment setting is not distinctive. The treatment settings most used in these study sites are university counseling centers and correctional institutions, perhaps indicating the ease of subject accessibility more than a selected choice of targeted population. Also disappointing is the fact that almost one-half of the studies, for whatever reason, neglected stating the theoretical orientation of the treatment (Table 3). Thirty-three percent of the studies reporting the theoretical basis of treatment are behavioral or cognitive-behavioral in orientation. Doctoral-level psychologists and trainees each provide one-fourth of the leadership role in the psychotherapy groups. Again regrettably, the largest percent (33%) either did not report or the report was so unclear that professional degree could not be determined (Table 3).

Patient diagnosis received the most descriptive attention with nearly one hundred percent reporting some classification (Table 4). In one-fourth of the studies, group treatment is aimed toward alleviating concerns of patients whose primary problem is medically or physically based (e.g., cancer, herpes). Clients whose problems are anxiety or depression also account for one-fourth of the diagnoses.

Post-treatment Change Comparisons

Treatment type comparison with wait-list controls. Fifty of the 116 studies examined compared post-treatment change in clients participating in one or more of the four types of group therapy (see Figure 2) against change exhibited by clients who had been in wait-list control groups for a comparable period of time. Effect sizes from these comparisons are summarized in Table 5 and provide an estimate of the amount of improvement active treatment provides above and beyond that which can be explained by maturation and/or spontaneous remission effects (Cook & Campbell, 1979). As evidenced by values in the confidence interval column, both the overall effect size and those associated with each type of group demonstrate that group therapy produces more improvement when compared to the untreated controls. For instance, the overall effect size of .54 suggests that the average client in group treatment, ignoring treatment type, is better off than 71 percent of the clients in the untreated control group.

The overall variability of effect sizes across the different structural types of group treatment (.30 to .60) did not explain a significant amount of variance, $F(3,176) = 1.92$, $p = .12$. This means that while different average levels of improvement were associated with different types of group treatment, none were reliably different from one another. However, exploratory least significant difference contrasts revealed a trend ($p < .10$) favoring Type I group (therapist/topic) over Type III (therapist/interaction).

Group treatment type comparison. A related question was whether or not the structural treatment types differed in their effectiveness when they were directly compared with one another rather than the indirect comparison noted above with the wait-list control

groups. Forty-five comparisons were drawn from thirty-four studies that directly contrasted two or more of the four structural classifications: Type 1 (therapist/topic), Type 2 (client/topic), Type 3 (therapist/interactive), and Type 4 (client/interactive). In these direct comparisons, there were no reliable differences between the four treatment types and combinations of treatment types when set against the .05 level of significance (Table 6).

However, since the t distribution is particularly sensitive to small Ns, which are abundant in this comparison, the treatment types were also compared using an alpha of .10. In the latter case, two comparisons were found to be different. Type 1 group treatment showed greater improvement than Type 2, the distinction between these two groups is who appears to be directing the flow of discussion. Type 1 group (therapist/topic) also produced greater improvement than Type 4 (client/interactive). This first difference converges with the findings of the treatment type comparison with wait-list groups, where there was a trend for Type 1 groups to produce higher levels of improvement. The second difference contradicts the comparative findings of group type with wait-list controls, wherein Type 4 groups produced equivalent levels of improvement when compared to Type 1 groups (Table 5).

Outcome source comparison on post-treatment change. When the outcome measures used to track post-treatment change were evaluated, five separate rating sources were identified: independent observation, objective, self report, significant other, and therapist (Table 7). The majority of the studies relied on self-report measures; objective ratings were used most often in studies that examined group therapy being applied to medical problems. When the variability of effect sizes was partitioned and tested by outcome source an overall

main effect was found, $F(4,177) = 2.97$, $p = .02$. Post hoc t-tests revealed that therapist ratings produced significantly higher effect sizes than did self-report, independent observer, and objective ratings. These differences were significant after correcting for family-wise error rate inflation and account for twice as much variance as did treatment type (6.3%). Additionally, significant others produced higher ratings of improvement when compared with objective ratings.

Outcome content comparison on post-treatment change. Since level of improvement has been shown to vary systematically by the content of the outcome measure being used in some psychotherapy studies, effect size variability was examined with the five categories suggested by Lambert & Hill (1994): general, personality, social adjustment, somatic, and target complaint. No reliable overall relationship was found between the content area tapped by the outcome measure and the average effect size, $F(4,177) = 1.27$, $p = .13$ (Table 8). Although 4% of the variance is accounted for by this classification, the high degree of within category variability indicated by the large standard deviations result in no reliable differences between the different content dimensions of the outcome measures. However, exploratory post hoc comparisons again revealed a trend ($p < .10$) suggesting that higher levels of improvement were reported for measures that assess change in social adjustment when compared to the remaining four content categories.

Pre- Post-Treatment Change Comparisons

An additional focus of this study was an exploration of factors that might explain differences in pre-to-post improvement rates in group psychotherapy. Several treatment (setting, dosage, group size and composition), client (diagnosis, chronicity of disorder,

gender, age) and methodological (randomization, attrition) variables were examined to ascertain if they might account for higher or lower rates of improvement. Results indicated that reliable predictions regarding pre-to-post change could be made from two client (diagnosis and gender), one treatment (composition), and one methodological (attrition) variable.

Client population/diagnosis. A significant overall main effect was found for client diagnosis in the prediction of pre-to-post effect size variation for the active treatment group, $F(11,83) = 3.27$, $p < .001$. Simply stated, clients from different diagnostic groups, on average, obtained significantly different benefits from group treatment (Table 9). Post hoc analyses of the effect sizes for the active treatment condition revealed that clients with a depressive disorder exhibited more improvement than clients from nine other diagnostic groups (i.e., anxiety, criminal behavior, inpatient, medical, neurotic, normal, outpatient, sexual abuse, stress). As well, clients with an eating disorder diagnosis showed more improvement than clients from six other diagnostic groups (i.e., criminal behavior, medical, neurotic, normal, outpatient, and stress related diagnoses). No other reliable differences were found between the diagnostic groupings.

The comparison of client improvement or deterioration when they received no treatment (wait-list control group) and when change over time is generally attributed to maturation, spontaneous remission, or deterioration effects (Cook & Campbell, 1979) also yielded a significant overall main effect, $F(9,31) = 8.23$, $p < .01$. Post hoc comparisons revealed two significant contrasts. Clients being treated for a medical condition exhibited significantly more deterioration than did clients diagnosed with depression, eating disorder,

normal, and stress classifications. In addition, those with a criminal behavior classification also exhibited more deterioration than those with depressive diagnoses.

A second way to examine the effect sizes depicted in Table 9 is to assess whether each is significantly greater than zero using a one tailed t-test. Post hoc analyses of the effect sizes for the wait-list control pre-to-post change revealed that in only one case was there reliable improvement or deterioration. Clients receiving a medical diagnosis who went untreated showed reliable deterioration. In contrast, all but four (criminal behavior, general outpatient, substance abuse, and thought disorder) of the active treatment categories demonstrated reliable improvement. Those that did not exhibit reliable improvement generally had one or more effect sizes that were negative indicating that clients in some studies from these diagnostic groupings deteriorated while in group treatment.

A final way of understanding the effect sizes in Table 9 is to examine the difference between pre-to-post effect sizes for the active versus the wait-list groups for the same diagnostic group; this is depicted in the last column. This procedure seems justified since the wait-list studies are a subset of the active group treatment studies.

For instance, although the depression effect size has the largest absolute value (1.25) in the active treatment condition and thus appears to be the "best" candidate for group treatment, removing the average improvement seen in untreated depressives (.35) estimated by the wait-list control group places it in second position, superceded by eating disordered clients who do not show as much improvement (.15) when they are left untreated. Thus, the final column of Table 9 can be used as an index of the relative improvement expected from group treatment when one takes into consideration "natural" gains made by untreated clients.

Client gender. A second, but less robust client variable that predicted improvement in group treatment was client gender. When groups whose composition was comprised of both male and female clients were compared to all male and all female groups, members of the mixed groups showed more improvement than did clients in same gender groups (Table 10). However, this was only the case in those studies where direct comparisons were made between treatment types (i.e., Type 1 group vs Type 2, Type 1 vs Type 3); when all pre-post studies were examined, the finding was not supported.

Composition. A third variable of note relates to the composition of the group as defined by whether the group was homogeneous or heterogeneous regarding diagnosis or symptom complaint. Eighty-eight of one hundred and one studies in this analysis described the composition of the groups as homogeneous in construction (e.g., eating disorder, medical condition, depression). It is unclear whether clinical or research reasons motivated the composition decisions. The average improvement in members from homogeneous groups was higher than that of heterogeneous membership groups (Table 10).

Attrition. A final variable shown to be related to treatment outcome was attrition. Fifty-two studies report the percent of attrition realized in their investigation. When these percentages were correlated with the pre-post effect sizes from the same studies, a reliable negative association was found, $r(51) = .29$, $p < .05$. Specifically, the larger the attrition rates in a study, the smaller the pre-to-post improvement realized.

Discussion

The above findings provide the first quantitative estimate of the effectiveness of group therapy taken from a large number of investigations where the group format is the

predominant focus. The fact that 71 percent (ES=.54) of the clients in the active treatment group were better off than untreated controls provides support for group treatment being an efficacious treatment alternative. However, group therapy's effectiveness can be best understood when compared to effect size estimates derived from the general psychotherapy literature (Figure 3). For instance, Lambert and Bergin (1994) give an estimate of seventy-five percent (ES=.82) for the effectiveness of active psychotherapies when contrasted with no treatment controls. Their estimate can be considered reliable as it was based on an average drawn from 15 meta-analyses summarizing 1080 separate studies in the general psychotherapy literature.

Although the group therapy effectiveness estimate in this study is similar to the Lambert and Bergin (1994) estimate, it does fall well below other estimates found in the literature. One explanation of the lower effect size for group treatment resides in the combination of findings from the present and a past review. In the present analysis, a significant negative correlation was found between percent of attrition and the effect size. Thus, group studies with higher attrition rates tended to have smaller effect sizes. Additionally, Burlingame, Kircher, and Taylor (1994) report that the attrition rate in the group literature during the same time period for studies examined in this meta-analysis (1980-1992) was nearly twice as high as that reported in meta-analyses from the individual literature (e.g.,Shapiro & Shapiro, 1983). Thus, the higher attrition rates found in the group literature as a whole, when coupled with the tendency for attrition to weaken the effects of treatment (as suggested by Kazdin [1994] and found in this study), might easily have resulted

in the lower overall effect size for group treatment. Thus, when evaluating the effectiveness of group treatment, level of attrition seems to be an important consideration.

Although the main focus of the study was on the efficacy of group treatment, an equally important goal was to begin to understand the role and influence of structure in the group therapy process, and if possible, identify what contribution that structure makes to the overall effectiveness of treatment. As such, a beginning effort was made to classify the treatment groups according to what (who—therapist/client/group) was driving the therapeutic process and what was the focus of the discussion (topic/interaction). While there were no reliable differences in the treatment type comparison with wait-list controls, th e was a trend for the most structured groups (Type 1-therapist directed, topic centered) to produce greater improvement than the less highly, or differently, structured Type 3 group (therapist directed, interaction centered). Type 4 group (group directed, interaction centered) also had an equivalent level of average improvement as Type 1 groups when compared with no treatment controls. Stated differently, those groups in which the therapist was the strong force directing the group and where member discussion was on a specific topic and didactic in nature (such as anxiety reduction, weight reduction, disruptive thoughts, etc.), tended to have greater improvement than did groups with strong therapist direction but in which clients were responsive to one another and the discussion was more interactive. Although Type 1 groups indicated greater improvement rates when compared with wait-list controls, there were no significant differences between treatment type groups when they were directly compared with one another.

The finding that the level of improvement achieved by group treatment significantly varied depending on the source of the outcome measure finds mixed support from other reviews (Lambert & Hill, 1994; Robinson et al., 1990; Dush, et al., 1983; Miller & Berman, 1983; Shapiro & Shapiro, 1982a). For instance, significant other ratings did reliably exceed physiological measures as suggested by two previous reviews (Lambert & Hill, 1994; Shapiro & Shapiro, 1982a), as well, therapist ratings produced larger effects than self-report and significant other ratings (Lambert & Hill, 1994). However, self-report measures were not reliably higher than independent observer ratings as suggested by two previous reviews (Miller & Berman, 1983; Shapiro & Shapiro, 1982a). Moreover, two previous reviews (Robinson et al., 1990; Dush, et al., 1983) found no relationship with effect size, although that runs counter to this study. Since these previous reviews are based primarily on individual therapy studies, it is difficult to direct a parallel rationale in order to draft reasonable explanations for the mixed findings. Thus, the present source differences are only preliminary guides for interpreting groups' effectiveness.

No relationship was found between the content of the outcome measure and the level of improvement. This finding parallels that of Robinson et al, (1990) and counters two previous reviewers' conclusions (Lambert & Hill, 1994; Shapiro & Shapiro, 1982a). In contrast to the finding that targeted outcome measures yield higher effect sizes (as found in the general psychotherapy literature), there was a trend in this study suggesting that social adjustment measures yield the highest levels of improvement in group treatment. Although this finding is compatible with the social microcosm theory used to explain why group

treatment is effective (Yalom, 1995), it is nonetheless tentative since the estimate was based on seven comparisons and reached only trend status.

The second major thrust of this study was an examination of the treatment, client, and methodological factors that might explain differences in pre-to-post-treatment change. In addition to attrition, three other variables were reliably related to improvement or deterioration: client diagnosis, gender, and group composition.

Client diagnosis was the strongest variable, accounting for approximately 30% of the effect size variance. Depression was the most effectively treated diagnosis, with the average group patient improvement surpassing nearly 90% of the patients' pre-treatment levels of distress. While the average effect size of 1.25 is comparable to the average effect size reported for the treatment of depression (Lambert & Bergin, 1994—based on five separate meta-analyses (1.11), it must be remembered that the effect size in the present study reflects pre-to-post-treatment improvement. These estimates are generally inflated above effect sizes drawn from comparative analyses (active vs. no treatment). A more realistic, albeit crude, estimate of the comparative effectiveness of group treatment is the relative effect size difference in which wait-list control improvement rates are subtracted from the active treatment estimates, thus yielding a substantially smaller value (.90).

In addition to depression, the only other diagnostic group that yielded significantly higher improvement rates was eating disorders. The effect size in the present study (1.21) surpassed the only other meta-analysis of group treatment with this population. Specifically, Fettes & Peters (1992) reported a pre-post comparison effect size for bulimia of .75. The primary difference in these two estimates is that the present analysis used only experimental

or quasi-experimental studies, resulting in an increase in methodological rigor (Cook & Campbell, 1979) and, also, fewer studies.

Although several other observations could be made regarding the diagnostic categories, the final one for this presentation relates to deterioration. Specifically, it is interesting to note that medical conditions were the only diagnostic group in the no-treatment condition that showed reliable deterioration. Given the large number of studies contributing to this average, an ethical case could be made for the importance of patients in these settings receiving immediate treatment, or some alternative treatment, rather than being placed in a no-treatment control group.

The remaining two findings relate to the composition of the group. Homogeneity in group composition, with respect to diagnosis or symptom complaints, was related to greater levels of improvement, and accounted for five of pre-to-post-treatment effect size variances. This finding is noteworthy given the high number of studies contributing to the effect size estimate (N=101). However, the effect size may be reflective of, and inflated due to the high number of homogeneous depression and eating disorder studies (N=32) that contributed to the average effect size of the homogeneous category. As mentioned above, the two highest pre-post treatment effect sizes came from the depression and eating disorder diagnostic groups. These two diagnoses make up over one-third (36%) of the studies in the homogeneous composition category. Post hoc analyses did not bear out this explanation and suggest that the two factors (diagnosis & homogeneity) account for independent effect size variance. This finding may be important since the most recent reviews on composition

(Yalom, 1995; Piper, 1994) suggest that little empirical support exists for its relationship to treatment outcome (except those found in analogue studies).

Finally, heterogeneity of gender (combined, female/male groups) was related to higher levels of improvement. This finding also does not find support in the most recent reviews of client variables (Piper, 1994). Given that this relationship was not robust over all comparisons, it must be replicated in future studies before greater confidence can be placed in the value of combined gender group composition.

Summary

A summary of the primary findings of the study is provided in Figures 4 and 5. In both figures, the pie chart reflects 100 percent of the effect size variance. Thirty-six percent of the variance has been assigned to measurement error. This value was obtained by assuming that the average outcome measure used in group therapy research has a mean reliability coefficient of .80 (Burlingame, et al., 1994). The squared reliability coefficient provides an estimate of true score variance with that remaining being error variance (36%).

Figure 4 depicts that only 6.3 percent of post-treatment effect size variance can be reliably explained by the source of the outcome measure. If one includes, the trends found for treatment type and the content of the outcome measure, the total variance increases to 13.5 percent. Clearly, very little variance in post-treatment comparative improvement rates was predicted in the present analysis.

Figure 5 portrays a very different pattern. Forty-four percent of pre-to-post-treatment improvement can be predicted by a combination of three factors. Client population/diagnosis explains nearly one-third of the differential effectiveness of group therapy. The next highest

factor is attrition, which relates to the integrity of the study design in being able to detect a treatment effect (Kazdin, 1994). Finally, five percent is attributed to the homogeneity of diagnoses in group treatment. After cataloging measurement error, only 19% of the variance is left unaccounted.

Since this is the first large scale examination of the differential effectiveness of group psychotherapy, it is clear to us that these findings are only preliminary . However, the results are promising and, in many respects, correspond to findings in the general psychotherapy literature. The bottom line is that we now have a more refined representation of the effectiveness of group treatment, and hopefully, a better direction from which to guide future research and application.

References

Baer, R. A., & Nietzel, M. T. (1991). Cognitive and behavioral treatment of impulsivity in children: A meta-analytic review of the outcome literature. Journal of Clinical Child Psychology, 20, 400-412.

Bednar, R. L., & Kaul, T. J. (1994). Experiential group research: Can the cannon fire? In S. L. Garfield & A. E. Bergin (Eds.), Handbook of psychotherapy and behavior change (4th ed., pp. 631-663). New York: Wiley.

Bergin, A. E., & Garfield, S. L. (Eds.). (1994). Handbook of psychotherapy and behavior change (4th ed.). New York: John Wiley & Sons.

Budman, S. H., Simeone, P. G., Reilly, R., & Demby, A. (1994). Progress in short-term and time-limited group psychotherapy: Evidence and implications. In A. Fuhriman & G. M. Burlingame (Eds.), Handbook of group psychotherapy (pp. 370-415). New York: Wiley.

Burlingame, G. M., & Fuhriman, A. (1994). Epilogue. In A. Fuhriman & G. M. Burlingame (Eds.), Handbook of group psychotherapy (pp. 370-415). New York: Wiley.

Burlingame, G. M., & Fuhriman, A. (1995, June). The comparative effectiveness of group psychotherapy: A fifty-year perspective on outcome research. Paper presented at the Society for Psychotherapy Research 26th Annual Meeting, Vancouver, Canada.

Burlingame, G. M., Kircher, J. C., & Taylor, S. (1994). Methodological considerations in psychotherapy research: Past, present, and future practices. In A. Fuhriman & G. M. Burlingame (Eds.), Handbook of group psychotherapy (pp. 41-80). New York: Wiley.

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. Psychological Bulletin, 98(2), 388-400.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences (Rev. ed.). New York: Academic Press.

Cook, T., & Campbell, D. (1979). Quasi-experimentation: Design and analysis issues for field settings. Boston: Houghton Mifflin.

Dagley, J. C., Gazda, G. M., Eppinger, S. J., & Stewart, E. A. (1994). Group psychotherapy research with children, preadolescents, and adolescents. In A. Fuhriman & G. M. Burlingame (Eds.), Handbook of group psychotherapy (pp. 340-369). New York: Wiley.

Dies, R. (1994). Therapist variables in group psychotherapy research. In A. Fuhriman & G. M. Burlingame (Eds.), Handbook of group psychotherapy (pp. 114-154). New York: Wiley.

Dush, D. M., Hirt, M. L., & Schroeder, H. H. (1983). Self-statement modification in the treatment of child behavior disorders: A meta-analysis. Psychological Bulletin, 106 (1), 97-106.

Fettes, P. A., & Peters, J. M. (1992). A meta-analysis of group treatments for bulimia nervosa. International Journal of Eating Disorders, 11, 97-110.

Forsyth, D. R. (1990). Group dynamics (2nd ed.). Pacific Grove, CA: Brooks/Cole Publishing Company.

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.

Fuhriman, A., & Burlingame, G. M. (1994). Group psychotherapy: Research and practice. In A. Fuhriman & G. M. Burlingame (Eds.), Handbook of group psychotherapy (pp. 3-40). New York: Wiley.

Gatchel, R. J., Baum, A., & Krantz, D. S. (1989). An introduction to health psychology (2nd ed.). New York: Random House.

Gorey, K. M., & Cryns, A. G. (1991). Group work as interventive modality with the older depressed patient: A meta-analytic review. Journal of Gerontological Social Work, 16, 137-157.

Grossman, P. B., & Hughes, J. N. (1992). Self-control interventions with internalizing disorders: A review and analysis. School Psychology Review, 21, 229-245.

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. New York: Academic Press.

Hunter, J. E., & Schmidt, F. L. (1990). Methods of meta-analysis: Correcting error and bias in research findings. Newbury Park, CA: Sage.

Johnson, B. T. (1989). DSTAT: Software for the meta-analytic review of research literatures. Hillsdale, NJ: Lawrence Eribaum.

Kaplan, R. M., Sallis, J. F., & Patterson, T. L. (1993). Health and human behavior. New York: McGraw-Hill.

Kazdin, A. E. (1994). Methodology, design, and evaluation in psychotherapy research. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change (4th ed., pp. 19-71). New York: Wiley.

Klein, R. H., Brabender, V., & Fallon, A. (1994). Inpatient group therapy. In A. Fuhriman & G. M. Burlingame (Eds.), Handbook of group psychotherapy (pp. 370-415). New York: Wiley.

Lambert, M. J., & Bergin, A. E. (1994). The effectiveness of psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change (4th ed., pp. 143-189). New York: Wiley.

Lambert, M. J., & Hill, C. E. (1994). Assessing psychotherapy outcomes and processes. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change (4th ed., pp. 72-113). New York: Wiley.

Miller, R. C. & Berman, J. S. (1983). The efficacy of cognitive behavior therapies: A quantitative review of the research evidence. Psychological Bulletin, 94, 39-53.

Nietzel, M. T., Russell, R. L., Hemmings, K. A., & Gretter, M. L. (1987). Clinical significance of psychotherapy for unipolar depression: A meta-analytic approach to social comparison. Journal of consulting and clinical psychology, 55(2), 156-161.

Piper, W. (1994). Client variables. In A. Fuhriman & G. Burlingame (Eds.), Hand book of group psychotherapy (pp. 83-113). New York: Wiley.

Robinson, L. A., Berman, J. S., & Neimeyer, R. A. (1990). Psychotherapy for the treatment of depression: A comprehensive review of controlled outcome research. Psychological Bulletin, 108(1), 30-49.

Rosenthal, R. (1983). Assessing the statistical and social importance of the effects of psychotherapy. Journal of Consulting and Clinical Psychology, 51(1), 4-13.

Rosenthal, R. (1991). Meta-Analysis: A review. Psychosomatic Medicine, 53, 247-271.

Rothstein, H. R., & McDaniel, M. A. (1989). Guidelines for conducting and reporting meta-analysis. Psychological Reports, 65, 759-770.

Russell, R. L., Greenwald, S., & Shirk, S. R. (1991). Language change in child psychotherapy: A meta-analytic review. Journal of Consulting and Clinical Psychology, 59(6), 916-919.

Shadish, W. R. (1992). Do family and marital psychotherapies change what people do? A meta-analysis of behavioral outcomes. In T. O. Cook, H. Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. J. Light, T. A. Louis, & F. Mosteller (Eds.), Meta-analysis for explanation: A casebook (pp. 129-208). New York: Russell Sage Foundation.

Shapiro, D. A., & Shapiro, D. (1982a). Meta-analysis of comparative therapy outcome studies: A replication and refinement. Psychological Bulletin, 92, 581-604.

Shapiro, D. A., & Shapiro, D. (1982b). Meta-analysis of comparative therapy outcome research: A critical appraisal. Behavioral Psychotherapy, 10, 4-25.

Shapiro, D., & Shapiro, D. (1983). Comparative therapy outcome research: Methodological implication of meta-analysis. Journal of Consulting and Clinical Psychology, 51, 42-53.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). The benefits of psychotherapy. Baltimore: The Johns Hopkins University Press.

Stinchfield, R., Owen, P. L., & Winters, K. C. (1994). Group therapy for substance abuse: A review of the empirical research. In A. Fuhriman & G. M. Burlingame (Eds.), Handbook of group psychotherapy (pp. 458-488). New York: Wiley.

Tillitski, L. (1990). A meta-analysis of estimated effect sizes for group versus individual versus control treatments. International Journal of Group Psychotherapy, 40(2), 215-224.

Toseland, R., & Siporin, M. (1986). When to recommend group treatment: A review of the clinical and the research literature. This Journal, 36, 171-201.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. Journal of Consulting and Clinical Psychology, 55(4), 542-549.

Yalom, I. D. (1975). The theory and practice of group psychotherapy (2nd ed.). New York: Basic Books.

Yalom, I. D. (1985). The theory and practice of group psychotherapy (3rd ed.). New York: Basic Books.

RULES FOR SELECTING RESEARCH ARTICLES: 1980-1992

Inclusion

- Groups focused on group psychotherapy as a primary treatment modality.
- Key words used: counseling, psychotherapy, interpersonal/interactive process group, insight oriented, or specific orientations (e.g. cognitive-behavior, psychodrama, psychodynamic).
- Groups could be theme (grief, divorce) or diagnosis (depression, eating disorder, cancer) specific.
- Study had to focus on specific outcome such as symptom reduction. At times process variables (cohesion, insight, HIM) were used as outcome variables.

Exclusion

- Delete task groups and large group awareness training (EST, Lifespring, T-groups).
- Delete groups that are primarily didactic or psycho-educational.
- Delete investigations focusing on specific techniques (pregroup training, videotape feedback) or specific group processes (therapeutic factors).
- Delete studies that focus on specific client, therapist, or group variables.

Figure 1

META-ANALYSIS GROUP CLASSIFICATIONS

Category 1

Guiding force is therapist or manual
Treatment is instructive
Discussion focuses on topic or content
Members respond to specific subject matter or to specific actions or behavioral practice

Category 2

Moving force is client(s) or group
Treatment is structured around the topic or content
Discussion focuses on topic or content
Groups may or may not use a manual

Category 3

Guiding force is the therapist or manual
Therapist or a specific model of group therapy structures the treatment
Discussion promotes interactive, responsive group process
Discussion is on the client(s)—their reactions, behaviors, and feelings, with evidence of a here and now orientation

Category 4

Moving forces are the client(s) or group
The unique "social microcosm," created by the composition of group members, structures the treatment
Discussion promotes interactive, responsive group process
Discussion is on client(s) or group-as-a-whole—their reactions, behaviors, and feelings, with evidence of a here and now orientation

Category 5

Hospital, residential daily, therapeutic regimen offerings (e.g., milieu, individual, family, recreation, etc.)

Figure 2

Portrayal of average effect sizes that contrast wait-list control, group treatment, and general psychotherapy estimates

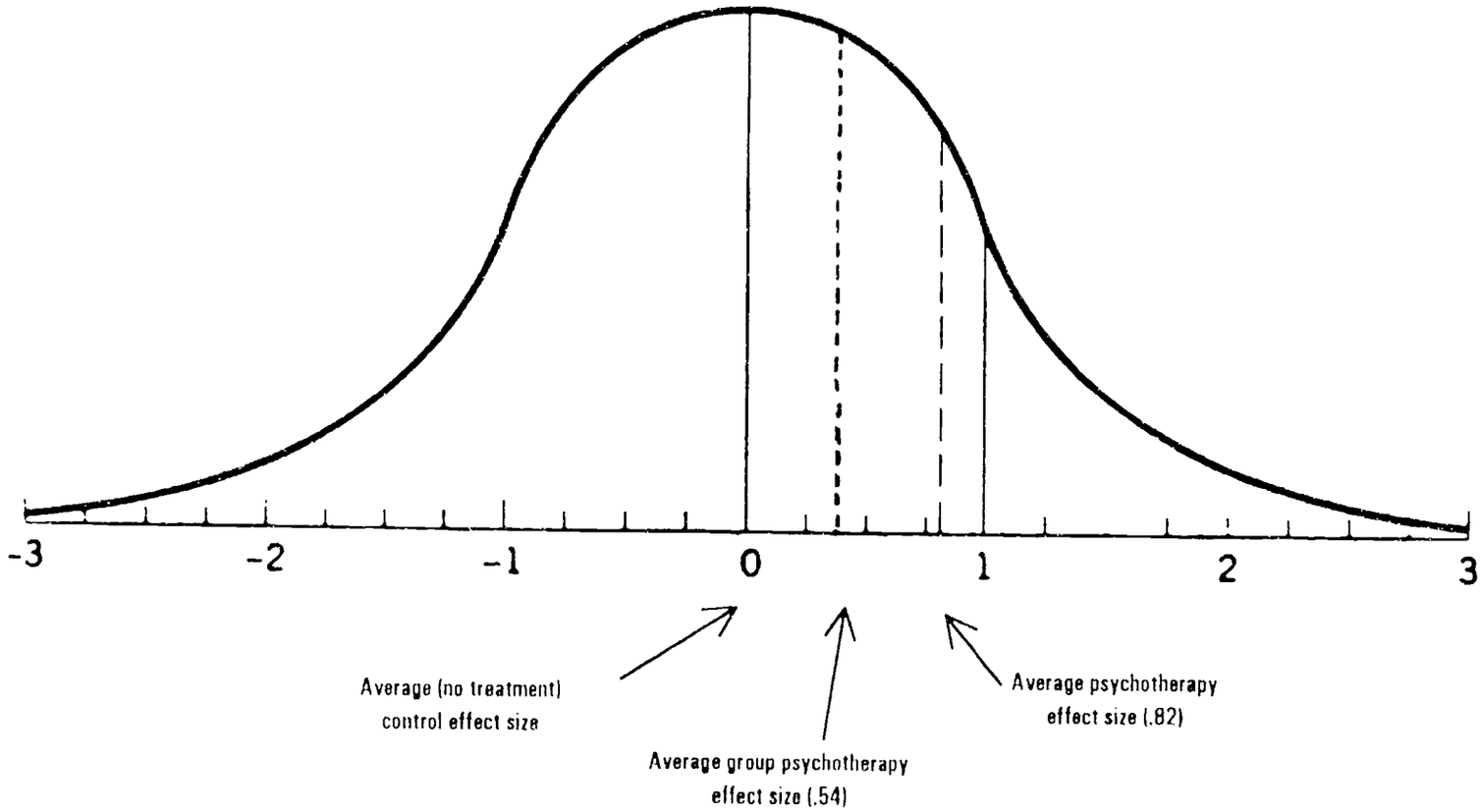
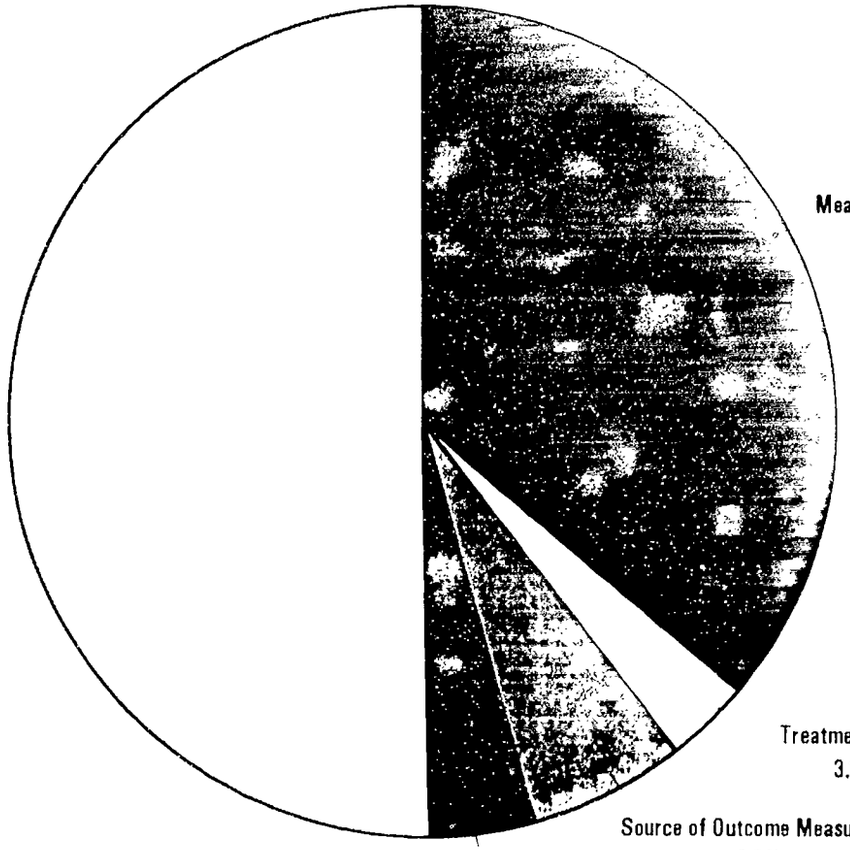


Figure Adapted from Lambert and Bergin (1994)

Figure 3

Accounting for post-treatment effect size variance: active vs wait-list control change (n=50)



* Trend only with p .10

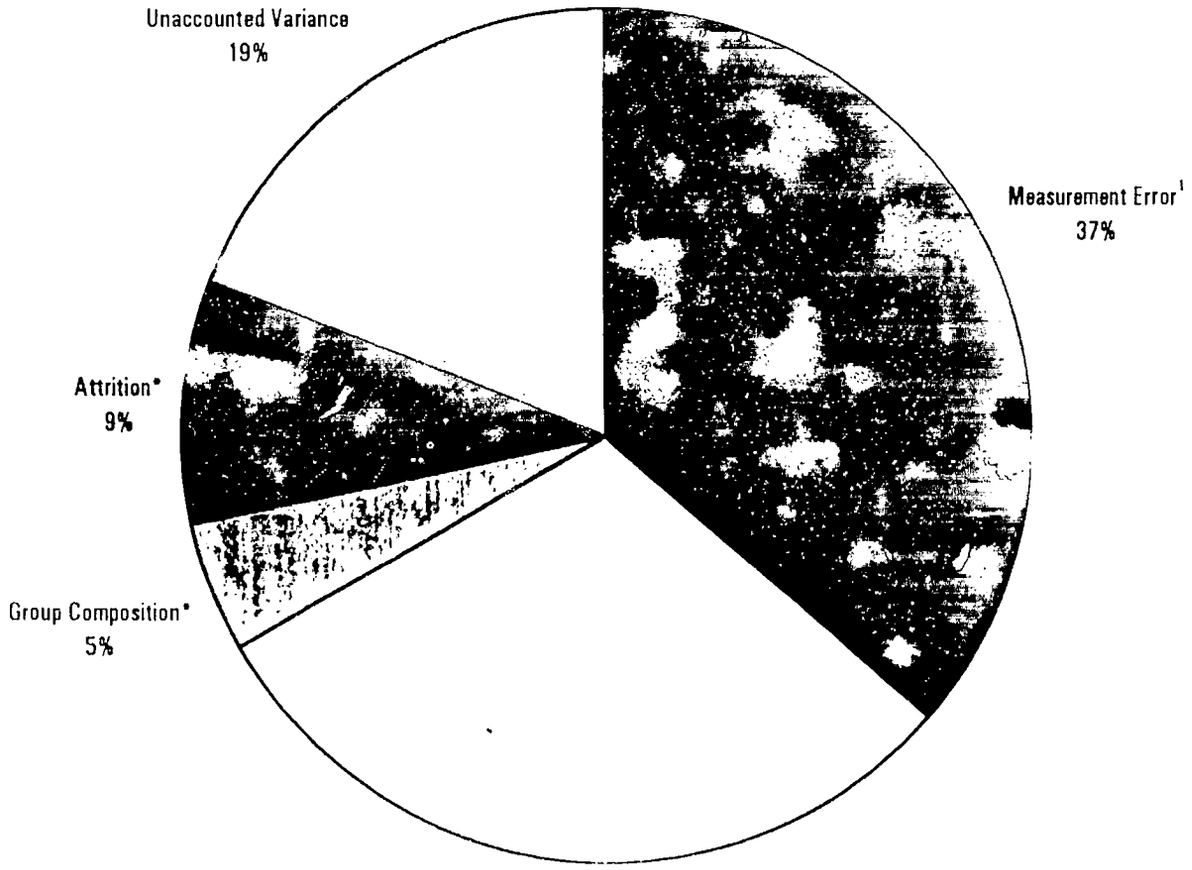
** Significant at p .05

[1]Based on an average reliability estimate of .80

13.5% of effect size variance is accounted for by group type and the source/content of outcome measure

Figure 4

Explaining pre-to-post treatment improvement rates (n approximates 100)



* $p < .05$

** $p < .001$

[1] Based on an average reliability estimate of .80

44% of effect size variance is accounted for by diagnosis, composition, and attrition

Figure 5

Table 1

Comparative Meta-analyses: Differential Effectiveness of Group vs. Individual Treatment—No Difference

Author	Tx Orientation	Patients	Effect Size	
			Individual	Group
Smith et al. 1980	heterogenous	heterogenous	.87	.83
Shapiro & Shapiro 1982	heterogenous	heterogenous	1.12	.89
Miller & Berman 1983	cognitive behavioral	heterogenous	.93	.79
Casey & Berman 1985	heterogenous	heterogenous	.82	.52
Weisz et al. 1987	heterogenous	child/adolescent	1.04	.62
Robinson et al. 1990	heterogenous	heterogenous	.83	.84
Tillitski 1990	heterogenous	heterogenous	1.35	1.35
Baer & Neitzel 1991	cognitive behavioral	children	.44	.79

Table 2

Comparative Meta-analyses: Differential Effectiveness of Group vs. Individual Treatment and Specific Populations—Reliable Differences

Author	Tx Orientation	Patients	Effect Size	
			Individual	Group
Dush et al. 1983	cognitive behavioral	heterogenous	.93 .71	.58 .36
Nietzel et al. 1987	cognitive behavioral	unipolar depression	10.06	12.47*
Russell et al. 1991		children	1.31	.10
Grossman & Hughes 1992		language deficiencies	.40	.80
Fettes & Peters 1992		bulimia nervosa		.75
Gorey & Crynes 1991		depressed elderly		.68

*Indicates average number of depressive symptoms

Table 3

Component Characteristics in Meta-Analytical Study

Treatment Setting	Percent
University Counseling Center	16.5
Hospital	2.5
Community Mental Health	1.3
Vetran's Administration Center	2.5
Correctional Institution	8.8
Outpatient Clinic	5.0
Unclear	15.0
Missing	43.8
Total	100.0

80 valid cases; 0 missing cases

Theoretical Orientation	Percent
Behavioral	10.0
Cognitive/Behavioral	23.8
Eclectic	7.5
Psychodynamic	7.5
Nondirective	5.0
Not Stated	46.3
Total	100.0

43 valid cases; 37 missing cases

Professional Degree	Percent
Psychologist	23.8
Social Worker	3.8
Master's (Unspecified)	8.8
Nurse	2.5
Trainee doctoral	22.5
Trainee master's	3.8
Unclear	21.3
Missing	13.8
Total	100.0

80 valid cases; 0 missing cases

Table 4

Component Characteristics in Meta-Analytical Study

Population/Diagnosis	Percent
Anxiety	12.5
Criminal Behavior	8.8
Eating Disorder	10.0
Inpatient (Unspecified)	1.3
Depression	10.0
Neurotic	3.8
Normal	2.5
Outpatient (Unspecified)	3.8
Medical Conditions	23.8
Sexually Abused	8.8
Stress	11.3
Thought Disorder	1.3
Missing	2.5
Total	100.0

78 valid cases; 2 missing cases

Table 5

Treatment Type Comparison with Wait-list Controls: Post-treatment Change

Treatment Type*	Data Points	Average Unweighted Effect Size	SD	95% Confidence Interval**
Therapist/Topic	125	.60	.66	.48 to .71
Client/Topic	18	.34	.35	.12 to .58
Therapist/Interaction	19	.30	.60	.02 to .59
Client/Interaction	18	.60	.37	.43 to .79
Overall	180	.54	.62	.45 to .63

*3.2% of effect size variance is accounted for by treatment type categories.
**If the values in the confidence interval include zero, then active treatment is considered to be no more efficacious than the wait-list controls using an alpha=.05.

Table 6

Direct Comparison of the Differential Effectiveness of Varying Types of Group Treatment

Group Treatment Comparison	N	Mean* Effect Size	SD	95% Confidence** Interval	90% Confidence Interval
1 vs 2	14	.27	.55	-.04 to .59	.01 to .53*
1 vs 3	6	-.36	.59	-.98 to .25	-.84 to .12
1 vs 4	6	.50	.59	-.12 to 1.12	.01 to .99*
2 vs 4	4	.14	.28	-.31 to .59	-.19 to .47
5 vs 5 + 1	3	-.36	.12	-.66 to -.06	-.57 to -.16
5 vs 5 + 3	6	-.22	.70	-.95 to .51	-.80 to .35
5 + 1 vs 5 + 3	3	.32	.28	-.37 to 1.02	-.15 to .80
5 + 3 vs 5 + 4	3	-.18	.54	-1.51 to 1.16	-1.08 to .73

*A positive effect size favors the first type of group treatment in the comparison, while a negative effect suggests higher levels of improvement in the second group type.
**If the values in the confidence interval include zero, then no reliable difference can be assumed to exist between the treatments being compared.

Table 7

Outcome Source Comparison with Wait-list Controls: Post-treatment Change

Source	Data Points	Unweighted Effect Size	SD	95% Confidence Interval
Independent Observation	15	.55	.95	.03 to 1.07
Objective	15	.28	.41	.05 to .51
Self	140	.53	.59	.43 to .63
Significant Other	9	.79	.36	.52 to 1.07
Therapist	3	1.48	.41	.46 to 2.50
Total	182	.54	.62	.45 to .63

6.3% of effect size variance is accounted for by source categories.

Table 8

Outcome Content Comparison with Wait-list Controls: Post-treatment Change

Content	Data Points	Unweighted Effect Size	SD	95% Confidence Interval
General	48	.47	.63	.29 to .65
Personality	30	.57	.54	.37 to .77
Social Adjustment	7	1.01	1.15	-.06 to 2.08
Somatic	22	.36	.42	.17 to .54
Target	75	.58	.61	.44 to .72
Overall	182	.54	.62	.45 to .63

4% of effect size variance is accounted for by content categories.
$F(4,177)=1.27$, $p=.13$, LSD 4 vs 5, 1 vs 5

Table 9

Client Diagnosis: Pre to Post Improvement or Deterioration Based on Average Effect Sizes

	Wait-list Control Groups			Active Group Treatment			Relative Effect Size Difference
	N	Mean Effect Size	SD	N	Mean Effect Size	SD	
Eating Disorder	6	.15	.18	13	1.21**	.68	1.06
Depression	4	.35	.40	19	1.25**	.71	.90
Criminal Behavior	1	-.35***	--	4	.41	.31	.76
Medical Condition	12	-.18**	.19	16	.49**	.38	.67
Anxiety Disorder	4	.14	.21	8	.73**	.18	.59
Sexually Abused	2	.02	.03	4	.55**	.28	.53
Outpatient	3	.07	.12	5	.58	.62	.51
Neurotic	2	.05	.07	5	.52**	.24	.47
Stress	5	.25	.33	7	.54**	.43	.29
Normal	2	.28	1.0	5	.40**	.21	.12
Inpatient	--	--	--	5	.73**	.28	--
Personality Disorder	--	--	--	1	1.17***	--	--
Substance Abuse	--	--	--	2	1.16	.24	--
Thought Disorder	--	--	--	4	.91	.70	--
Total	41	.13	.35	98	.90**	.62	

*Negative values indicate deterioration from pre-treatment while positive values indicate improvement.

**Values are significantly greater than zero at $p < .05$.

***Not able to calculate confidence interval.

30.2% of effect size variance is accounted for by client population/diagnosis.

Table 10

Client and Treatment Variables that Predict Pre to Post Treatment Improvement and Deterioration

	Gender of Group	N of Studies	Mean Effect Size	SD	F-Ratio	F-Prob
Type 1 versus Type 2 groups	Mixed	5	.80	.61	11.75	.006
	Female	7	.00	.05		
Type 1 versus Type 3 groups	Mixed	2	.16	.06	105.8	.002
	Male	3	-.90	.13		
Homogenous Groups	--	88	.84	.58	4.74	.03
Heterogeneous Groups	--	13	.47	.37		

14% of effect size variance is accounted for by gender and composition.