1989

A Meta-Analytic Review of School-Based Psychotherapy: A Re-Evaluation Concerning Methods and Procedures

Jill Carmody
Loyola University Chicago

A META-ANALYTIC REVIEW OF SCHOOL-BASED PSYCHOTHERAPY: A RE-EVALUATION CONCERNING METHODS AND PROCEDURES

by

Jill Carmody

A Thesis Submitted to the Faculty of the Graduate School of Loyola University of Chicago in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

May

1989

ACKNOWLEDGMENTS

The author would like to express her thanks and appreciation to Dr. Joseph Durlak for his guidance and support throughout this project. Dr. Durlak was truly instrumental in initiating the author's interest in the meta-analytic procedure and was extremely dependable in providing assistance, direction and inspiration to the author during the development of this report. Thanks are also extended to Dr. Fred Bryant for his helpful suggestions regarding the format and style of this report, for his advice concerning the statistical treatment of the data, and for his positive academic influence.

In addition, the author wishes to acknowledge Anne Wells and Claudia Lampman-Petraitis for their valuable periodic input during the construction of this project.

Finally, special gratitude is extended to the author's parents, Maurice and Mary Ann, and to Dan, for their unconditional support and encouragement throughout all of the author's academic pursuits.

VITA

The author, Jill Anne Carmody, is the daughter of John Maurice Carmody and Mary Ann (Howley) Carmody. She was born July 16, 1964, in St. Joseph, Michigan.

Her elementary education was obtained in suburban Palos Heights, Illinois. Her secondary education was completed in 1982 at Mount Assisi Academy, Lemont, Illinois.

In September, 1982, Ms. Carmody entered Loras College in Dubuque, Iowa, receiving the degree of Bachelor of Arts in Psychology in May, 1986. While attending Loras College, Ms. Carmody became an active member of the College Peer Counseling Services and, in 1986 she became a student member of the American Psychological Association.

In September, 1986, Ms. Carmody began graduate study at Loyola University of Chicago, in the Applied Social Psychology Program, and completed her Master of Arts degree in May, 1989.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CONTENTS OF APPENDICES

INTRODUCTION

The school has become an increasingly popular setting for the administration of various types of psychotherapy to children. In conjunction with this growing prevalence, assessments of the effectiveness of school-based psychotherapy has also increased. Due to this growing variety of school-based psychotherapy literature, many qualitative reviews have been produced, examining and summarizing the relative effectiveness of a limited variety of such treatments.

However, difficulties were found among many qualitative reviews of the school-based psychotherapy literature. Such reviews were found to be very limited in scope, susceptible to a subjective sampling of the literature, and often difficult for the reviewer himself/herself to summarize due to varying design features in the studies chosen. Thus, these qualitative reviews, as well as others, often did not summarize the literature efficiently.

An alternative technique to the qualitative review is meta-analysis. This technique allows for a considerably larger sample of studies to be reviewed and summarizes them in quantitative terms. Thus, relative effectiveness can be ascertained and summarized efficiently. Furthermore, meta-

analysis provides an overall value indicating, in general, how effective the intervention is.

In an effort to summarize the school-based psychotherapy literature, a previous meta-analysis was performed by Prout and DeMartino in 1984. Their meta-analysis, however, was limited in many respects which, in turn, poses threats to the validity of the conclusions which they offered. The present study examines these limitations to Prout and DeMartino's review by first highlighting the primary threats to validity to which a meta-analysis may be susceptible, and then investigating the weaknesses of Prout and DeMartino's meta-analysis in terms of these threats. Finally, in an effort to overcome these shortcomings, alternative procedures are offered and thus applied, which bestows the primary purpose of this meta-analysis.

REVIEW OF RELATED LITERATURE

School-based psychotherapy has long been utilized in schools as a means for modifying behavior and adjustment problems as well as preventing them. Stemming from the child guidance movement of the 1920's, mental health professionals began to focus their attention on children in the schools. Referrals to external mental health resources grew in frequency and finally by the early 1960's, mental health professionals began entering the school setting offering a variety of consultation programs (Durlak, 1983).

The need for psychological services in the schools is in even greater demand and growth today. The prevalence of psychotherapy in the schools is emphasized by Achenbach (1982) who states, "Probably more troubled children are dealt with in educational settings than all other settings combined" (p. 77). Weiner (1982) estimated that 20% to 30% of the children entering elementary school possess behavior problems ranging from moderate to severe, and half of these children require professional treatment. Furthermore, the prevalence of school maladjustment has been estimated to range from 2% to 30% depending on the criteria used to judge the degree of maladaptation (Durlak, 1983). Therefore, the school provides an appropriate setting in

which to offer therapy to children and to conduct either primary or secondary prevention programs.

Evaluating the effectiveness of psychological services provided in schools is becoming an area of great concern and a recommended focus for future research (Kratochwill, Feld, & Van Someren, 1986). Although the literature contains many studies assessing the effects of various forms of therapy applied in school settings, we do not have a clear, systematic and empirical evaluation of the effects of school based treatment. This is so because of the limited focus of many reviews and the problems posed by research studies that make conclusions and interpretations difficult.

Many recent narrative reviews have evaluated specific dimensions of school-based psychotherapy. Prout and Harvey (1978), for example, evaluated studies of desensitization procedures applied toward school-related problems. The authors concluded that the reviewed studies demonstrated the efficacy of desensitization procedures, particularly when combined with in vivo and operant techniques for reducing such school-related problems. Expressing confidence in these results, Prout and Harvey (1978) stated that "the combined approach is both theoretically and clinically sound" (p. 538).

However, other reviews of school-based psychotherapy studies have not been able to provide such strong

conclusions. Many difficulties arise when reviewing the literature that limit the conclusions reviewers can draw from the recruited studies. One of the primary difficulties is selecting a representative sample of related studies. The problem of subjectivity in selecting studies poses a major threat to the validity of the literature review (Cooper & Rosenthal, 1980); thus it is important to select a sample of literature that is representative of its population. However, this may require obtaining a large sample of studies which, in turn, may be difficult for the narrative reviewer to handle. As the reviewer encounters an increasing number of related studies, he/she "must rely on an extraordinary ability to mentally juggle relationships among many variables" (Light & Pillemer, p. 4, 1984). As a result, the reviewer tends to cover only a small sample of the literature, allowing for an analysis of the literature that is easier to comprehend.

Another primary difficulty arises when the research for the review contains a variety of methodological flaws. Narrative reviewers have no way of evaluating empirically how different methodological features relate to therapy outcome. Thus, when the reviewed literature contains poorly designed investigations, more narrative reviewers are apt to withhold or greatly qualify their final conclusions until more rigorous research is produced. Shaw and Wursten's (1965) review of studies on group procedures

used in schools serves as an example. Covering the period from 1953 to 1963, Shaw and Wursten (1965) distinguished among three types of group procedures: direct treatment with the student, indirect treatments offered through teacher consultation, and indirect treatments offered through parent counseling. Based on their review, Shaw and Wursten (1965) were able to make a general conclusion that most of their studies reported "successful" outcomes, yet were apprehensive about accepting such results because many studies possessed "inadequate controls, inadequate statistical procedures, and inadequate outcome criteria" (p. 32). They attributed this to publication bias, where studies which produce significant or favorable outcomes are more likely to get published than those that are nonsignificant. Shaw and Wursten (1965) included a small proportion of unpublished studies in their review; however, with the larger proportion being published studies they felt they could not ignore the publication bias effect "as a possible contributing factor to the preponderance of 'successful' outcomes reported" (p. 32). The relationship between design quality, type of study (published or unpublished), and type of outcome (significant or nonsignificant) would be interesting to assess from a collection of related studies. However, this would be difficult to do so in a typical narrative literature review.

Similarly, a review by Henry and Kilmann (1979) on group counseling with high school students resulted in limited conclusions. Henry and Kilmann (1979) stated, "Overall, the poor quality of the research suggests that the case for high school group counseling has not yet been demonstrated" (p. 44).

Reviews involving studies of other school-based psychotherapies have experienced similar design complications. For example, Hobbs, Moguin, Tyroler, and Lahey (1980) and Gresham (1985) evaluated the clinical utility of cognitive behavioral therapy in treating children's school problems. Unfortunately, conclusions were impaired due to inherent methodological limitations found in many of the studies included in the reviews. Poor study characteristics such as invalid outcome measures, inadequate descriptions of treatment procedures, and unspecified subject characteristics precluded any further interpretations of cognitive behavioral therapy's techniques.

Other difficulties narrative reviewers encounter when summarizing a group of related studies are dissimilar study characteristics (i.e. types of therapy, subject characteristics, outcome measures) and mixed results. Such characteristics often hinder the reviewer from drawing general or specific conclusions. For example, Rosenbaum and Drabman (1979) analyzed studies employing self-control

training for children to assist them in appropriately managing their own academic and social behavior. Rosenbaum and Drabman (1979) focused on four elements of the procedure: self-recording, self-evaluation, self-determination contingencies, and self-instruction. In sum, they found that self-recording produced modest and short-term changes, but could be improved if used in conjunction with reinforcement contingencies. This was also found in another review (Gresham & Lemanek, 1983). However, studies demonstrating the value of self-determined contingencies as opposed to externally-determined contingencies for reducing disruptive classroom behavior provided mixed results, with some supporting the former intervention and others supporting the latter. This result was also found among the studies of self-instructional training. In an effort to resolve such diverse outcomes, Rosenbaum and Drabman (1979) suggested that "future studies (should) attempt to identify variables resulting in greater magnitude of behavior change" (p. 472).

Diverse design features also presented problems in reviews attempting to determine which of the various components of cognitive behavioral training improve children's behavior. Gresham and Lemanek (1983), for example, reviewed cognitive behavioral training, focusing on the techniques of modeling, coaching, a combination of techniques, and self-control training. Results indicated

that children in the treatment group significantly improved their rates of social interaction over controls. However, due to the differences across studies, Gresham and Lemanek (1983) noted that treatment effects may have depended on the peer orientation of target children, if the modeling film was narrated in the first person or third person, and the degree of model-observer similarity. In addition, Gresham and Lemanek (1983) noted that the outcome measures used were generally global rates of social interaction, which tends to demonstrate the overall rate of peer-oriented behavior rather than its quality, which may be a more important variable. Gresham and Lemanek's (1983) review suggests the need to determine not only what components of an intervention contribute to favorable results and which contribute to unfavorable results, but also what degree or magnitude of change is achieved by each.

As the above reviews have demonstrated, qualitative reviews of the literature are often limited by several problems that may be inherently difficult to overcome. For example, since a large sample of the literature is often difficult for the reviewer to summarize, a small and perhaps unrepresentative group of studies is reviewed. As a result, the qualitative review is open to selection bias in terms of the data evaluated. Furthermore, methodological limitations in the studies reviewed are

often encountered and cannot easily be reconciled, leading to in tenuous conclusions. In addition, studies may be theoretically similar yet contain unique design features, making it difficult to determine the relative effectiveness of alternative therapies. Problems also arise when studies share similar design characteristics yet produce opposing outcomes.

Finally, relationships of magnitude cannot be assessed from the qualitative review. For example, cognitive behavior therapy is made up of many techniques such as modeling and coaching. It would be advantageous for the researcher as well as the practitioner to discover how successful each technique is with differing populations as assessed by different outcome measures.

In summary, the traditional literature review process is open to an unsystematic, subjective, and hence, biased cummulation of studies (Light & Pillemer, 1984). Furthermore, magnitude of impact as well as relative impact cannot be assessed because of the qualitative nature of the review. Since these obstacles often stand in the way of accurately answering important research questions, such as the one proposed here (how effective is school-based psychotherapy?), an alternative approach was needed, one that could assist in overcoming the limitations associated with the narrative review.

META-ANALYSIS

One relatively new alternative that can quantitatively integrate and evaluate the literature while overcoming some of the problems associated with the conventional literature review is meta-analysis. Meta-analysis is a systematic and quantitative approach to aggregating the findings of primary research and can be a powerful approach for summarizing the characteristics and corresponding results of many related studies. Promulgated by Smith and Glass (1977), this method requires the use of all relevant primary studies found in the literature. Based on an explicit set of criteria, a representative sample of theoretically relevant studies is chosen for review. From the sampled literature, study characteristics (such as type of treatment, subject profiles, outcome measures, etc.) and statistical results are extracted and recorded. Effect sizes pertaining to the magnitude of change achieved on each outcome measure within each study can then be computed. Finally, these ES's are averaged thus yielding an overall number that indicates, in general, the impact of the intervention.

ADVANTAGES OF META-ANALYSIS

One of the main strengths of meta-analysis is its ability to summarize quantitatively in one common metric the research findings from a large collection of studies. With the ability to account for an enormous quantity of

related studies, a more extensive amount of essential information can be included in the review. As a result, the meta-analysis can summarize all selected research on a particular topic concisely.

Because meta-analysis is a quantitative approach to integrating study findings, it also has the ability to assess the magnitude of the treatment effect. Rather than indicating merely that a particular intervention is effective, the effect size produced from the meta-analysis establishes the specific degree to which the intervention is effective. Generally, the higher the effect size, the better the treatment. Because of this, meta-analysis is less conservative and more exact than qualitative (narrative) reviews. For example, a qualitative review would regard a positive yet nonsignificant result as failing to support the hypothesis, whereas meta-analysis enables the reviewer to quantify in a common metric the degree to which the result does support the hypothesis. In addition to providing magnitude, the effect size presents a direction; for example, some treatments may produce positive effects but others may produce negative effects. Meta-analysis permits a specific assessment of the proportion of both positive and negative effects.

Another advantage of the quantitative aspect of meta-analysis is that main effects and interactions involving variables of theoretical or empirical interest can be

evaluated. Interactions can be assessed by breaking down the overall effect size into prescribed subcategories, usually based on the research questions of interest. Relative impact can then be investigated by partialing out different variables of interest (type of treatment, nature of problem, type of measure, age of child, etc.) and examining their corresponding effect sizes. From this, research questions such as: "Under what circumstances does the treatment work best?" may be answered (Light & Pillemer, 1984, pp. 156-157). In addition, the reviewer can explore how various treatments influence a particular outcome measure or how various outcome measures impact on a particular treatment (Bryant, 1986).

Unlike the narrative review which generally organizes the studies' confounds in relation to their overall conclusions, interactions from a meta-analysis can provide insight into how methodological quality affects study results and how the two are related to a study's source (published or unpublished). Bangert-Drowns (1986) stated that "This is precisely what meta-analysis hopes to answer: are some regular patterns discernible in a body of studies on a given topic that show divergent outcomes?" (p. 388).

CRITICISMS AND CONCERNS OF META-ANALYSIS

Because it is a relatively new technique in the field of psychology, meta-analysis has not gone without criticism

nor challenge. Meta-analysis can suffer the same methodological problems as that of primary research (see Cook & Campbell, 1979), if not conducted properly (Bryant, 1986; Bryant & Wortman, 1984; Glass, McGaw & Smith, 1981; Wortman, 1983). These can be classified according to the four main types of validity: external, internal, construct, and statistical conclusion validity, which are discussed below.

External Validity. External validity refers to the extent to which the results of a study can be generalized to different populations, settings, or time periods (Cook & Campbell, 1979). Such generalizability is limited when only published studies are included in a meta-analysis. Thus, it is generally recommended that meta-analysts include published as well as unpublished studies (Glass, McGaw, & Smith, 1981; Light & Pillemer, 1984; Rosenthal, 1979; Strube & Hartmann, 1982; Sweeney, Anderson, & Bailey, 1986). Such sampling from a variety of sources may improve the external validity by enhancing generalizability and representativeness.

Furthermore, it is generally recommended that unpublished studies be included in order to avoid publication bias. Restricting the sample so that the meta-analysis includes only published studies, which seemingly manifest a higher proportion of significant results than unpublished studies, may inflate the overall ES thus

leading to erroneous conclusions.

Internal validity. One criticism relates to the internal validity of meta-analysis, that is, the degree to which one can infer a valid causal relationship between two variables (Cook & Campbell, 1979). Such a threat arises when poorly designed studies offering weak causal inferences are included in a meta-analysis. This criticism, which was primarily directed toward Smith and Glass's (1977) research synthesis on psychotherapy, came from Eysenk (1978), who opposed the endeavor of including studies of inferior design quality in the meta-analysis (purportedly what Smith and Glass had done). His reference to the axiom "garbage in-garbage out" reflects Eysenk's (1978) skepticism toward mixing flawed studies with higher quality studies and thus producing confounded results. Others share this concern as well (e.g. see Bryant & Wortman, 1984; Gallo, 1978; Kazdin & Wilson, 1978; Rachman & Wilson, 1980). However, throwing out studies on the basis of their design quality calls for making subjective judgments and thus introduces the possibility of bias. Because of this and because including such low quality studies will increase the data base, Smith, Glass, and Miller (1980) recommend including all relevant research in the meta-analysis.

In an effort to reduce these problems associated with internal validity, Strube and Hartmann (1982) proposed

that the studies entering into the meta-analysis should be weighted according to their quality. Thus, an evaluation of how design quality relates to study outcomes can be assessed. If studies of differing quality do not yield significantly different findings, then concerns about including studies of different design quality are greatly lessened. Mansfield and Busse (1977) suggested a procedure similar to Strube and Hartmann's (1982), but recommend throwing out studies of extremely low quality and weighting the remainder.

Construct validity. Another main concern about the meta-analytic procedure relates to construct validity. This refers to the degree to which the outcome measures are valid representations of the independent variables (Cook & Campbell, 1979). For example, the process of classifying a large variety of therapies into broad categories such as behavioral and nonbehavioral is analogous to lumping together "apples and oranges" (Gallo, 1978; Presby, 1978; Wortman, 1983). Presby claimed that this process ignores "important differences among the nonbehavioral therapies, for example," and thus these differences are canceled out, leading to erroneous conclusions about the different therapies (Presby, 1978, p. 514). According to Cook and Campbell (1979), in primary research "the dependent variables should not be dominated by irrelevant factors that make them measures of more or less than was intended"

(p. 61). This statement can be applied toward meta-analysis as well.

To overcome this problem, suggestions have been made to explicitly a priori specify the study's scope such as the treatments, outcome measures, subject population, and control/comparison groups of interest, and to account for the various forms of treatment separately (Bryant, 1986; Wortman, 1983). Thus, not only can effects from the "superclasses" of behavioral and nonbehavioral treatment be assessed, but also the various forms of treatment that make up the superclasses can be examined individually.

Statistical Conclusion Validity. This type of validity refers to the proper use of statistics in detecting cause and effect relationships (Cook & Campbell, 1979). One of the main concerns about the statistical conclusion validity of meta-analysis is how multiple measures within a single study should be dealt with when calculating effect sizes. A study may measure an outcome using a number of different instruments, each measuring a separate construct (such as cognitive ability, social ability, etc.), or each measuring the same construct. In addition, such instruments may be employed at multiple points in time. The general consensus is that such multiple measures are not independent and if analyses are based on ES's calculated for each outcome measure, then studies with multiple measures may contribute more to the

overall ES than studies with single measures. Thus, calculating an effect size for each outcome measure or each comparison may lead to repeated bias (Light & Pillemer, 1984).

One suggestion has been to classify the outcomes (or weight them) according to what they measure, such as cognitive ability, social ability, etc. (Strube, 1981; Strube & Hartmann, 1983). Another recommendation has been to average the effect sizes for multiple outcome measures within each study; thus the study becomes the unit of analysis (Rosenthal, 1984).

Although the meta-analytic procedure still has some challenges to overcome, it can be potentially superior to the traditional qualitative review because meta-analysis is, as stated by Fiske (1983) "more scientific and because (it) more closely approximate(s) the ideal in scientific work. As in the best of science, all steps are explicit" (p. 69). Nevertheless, meta-analysis is a new procedure in the field of psychology and thus "Its methods have not been perfected" (Fiske, 1983, p. 69). Thus, conclusions drawn from such meta-analyses are not necessarily or readily accepted by others. This brings us to the purpose of the present study.

A recent meta-analysis on the effectiveness of school-based studies of psychotherapy was conducted by Prout and DeMartino (1986). Using both "standard and

computerized search procedures" Prout and DeMartino (1986) located 33 published studies, which enabled them to calculate a total of 119 effect sizes. Based upon an average effect size of .58 from the 33 investigations, Prout and DeMartino concluded that school-based psychotherapy is "moderately effective" (p. 289). Unfortunately, several methodological problems in their review preclude accepting their major conclusion and the others they offered. These methodological complications will be examined and discussed below according to the threats to validity they pose. Following this, explicit procedures designed to overcome these problems to permit a more valid assessment of school-based psychotherapy will be provided.

METHODOLOGICAL PROBLEMS WITH PROUT AND DEMARTINO'S REVIEW

External validity. Prout and DeMartino (1986) evaluated an inadequate sample of the literature. As noted above, Prout and DeMartino (1986) based their evidence on a review of only 33 studies, whereas a careful search of the literature appears to reveal a significantly larger sample of relevant studies. Restricting a review's sample size increases the probability of making a Type II error (Cook & Campbell, 1979). Also, because it is difficult to determine if their sample of studies was a representative one, the magnitude of the resulting effect size may be a biased depiction of school-based psychotherapy studies.

Although it is difficult if not impossible to sample all existing studies for a meta-analysis (Feldman, 1971), efforts should be made to sample the literature as thoroughly as possible to increase the generalizability of findings from the meta-analysis.

A second methodological problem relating to the external validity of their study was that Prout and DeMartino (1986) did not thoroughly specify their search procedures, making replication difficult. For example, they did not identify the journals, articles, texts, and abstracts that were searched, nor did they include details about their procedures for the computerized literature search. Since a thorough literature search is an essential precondition for obtaining a representative sample of relevant studies (Arkin, Cooper, & Kolditz, 1980; Rosenthal, 1979), the procedures used should be made explicit, particularly for the purpose of replication (Fiske, 1983).

A third threat to the external validity of their meta-analysis exists because Prout and DeMartino (1986) included only published studies in their evaluation, thus increasing the probability of publication bias. Significant results are more likely to get published (Bakan, 1967; Cook & Leviton, 1980; Greenwald, 1975; Smart, 1964; Sterling, 1959). Comparing results from a number of reviews, Smith (1980) reported that published

studies produce one-third larger effect sizes than unpublished studies. As a result, resting conclusions on the results of published studies only is likely to yield an overestimation of treatment effects. Therefore, as pointed out above, it has been recommended that both published and unpublished studies be used in the meta-analysis to increase the generalizability of the findings.

Internal validity. A potential threat to the internal validity of Prout and DeMartino's study (1986) relates to their definition of treatment. Prout and DeMartino (1986) chose not to include studies which involved parent counseling or teacher consultation. However, they stated that "In some cases, studies may have included indirect interventions as concurrent or adjunct treatments" (p. 287). From this statement, it is difficult to determine if such studies were in fact included in Prout and DeMartino's (1986) review, and if so, they did not disclose any further information about the studies, such as what type of and how many "indirect interventions" were used, with what other treatments were they used in conjunction, and what effect size did they yield. Providing this information enables other reviewers to assess the effectiveness of such treatments.

Construct validity. Ambiguous methodological descriptions within Prout and DeMartino's (1986) review introduced some complications. One difficulty involves

their definition of "school-based" studies. In order to be included in Prout and DeMartino's (1986) meta-analysis, studies had to involve "direct, active intervention by a professional helper (e.g. school psychologist, counselor), and be conducted in a school or deal with a school-related problem"(p. 287). This latter phrase is problematic because it allows for a potential violation of the school-based focus of their meta-analysis. For example, since children with attention deficit disorders frequently have difficulty managing their school behavior, virtually any treatment offered to such children in any setting could be included. Prout and DeMartino's (1986) definition of "school-based" is unclear and poses a threat to the construct validity of the independent variable in their review (Cook & Campbell, 1979; Glass, McGaw, & Smith, 1981; Wortman, 1983). Because of this, it would be tenuous to attribute their results for psychotherapeutic interventions specifically to the school setting.

A second construct validity problem relates to Prout and DeMartino's (1986) application Meltzoff and Kornreich's (1970) definition of psychotherapy:

> The informed and planful application of techniques derived from established psychological principles, by persons qualified through training and experience to understand these principles and to apply these techniques with the intention of assisting

> individuals to modify such personal characteristics as feelings, values, attitudes and behaviors which are judged by the therapist to be maladaptive or maladjustive. (p. 6)

Although the above definition is an acceptable one for psychotherapy, Prout and DeMartino (1986) actually included studies of normal children in their review. That is, studies of developmental counseling and prevention-oriented programs for school children who did not manifest any maladaption or maladjustment were also evaluated. Not only did Prout and DeMartino (1986) fail to distinguish the types of target populations in the studies they reviewed, but also they did not present effect sizes separately for children with and without problems. Such confounding of target groups obscures potentially important interpretations and conclusions that may be drawn from school-based interventions. The current review included studies of both maladaptive and normal children but coded for the existence of child problems or lack of and assessed treatment effects as a function of the child's adjustment level.

Statistical conclusion validity. Prout and DeMartino (1986) computed separate effect sizes for each outcome measure for each study, thus treating them (i.e., outcome measures) independently. They computed 119 ES's from their pool of 33 studies. It is now generally recommended that

when a single study reports multiple outcome measures for a treatment, each outcome's effect be pooled; thus each study will yield only one average effect size, as discussed earlier (Light & Pillemer, 1984; Rosenthal, 1984). Prout and DeMartino were aware that their procedure creates problems and "at the judgment of the investigators" combined the data in studies with many outcome measures into only a few effect sizes. Nevertheless, they did not specify their procedures for doing this. Therefore, this approach appears unsystematic and allows for individual studies to contribute differentially to the overall results.

Secondly, Prout and DeMartino (1986) did not mention how they calculated effect sizes beyond the standard formula involving means and standard deviations. Since some studies do not always provide these basic data needed to calculate the effect size, other methods have been developed to use in such situations, for example when the study provides only the *F* statistic or the *t* statistic. Furthermore, the effect size may differ depending on the method by which it is calculated (for example, see Strube, 1981). Therefore, it is critical to report how effect sizes were calculated in different circumstances, and to assess the importance of these calculations in terms of study outcomes.

A third statistical conclusion validity limitation

refers to Prout and DeMartino's (1986) conclusion that group (average effect size = 0.63) and behavioral (0.65) therapies were more effective than individual (0.39) and nonbehavioral (0.40) interventions, respectively. Standard deviations were not reported for these categories and Prout and DeMartino apparently rested their conclusions on visual inspection of the data. In effect, they failed to capitalize on the main advantage of meta-analysis, namely that conclusions and interpretations are offered based upon statistical analyses of study results and characteristics rather than subjective judgments.

The purpose of the present review is to re-evaluate the effects of school-based psychotherapy using a representative sample of studies and following generally accepted meta-analytic procedures. In doing so, the following general questions are posed:

1. What is the overall effect of school-based psychotherapy?

2. How does effectiveness vary as a function of the theoretical nature of the treatment or its method of administration. That is, do behavioral and nonbehavioral therapies achieve similar results? Do differences appear when treatment is administered in groups to school children as opposed to individually?

3. Do design features influence therapeutic efficacy? That is, how does the overall experimental

quality of the research design (such as type of control group) and characteristics of the outcome measures (normative or non-normative) relate to outcome?

4. Does the effectiveness of school-based psychotherapy vary as a function of subject characteristics such as type of problem, age, and sex?

HYPOTHESES

In response to some of these questions, the following experimental hypotheses are offered based on the findings from meta-analytic reviews in general and the results of narrative reviews of therapy with children.

1. The overall effect of school-based psychotherapy will not be significantly different from the overall effect of psychotherapy with adults as reported in the meta-analytic review of Smith and Glass (1977). This hypothesis is corroborated by the findings of other meta-analyses of psychotherapy with children and adults, where similar overall effect sizes were obtained regardless of client age. For example, other meta-analyses of psychotherapy with children yielded overall effect sizes of 0.71 (Casey & Berman, 1985) and 0.79 (Weisz, et al., 1987). Similarly, meta-analyses of psychotherapy with adults yielded overall effect sizes of 0.93 (Shapiro & Shapiro, 1982) and 0.68 (Smith & Glass, 1977). Therefore, as stated above, it is expected that the current meta-analytic review of psychotherapy with children will yield an effect size

similar to that obtained by Smith and Glass (1977) in their meta-analytic review of psychotherapy with adults.

2. Behavioral treatment will yield significantly higher effect sizes than nonbehavioral treatment. This effect was demonstrated in other recent meta-analyses (Casey & Berman, 1985; Weisz et al., 1987), thus, this hypothesis was formulated based on such evidence in the literature.

3. Higher quality studies, that is studies which manifest eminent design quality such as random assignment to groups and normed outcome measures, will produce lower effect sizes relative to studies which manifest poor design quality. Studies which may be biased due to improper selection practices, use of nonnormative outcome measures, and attrition have been shown to produce larger estimates of outcome than those studies which employ random assignment to groups, normed outcome measures, and little or no attrition (Foulds, 1958; Mansfield & Busse, 1977; Wortman, 1983; Wortman & Bryant, 1985). This hypothesis was formulated based on such information.

4. Treatment effects will vary as a function of the child's adjustment level. The ordering of effects sizes from highest to lowest is expected to be: children with moderate to severe problems, children with mild problems, and normal children. No specific predictions are offered with respect to children whose problems are of an unknown

clinical nature. The rationale behind this hypothesis is that diagnostic measures are pathologically oriented, thus only serving those whom exhibit some degree of pathology. Furthermore, changes in normal children, such as self-esteem, are harder to achieve than changes in children with mild problems, such as anxiety/phobias. Therefore, it is presumed that children exhibiting moderate problems will be more susceptible to change, and to a larger degree of change, than those judged as having mild or no problems.

5. It is expected that treatment effects involving comparisons to attention-placebo controls will yield significantly lower effects sizes than comparisons which are made to no-treatment controls. The placebo control group's main purpose is to ascertain whether or not the attention received or expectations assumed by the subjects significantly contribute to the subjects' improvement. Studies employing both placebo and no treatment controls have found that when the treatment group is compared to the former the rate of improvement is smaller than comparisons to the no treatment controls (Landman & Dawes, 1982; Smith & Glass, 1977). This effect is expected in the current review as well.

6. Unpublished studies will yield significantly lower effect sizes than published studies. This hypothesis is supported by evidence that published studies produce larger estimates of outcome than unpublished studies

(Glass, McGaw, & Smith 1981; Lane & Dunlap, 1978; Smart, 1964). As a result, it is surmised that publication policies are biased toward studies with significant findings, and thus published studies will inflate the actual effect size.

METHOD

There are several important differences between this review and Prout and DeMartino's (1986) regarding the review process and analytic procedures. Whenever these occur, they are noted.

QUALIFICATIONS AND CRITERIA

Studies qualified for this review if they meet each of the following five criteria.

1. Year of study's completion. Studies completed through the years 1962 to 1982 inclusive were reviewed.

2. Treatment vs. control comparison. Studies qualified if they contained at least one experimental group that was compared to at least one control group.

3. Who administered the treatment. Studies qualified if the treatment was administered by mental health professionals (those with a Ph.D. or M.A. in psychology, M.D. in psychiatry, social workers, and school guidance counselors) or professional trainees (graduate students in psychology, interns, practicum students, and psychiatric residents). Also, since Prout and DeMartino (1986) included studies in which treatments were conducted by one of the authors, and those in which the only description of the change agent was "experimenter," this

review included such studies as well.[1] Presumably, these treatments were conducted by professionals. However, studies involving parent counseling or teacher consultation were excluded. Only studies involving direct treatment to children (that is, therapy administered specifically by the mental health professional to the child, as opposed to indirect therapy involving parents and teachers appropriately trained in mental health skills) were included.

4. Treatment context. This review included studies directed at modifying children's school adjustment. Prout and DeMartino (1986) included studies if they were conducted in the school or if the studies dealt with "a school-related problem." To be more explicit, the present review only included studies conducted in school settings.

5. Nature of the problem. Prout and DeMartino (1986) included studies of normal and maladapting children, developmental counseling interventions, and prevention-oriented programs. Such studies were included in this review as well, but the child's adjustment level was coded and analyzed in relation to outcome.

6. Age/Grade. The grade levels, which were not specified in Prout and DeMartino's (1986) review, include preschool, elementary school, junior high, and high school.

LITERATURE SEARCH

Published studies were located by manually searching the contents of 14 journals in which school psychotherapy studies seemed most likely to appear (see Appendix A). Also examined were several research reviews of the child therapy literature apparently overlooked by Prout and DeMartino (1986) and two recent meta-analyses covering child therapy (Casey & Berman, 1985; Weisz, Weiss, Alicke, & Klotz, 1987). Finally, reference lists of other identified studies were also examined. This search procedure was very tedious but was considered necessary in order to uncover an adequate sample of the published literature.

To evaluate the possibility of publication bias, a representative sample of unpublished doctoral dissertations was searched, both manually and by computer. A manual search of all Dissertation Abstracts from 1962 to 1982 was considered too impractical. Therefore, the following procedure was used to obtain a representative sample of dissertations. First, a computer search of Dissertation Abstracts was conducted using 43 search terms. From this, a large number of potentially relevant studies was identified. To determine the accuracy of this search, the computer-generated citations were checked against a sample of Dissertation Abstracts that was searched manually. This was done by randomly selecting one year from each of three

decades covering the review period, resulting in an investigation of the years 1967, 1972, and 1982. For only one of these years (1972) the entire abstracts was searched. A random sample of ten issues of the Abstracts from each of the other two years was selected and examined. At this point, a comparison of the relevant studies found by the computer and manual searches was made, revealing the proportion of computer citations that were inappropriate (false positives). Generally, this occurred because the cited studies did not include therapy outcome studies, empirical assessments of outcome, or involved within subjects designs. The manual search uncovered three times as many relevant unpublished dissertations than the computer-generated search (false negatives). These findings remained generally consistent across the three years of the manual search. Thus, based on the number of false negatives and false positives produced by the computer search, an estimate of 300 relevant unpublished dissertations appeared during the years 1962 to 1982.

Because the computer search provided a fair approximation of both the distribution and the total number of relevant dissertations during the review period, a sample of unpublished dissertations was obtained in the following manner. An initial 15% random sample of dissertations was drawn from the original computer list, along with an additional 10% random sample of replacement

studies. The manner in which the replacement studies were used will be introduced following the discussion below concerning the procedures for obtaining dissertations.

Dissertations for review were secured as follows. The review period was stratified into four-year intervals (1982-1979, 1978-1975, 1974-1971, etc.) and studies were randomly selected accordingly. First, one of the randomly chosen studies was inspected in Dissertation Abstracts to ensure its relevance and was included if appropriate. Following this, two additional studies were selected by manually surveying entries on adjacent pages of Dissertation Abstracts, until two additional relevant studies were found. If the initial randomly selected study was judged inappropriate and/or irrelevant, three additional entries on adjacent pages were surveyed.

The selected dissertations were then obtained through interlibrary loan. As each study was received, it was inspected further to verify whether it met the qualifications for inclusion. Occasionally studies were judged nonuseable, either because the dissertation eventually became published, the effects of the treatment were qualitatively rather than quantitatively assessed, the empirical data were reported in such a way that no effect sizes could be calculated, or the design did not include a usable control group, However, the primary reason for eliminating studies from the initial sample was because

some institutions did not offer their dissertations for interlibrary loan. It was in anticipation of this practice that the 10% replacement sample was developed so that all nonuseable dissertations resulting from the first sample could be replaced with dissertations taken from the second sample. Replacement and initial studies were matched according to year of completion. The list of both the published and unpublished studies selected for review appears in Appendix A.

CODING OF STUDIES

Each study was coded on 44 variables, which were divided into eight major categories (The coding scheme appears in Appendix B). The categories include (1) basic identifying data (year of publication, source of study - published or unpublished); (2) design features (type of design, assignment to groups, sample size); characteristics of (3) the subject populations, (4) the therapists, (5) comparison groups, (6) treatments and (7) outcome measures; and finally, (8) how effect sizes were calculated.

Characteristics of the subject population were assessed by coding children's presenting problems. It was considered necessary to assess subject characteristics and degree of problem severity in order to determine whether treatment impact varies as a function of these variables. For example, it may be found that boys respond better to a

particular treatment than girls, or that children with clinical problems respond more/less favorably to a particular treatment than children with mild problems. To assess this information, a two-step process was required. First, problems were coded according to their general seriousness: (1) none; (2) mild; (3) those of uncertain nature or degree; and, (4) those reflecting moderate to severe school maladjustment (see Appendix B). Second, the problems of children falling into one of the latter three categories were further classified into one of 13 categories reflecting the continuum of internalizing and externalizing symptomology (see # 19 in Appendix B).

CALCULATION OF EFFECT SIZES

Whereas Prout and DeMartino (1986) used the standard deviation of the control group to calculate ES, this review used the pooled standard deviation in the denominator (Cohen, 1977). Thus, the following formula was applied for each study supplying the appropriate information:

$$\frac{Mt - Mc}{\sqrt{\dfrac{(Nt - 1) \times SDt^2 + (Nc - 1) \times SDc^2}{Nt + Nc - 2}}}$$

where Mt is the mean of the treatment group, Mc is the mean

of the control group, Nt and Nc, and SDt and SDc are the sample sizes and standard deviation of the treatment and control groups, respectively. Prout and DeMartino (1986) provided no details on the calculation of ES's when means and SD's were not available. When such data were unavailable for this review, estimates of effect sizes were computed following the procedures described by Glass, McGaw and Smith (1981, chapter 5).

Two exceptions to these procedures are made. The first exception pertains to studies which include more than two groups, provide the group means and provide information from the ANOVA summary table (specifically the mean square F values between groups). Given this situation, ES's were calculated using Holmes' (1986) equation # 27 because of the ease of calculation.

The second exception includes Wortman and Bryant's (1985) adjusted effect size, to be used when treatment and control groups are nonequivalent. This equation corrects for an otherwise biased estimate of the treatment's impact by calculating two effect sizes, one for posttreatment and one for pretreatment scores. The corrected effect size is then computed by subtracting the pretreatment from the posttreatment ES. This formula was applied when the experimental and control groups differed at pretreatment. For example, this formula was used when a treatment group was comparatively inferior to controls at posttreatment,

but nevertheless made substantial progress compared to controls given pretreatment comparisons. If a finding for an outcome measure was not reported, or if it was described as nonsignificant and no further information was provided, the effect size for that measure was estimated as zero.

Overall, twelve different methods for calculating or estimating effect sizes were used, depending on the information provided in the study. Each method was coded accordingly because, as Strube and Hartmann (1983) noted, "the results of a meta-analysis may vary depending on the specific techniques used (to calculate an ES)" (p. 21). Prout and DeMartino (1986) computed 119 ES's from 33 studies, resulting in an average of three to four separate effect sizes per study. Thus, it appears that they treated each outcome measure within a study as independent when in fact they were dependent. Because Prout and DeMartino (1986) did not state their procedures for managing multiple outcome measures, the studies with such measures may have contributed disproportionately to their overall effect size. In an effort to deal with this issue, this review provided two overall effect sizes, one based on each outcome measure as the unit of analysis and one based on the study as the unit of analysis. This latter procedure corrects for nonindepenence within studies and is recommended by other meta-analysts as well (Light & Pillemer, 1984; Landman & Dawes, 1982). Furthermore,

guided by the research hypotheses, this review pooled measures within a study comparison that were conceptually related. For example, when examining the relationship between the type of control group and effect size, comparisons within each study were pooled within each type of control group. Such pooling occurred only when measures within a comparison (e.g. attention-placebo control) were conceptually similar, and the methods by which the ES's were calculated were the same. Thus, identically coded measures within a comparison were combined and the effects sizes were averaged. This procedure is further explicated in the subsequent section.

RESULTS

OVERALL EFFECT SIZE

A total of 212 studies of school-based psychotherapy were analyzed, which produced 753 posttest effect sizes. Thus, there were an average of 3.75 comparisons per study (similar to Prout & DeMartino's 3.6). The overall average effect size of school-based psychotherapy produced from these data was .47 with a standard deviation of .70 (different from Prout & DeMartino's ES of .58). Based on this, it can be stated that the average person in the experimental group is better off than 68% of those in the control group. Furthermore, effect sizes ranged from -1.69 to +5.40 with negative ES's comprising only 10.4% of the total sample. Table 1 provides the effect sizes grouped into increments of .10 and their respective frequencies.

The generally recommended procedure of calculating one effect size per study was completed next. Using this procedure, a total of 206 effect sizes (one per study) were produced yielding an average overall effect size of .58 with a standard deviation of .72. (Six of the 212 studies were strictly follow-up studies with no posttest information and thus were excluded.) Based on these data, it can be stated that the average person in the treatment group is better off than 72% of those in the control group.

The ES's from the above two procedures are significantly different, $t(957) = 1.96$, $p < .05$, two-tailed.

The distribution of the ES's indicated a large frequency of zero ES's ($N = 172$) as well as a small, but obvious proportion of outliers (see Table 1). A large percentage of zero Es's was expected primarily because, while coding the studies, it was found that several authors reported nonsignificant findings but did not provide any other statistical data with which to calculate an ES. As a result, the ES was estimated as zero. Thus, a conservative estimate of the treatment effect was used when such information was reported.

Because the largest ES was +5.40 it was necessary to examine this as well as other outliers. Outliers were defined as cases which fell beyond 3.5 standard deviations of the mean. Thus, any ES of 3.00 or greater fell into this category, resulting in a total of seven outliers. Each study yielding the outlier was carefully examined for unusual features relating to subjects, treatments, analyses, or outcome measures. Since no peculiar features in these areas were discovered that might contribute to the comparatively large ES's, all seven outliers were included in all analyses.

However, because these outliers increase the variances for each analysis, and because the presence of a large proportion of zeros affects the normality of the

Table 1

Distribution of Effect Sizes

Frequency	ES	Frequency	ES
1	-1.64	16	1.16
1	-1.14	10	1.26
1	- .94	11	1.36
4	- .84	16	1.46
2	- .74	5	1.56
1	- .64	9	1.66
3	- .54	2	1.76
5	- .44	7	1.86
12	- .34	5	1.96
13	- .24	5	2.16
14	- .14	2	2.26
193	- .04	3	2.36
48	.06	1	2.46
45	.16	1	2.56
51	.26	3	2.66
44	.36	1	2.76
43	.46	1	2.86
29	.56	1	3.06
34	.66	2	3.36
29	.76	1	3.66
31	.86	1	4.16
22	.96	1	5.06
22	1.06	1	5.36

distribution, some concerns arose regarding the results of the data analyses. With these points in mind, nonparametric tests (such as the Mann-Whitney U and the Kruskall Wallis) were applied to the data as well as parametric tests and a comparison was made between these two methods. However, this comparison revealed no significant differences in the resulting outcomes. Thus the results from parametric tests are reported here.

Furthermore, in addition to the initial parametric analyses, alternative analyses were conducted by using the same procedures but excluding the seven outliers as well as omitting all the zero ES's. Whenever results from these alternative analyses produced significantly different findings from the original analyses, the data are provided.

An overview of some of the characteristics of the reviewed studies is presented in Table 2. The average age of the child receiving school-based psychotherapy within this meta-analysis was 10.45 years (SD = 3.13), with a mode of 10.5 and a range of 3.8 to 18.0 years of age. Based on this age variable, four overall grade level categories were constructed, indicating that the majority of the sample (67%) was of elementary school age. The breakdowns for race could only be partially determined since 73% of the total sample of studies did not report this information. Nevertheless, the data from those who did provide this information are reported in Table 2. In terms of

Table 2

Sample Characteristics and Demographics

	N	%	$\bar{X}$	Mode	Range	SD
Age	205	----	10.45	10.5	3.8 - 18	3.13
Sample N[a]	211	----	93.5	24.0	8 - 1675	219.2
Length of Trmt	222	----	39.6	30.0	5 - 350	33.4
N of Trmt Sessions	222	----	15.4	10.0	1 - 160	19.5
Grade						
Preschool	8	3.8				
Elementary	142	67.0				
Jr. High	26	12.3				
High School	36	17.0				
Race						
Major./All White	27	12.7				
Major./All Minor.	18	8.5				
Mixed	13	6.1				
Unknown	154	72.7				
Problem Severity						
None	32	15.1				
Mild	31	14.6				
Moderate/severe	63	29.7				
Uncertain	86	40.6				

[a]Eliminating 6 studies with an unusually large sample size (N > 915) reduces the mean to 62.32 and the SD to 59.61.

children's presenting problems, 15.1% of the studies used subjects with no problems, 14.6% had children with problems that were mild in nature, and 29.7% were diagnosed with moderate to severe problems. The remaining 40.6% of the studies provided insufficient information so that the children's problems were coded as uncertain or unknown in nature.

Other sample descriptions include size of treatment group, length of treatment and number of treatment sessions. The size of the treatment group contained an average of 93 subjects with a mode of 24. The average treatment session lasted 39.65 minutes, with a mode of 30 and a range of 5 to 350 minutes, while the mean number of treatment sessions was 15.42 with a mode of 10 and a vast range from 1 to 160 sessions (see Table 2). In sum, the average person is this sample was 10.75 years old and participated in approximately 15 treatment sessions, each lasting about 40 minutes. With this general overview of the sample characteristics in mind, the next sections evaluate each of the experimental hypotheses.

SCHOOL-BASED VS. ADULT PSCYHOTHERAPY

The first hypothesis to be tested refers to the null hypothesis of no difference between the mean effect of school-based psychotherapy and psychotherapy with adults (primarily) as described by Smith and Glass (1977). This result is reported in Table 3. The headings in the table

refer to the form of the school-based data. The unaggregated form is essentially the data in their original, most basic form, with 784 posttest outcome measures, each with a corresponding effect size. Smith and Glass's data were of this nature as well, evidenced by their inclusion of 375 studies which yielded 833 effect sizes. A comparison of the unaggregated current data to that of Smith and Glass's indicated that the overall average effect size of .47 from the 753 effect sizes of school-based psychotherapy was significantly different from and smaller than Smith and Glass's average of .68 from their 833 effect sizes of adult psychotherapy, $t(1584) = 6.18$, $p < .01$, two-tailed.

The aggregated school-based data were also used to test this hypothesis, as reported in Table 3, principally because it has been recommended that the data be pooled in such a way. The aggregated data contain an average of the multiple ES's within each study, resulting in one mean ES per study. Thus, the comparison of the aggregated school-based mean ES of .58 produced from 206 ES's was not significantly different from Smith and Glass's adult psychotherapy mean of .68 produced from 833 effect sizes, $t(1037) = 1.82$, n.s. The mean from the aggregated data therefore supports the null hypothesis of no difference between school-based and adult psychotherapy.

These disparate findings primarily result from the

Table 3

School-based Psychotherapy ES VS. Smith and Glass's (1977) Adult Psychotherapy ES

Unaggregated Data

	ES	SD	n	
School-based	.47	.70	753	$t = 6.18^*$
Adult	.68	.67	833	

* $p < .01$

Aggregated Data

	ES	SD	n	
School-based	.58	.72	206	$t = 1.82$ n.s.
Adult	.66	.67	833	

different forms of the school-based data. It is difficult to verify which group of data accounts for a more worthy test of the hypothesis. Although the aggregated data set ensures independent ES's, the form of the unaggregated data is statistically comparable to Smith and Glass's (1977). Landman and Dawes (1982), however, did perform a similar meta-analysis on a subsample (N = 42) of Smith and Glass's data, treating the data in aggregated form (the study was the unit of analysis). As a result, the aggregated ES resulting from Landman and Dawes' re-analysis could be compared to the aggregated ES of this meta-analysis. The aggregated effect size in their study was .90, which is significantly different from this study's aggregated effect size of .58, $t(247) = 22.22$, $p < .001$. Clearly, Landman and Dawes' ES of .90 is considerably larger than Smith and Glass's ES of .68. This difference may be due in part to aggregating the data or to the lack of dissertations in Landman and Dawes' subsample; dissertations were not included because they were too difficult to obtain. Thus, even though a similar comparison could be made between the present study's aggregated ES and Smith and Glass's aggregated ES through Landman and Dawes' re-analysis, unpublished studies were left out which may have increased the ES and thus provided for an inaccurate comparison with the present school-based ES.

Because a comparable match of the present data to

Smith and Glass's cannot be made, I am more inclined to place priority on the first test of this hypothesis. Thus, the data indicate a difference between school-based and adult psychotherapy, with school-based psychotherapy producing a significantly lower ES than Smith and Glass's ES of adult psychotherapy.

The remaining hypotheses do not include comparisons between these data and other's data, such as Smith and Glass's. Therefore, the data will be aggregated across the variables of interest since this procedure increases the independence of the resulting ES's. This will be explicated further as each test for each hypothesis is presented.

ATTENTION PLACEBO VS. NO-TREATMENT CONTROLS

The second hypothesis to be tested referred to finding significant differences among the control groups employed. In particular, it was hypothesized that the attention-placebo control group would yield a significantly larger effect size compared to the no-treatment control group. The data were analyzed in aggregated form so that for each study the effect sizes were averaged across each type of control group. For example, if one study utilized three outcome measures and two types of controls such as no treatment and attention placebo, a total of six effect sizes would initially be calculated. However, when the data are aggregated, a total of two effect sizes would

result, one for each type of control group. The total number of control group effect sizes produced from the data analyzed in aggregated form was 232. Thus, more than one type of control group was utilized within 30 studies. The mean's and N's per group are provided in Table 4.

A one-way ANOVA was used to test this hypothesis which resulted in nonsignificant differences among the three control groups, $F(2, 229) = .80$, n.s. Therefore, the type of control group used did not appear to affect the ES obtained.

BEHAVIORAL VS. NONBEHAVIORAL TREATMENT

It was hypothesized that behavioral treatment would yield significantly higher ES's than nonbehavioral treatment. Some of the behavioral treatment applied in this meta-analysis included modeling, rational-emotive therapy, self-instruction training, desensitization, cognitive-behavior therapy and covert reinforcement, whereas nonbehavioral treatment included client-centered therapy, values clarification, transactional analysis, and affective counseling. These two major categories of treatment (behavioral and nonbehavioral) were aggregated across the four control groups since type of control group did not influence the results. Therefore, if a study applying behavioral treatment contained four outcome measures and utilized two types of control groups, initially producing a total of eight ES's, it would be

Table 4

Mean Effect Sizes, Standard Deviations, and Sample Sizes as a Function of Type of Control Group

Type of Control Group	M	SD	n
No Treatment Control	.61	.78	159
Attn Placebo Control	.48	.51	63
Waiting-List Control	.60	.48	10

Note. $F = .80$, n.s.

reduced to one average ES corresponding specifically to the behavioral treatment. Similarly, if a study applied both behavioral and nonbehavioral treatment, four outcome measures and two types of controls, the data would be aggregated to produce two average effect sizes (from 16), each ES corresponding to the type of treatment.

Using the aggregated data set, there were a total of 104 studies of behavioral therapy and a total of 114 studies of nonbehavioral therapy. Given that only 206 studies were used and that the total N for type of treatment was 218, 12 studies applied both treatments. The average ES for behavioral treatment was .85 with a standard deviation of .81, while nonbehavioral treatment produced an average ES of .31 with a standard deviation of .54. A t-test was performed on these means, revealing a significant difference between the two therapies in support of the hypothesis, $t(216) = 5.88$, $p < .001$, one-tailed.

Since the value of behavioral treatment is almost three times greater than nonbehavioral treatment, the remaining hypotheses were examined by taking the dual classification of treatment into account. That is, two-way analyses of variance were used instead of t-tests or one-way ANOVA's that would normally be required to test each remaining hypothesis. Since results had already indicated that there was a significant effect for type of treatment, interest in the subsequent ANOVA's focused on the possible

interactions between treatment and other variables.

PUBLISHED VS. UNPUBLISHED STUDIES

It was hypothesized that published studies would produce significantly higher values than unpublished studies. A two-way ANOVA was used to test this hypothesis, with treatment (behavioral vs. nonbehavioral) as one independent variable and source (published vs. unpublished) as the second independent variable. Studies from books were eliminated from this analysis because there were only six studies of this category. Main effects were significant for both type of treatment and source of study, $F(208) = 30.23$, $p < .001$, and $F(208) = 5.05$, $p < .03$, respectively. But, the analysis of variance indicated a nonsignificant interaction, $F(1, 208) = .87$, $p < .35$. Cell means for type of treatment and source of study are provided in Table 5. Looking at the treatment means, an obvious difference exists, with behavioral treatment yielding a larger value than nonbehavioral, as predicted. An inspection of the source means indicates a large difference between the two variables, with published studies generating a significantly larger ES ($M = .66$) than unpublished studies ($M = .36$). In conclusion, the results support the hypothesis that source of study makes a difference, with published studies producing significantly higher ES's than unpublished studies.

Table 5

Mean Effect Sizes, Standard Deviations, and Sample Sizes for Type of Treatment by Source of study

	Treatment					
	Behavioral			Nonbehavioral		
Source	*M*	*SD*	*n*	*M*	*SD*	*n*
Published	.94	.85	82	.37	.37	78
Unpublished	.57	.54	21	.21	.18	31

Note. Main effect for source, $F(1,208) = 5.05$, $p < .03$.
Main effect for treatment, $F(1,208) = 30.23$, $p < .001$.

TYPE OF PRESENTING PROBLEM

The fifth hypothesis stated that children with moderate to severe problems will produce higher effect sizes than children with mild problems, while, in turn, children with mild problems will produce larger effect sizes than children with no problems. Before testing this hypothesis, type of presenting problem was aggregated within type of treatment, resulting in a total N of 218.

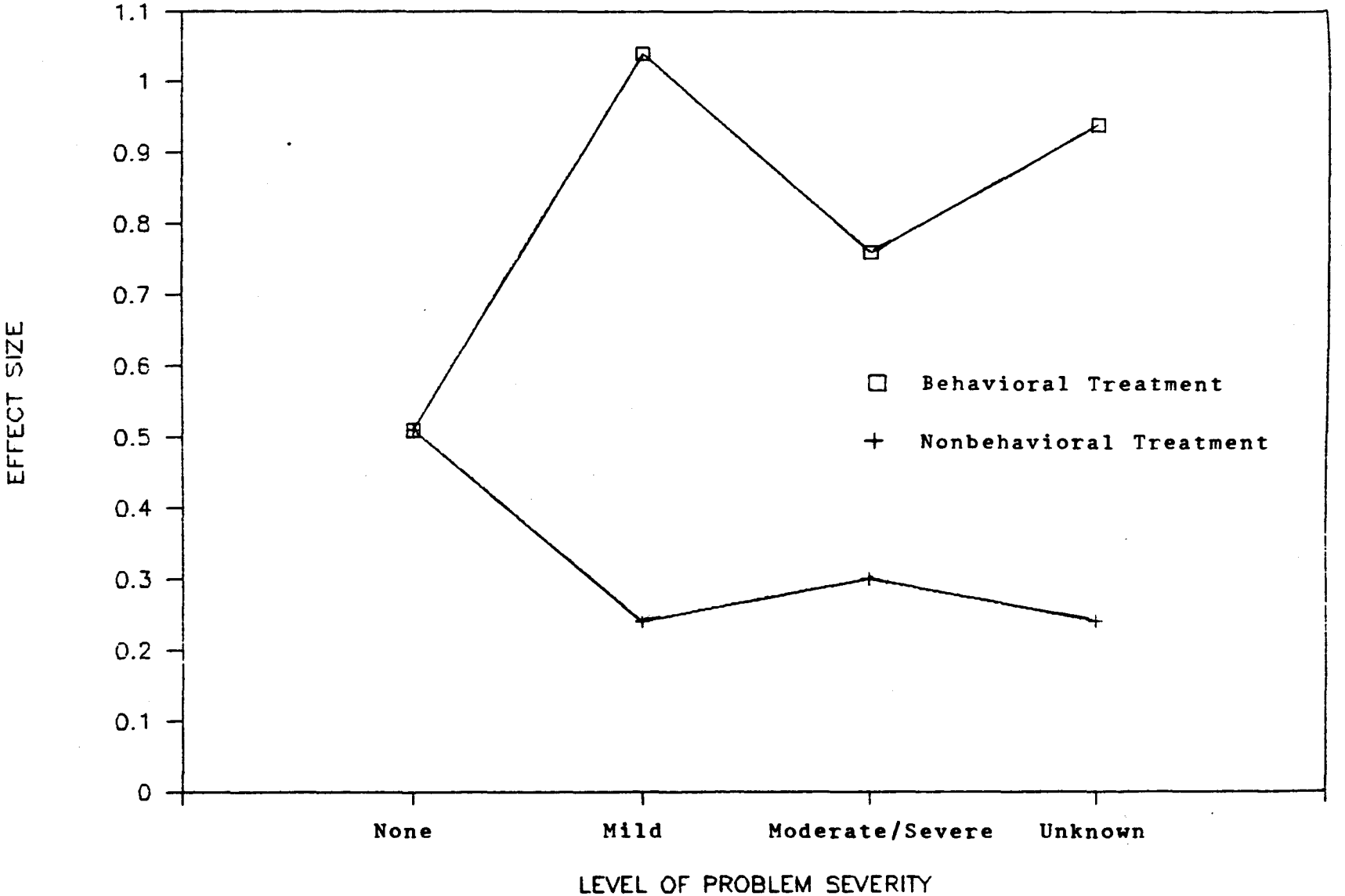
To test this hypothesis a two-way ANOVA was performed, with type of treatment as one independent variable and presenting problem as the other independent variable. A main effect was found for the type of treatment but not for the levels of problem severity, $F(1, 210) = 33.87$, $p < .001$, $F(3, 210) = .28$, n.s., respectively. This analysis revealed an interaction between the two variables which approaches significance, $F(3, 210) = 2.53$, $p < .06$. The cell means, SD's and N's are provided in Table 6 and presented graphically in Figure 1. Since the interaction revealed a trend, a simple main effects analysis was performed which revealed a significant difference occurring between behavioral and nonbehavorial treatment for subjects with mild problems, moderate to severe problems, and for subjects whose problems were of an unknown nature, p's < .003. It appears that children in these three categories benefit significantly more from behavioral treatment than from nonbehavioral treatment,

Table 6

Mean Effect Sizes, Standard Deviations, and Sample Sizes for Type of Treatment by Level of Problem Severity

Children's Level of Problem Severity	Treatment					
	Behavioral			Nonbehavioral		
	M	SD	n	M	SD	n
With Outliers and Zero Effect Sizes						
None	.51	.68	13	.51	1.06	22
Mild	1.04_a	.71	19	$.24_b$	.28	14
Moderate/Severe	$.76_a$	.79	30	$.30_b$	.35	33
Unknown	$.94_a$	.88	42	$.24_b$	.29	45
Without Outliers and Zero Effect Sizes						
None	.67	.70	10	.41	.33	18
Mild	.97	.42	18	.31	.28	12
Moderate/Severe	.87	.59	27	.32	.33	31
Unknown	.89	.63	38	.32	.37	40

Note. Row means with different subscripts are significantly different at the .003 probability level based upon a test for simple main effects.


1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
EFFECT SIZE
Behavioral Treatment
Nonbehavioral Treatment
None
Mild
Moderate/Severe
Unknown
LEVEL OF PROBLEM SEVERITY

whereas type of treatment does not make a difference for children who have no problems.

As noted in the bottom half of Table 6, the results differed when data from outliers and zero ES's were excluded. In the latter procedure, the interaction did not approach significance and there was no significant main effect for the type of problem, $F(3, 186) = .92$, $p < .43$; $F(3, 186) = .94$, $p < .94$, respectively.

QUALITY OF STUDY

It was hypothesized that those studies which manifested better design features would yield significantly lower effect sizes than those studies which demonstrate design features of poorer quality. Design variables included degree of attrition, group assignment procedure, and how subjects were selected for the study (source of Ss). Initially, a multiple regression analysis (MRA) was performed on these factors using $SPSS_X$. To do this, attrition was calculated into percentages while the variables within the latter two factors (group assignment procedure and source of subjects) were dummy coded. The results revealed, however, that the correlations between each quality variable and effect size were nonsignifcant. An inspection of the mean ES's for each of these variables (see Table 7) reveals that a large majority of the studies within this meta-analysis correspond to those features which constitute good study quality. Thus the range of

Table 7

Mean Effect Sizes, Standard Deviations, and Sample Sizes for the Four Variables of Study Quality.

	M	SD	n
Group Assignment			
Matching	.38	.43	12
Random	.62	.76	158
Available, Intact	.53	.60	18
Voluntary	.26	.30	6
Other	.28	.30	3
% of Attrition			
0 - 5%	.62	.77	158
6 - 10%	.45	.48	20
11 - 15%	.24	.20	13
16 - 20%	.73	.67	7
21% or more	.45	.44	7
Source of Subjects			
Problem-Oriented Msrmt.	.60	.72	126
Volunteers	.52	.49	22
Mixed/Other	.53	.80	49

quality was not evenly distributed; 86% of the studies used random assignment and matching procedures, 77-87% indicated 0-10% attrition, and 64% of the studies obtained their subjects based on problem-oriented measurements.

Since the degree of relationship between study quality and ES could not be ascertainable through the MRA due to the large percentage of studies manifesting appropriate design quality, a 2 x 2 (group assignment procedure x source of subjects) ANOVA was performed to determine if the differences among the effect sizes within each factor were significant and if they interacted. Results of this anaylsis revealed nonsignificant main effects for group assignment and source of subjects, $F(4, 185) = .67$, $F(2, 185) = .12$, respectively, as well as a nonsignificant interaction, $F(5, 185) = .48$. Thus the mean ES's corresponding to each group assignment procedure and source by which subjects were selected do not differ significantly from each other.

EXPLORATORY ANALYSES

Other analyses were performed as a means of investigating some post hoc questions for which no specific hypotheses were offered. Since a meta-analysis provides the researcher with a large supply of variables which can be tested in a number of different ways, exploring the data beyond the stated hypotheses was considered useful. Furthermore, such exploratory analyses enable other

researchers to profit more from the meta-analysis by shedding light on other variables which may be contributing to or supporting the results found for each of the above research questions.

For example, the previous data indicated behavioral treatment was superior to nonbehavioral treatment. Reasons for this superiority may lie in the measures used to assess the impact of such treatment. Behavioral treatment is often measured by some observable assessment such as number of truancies, on/off task behavior, verbal/physical aggression within the classroom, etc. Since these are observable events, they are often easier to measure (and perhaps easier to change) than those at the other end of the spectrum (e.g., unobservable events, such as self-esteem). As a result, type of measurement was explored in terms of how it affected the results of behavioral and nonbehavioral treatment.

The nine categories of outcome measures (listed in Appendix B, #33) were aggregated within each type of treatment so that if a behavioral treatment in one study, for example, was assessed by two achievement tests and three different independent behavioral observation measures, there would be two effect sizes for that study, each representing the aforementioned tools by which the behavioral treatment was assessed.

An inspection of the aggregated data revealed a small

N (3) for the nonindependent behavioral outcome measure, which generated a relatively large ES of .71. Because the cell size was so low, it was eliminated from the data analyses.

A two-way (treatment x type of outcome measure) ANOVA was performed, revealing a marginally significant main effect for outcome measure, $F(7, 387) = 1.84$, $p < .08$, and a marginally significant interaction, $F(7, 387) = 1.83$, $p < .08$. To examine the source(s) of this trend, a simple main effects was applied on these data. This analysis revealied that behavioral observations, normed rating scales (which includes behavioral checklists), and nonnormed rating scales, produced significantly higher ES's for behavioral treatment than for nonbehavioral treatment p's < .003. Trends emerged in the same direction (behavioral > nonbehavioral treatment) for achievement tests, $p < .07$, and objective performance measures $p < .10$. Table 8 reveals the cell means for each of these variables.

The next exploratory probe considered whether an outcome measure's degree of specificity influences the ES's corresponding to each type of treatment. For example, do general or specific measures of treatment impact produce larger effect sizes? To test this possibility, the level of specificity (general/specific) was aggregated within each type of treatment and a two way (treatment x level of specificity) ANOVA was performed. There were no

Table 8

Mean Effect Sizes, Standard Deviations, and Sample Sizes for Type of Treatment by Type of Outcome Measure

	Treatment					
	Behavioral			NonBehavioral		
Outcome Measure	*M*	*SD*	*n*	*M*	*SD*	*n*
Indep behav observation	1.02	1.01	34	.13	.22	13
Peer sociometric	.31	.54	9	.14	.38	18
Normed rating scale	.60	.79	35	.21	.26	44
Non-normed rating scale	.63	.67	51	.26	.34	72
Achievement test	.60	.71	14	.21	.44	16
Other performance measure	.66	.59	28	.84	1.83	8
School grades	.67	.93	12	.31	.52	28
Objective performance msr	1.29	1.73	8	.38	.62	13

Note. Two-way ANOVA revealed a nonsignificant interaction, $F(7, 387) = 1.83$, $p < .08$, and nonsignificant main effect for outcome measure, $F(7, 387) = 1.84$, $p < .08$.

significant main effects but a significant two-way interaction was obtained, $F(1, 315) = 4.40$, $p < .04$. Inspection of the cell means (see Table 9) revealed that specific measures (e.g., behavioral observations) of the impact of behavioral treatment produce higher effect sizes than general measures of the same treatment (e.g., GPA). This variable did not make a difference, however, for the nonbehavioral treatment.

Different results occurred when outliers and zeros were eliminated. This analysis revealed significant main effects for both the level of specificity, $F(1, 267) = 4.28$, $p < .001$, and type of treatment, $F(1, 267) = 63.70$, $p < .04$, but no significant interaction. These means and cell sizes are provided in the bottom half of Table 9.

Another post-hoc interest which prompted exploratory analysis involved an inspection of how the source of outcome measure influences the ES for each type of treatment. Again, the data were aggregated within each treatment and the cell means were inspected. Eight cells (behavioral and nonbehavioral treatments crossed with the source categories of parents, therapist, mixed and other) contained n's of less than four and, consequently, were eliminated. Following this, a 2-way ANOVA was applied revealing a marginally significant interaction between these two variables, $F(4, 343) = 2.16$, $p < .07$. To probe this trend, a simple main effects analysis was applied,

Table 9

Mean Effect Sizes, Standard Deviations and Sample Sizes for Type of Treatment by the Measure's Level of Specificity

Level of Measure's Specificity	Treatment: Behavioral M	SD	n	Treatment: NonBehavioral M	SD	n
With Outliers and Zero Effect Sizes						
Specific	$.91_a$	.80	101	$.34_b$	.57	105
General	.47	.62	44	.21	.34	69
Without Outliers and Zero Effect Sizes						
Specific	.91	.57	90	.36	.33	88
General	.68	.75	35	.29	.41	57

Note. Analyses with outliers and zero ES's produced a signficant interaction, $F(1, 315) = 4.40$, $p < .04$. Row means with different subscripts indicate a significant difference at the .04 probability level. Analyses without outliers and zero ES's did not produce a significant interaction but obtained significant main effects for both treatment and source, $F(1, 267) = 63.70$, $p < .001$, $F(1, 267) = 4.28$, $p < .04$, respectively.

revealing differences between the two types of treatment when the sources of measurement were independent observers, teachers or school, or the subject (self-report). Each of these three sources of measurement yielded significantly higher effects sizes for behavioral treatment than for nonbehavioral treatment, *p*'s < .005 (see Table 10).

Again, a difference was found between this analysis and the alternative analysis which excluded outliers and zero effect sizes. The latter analysis revealed only a significant main effect for the type of treatment $F(1, 269) = 53.44$, $p < .001$. A nonsignificant main effect was found for the source of the outcome measure, $F(4, 269) = .91$, and a nonsignificant interaction was indicated as well, $F(4, 269) = .77$, n.s. The effect sizes are provided in Table 10.

Finally, the last area of speculation concerned differences in the method by which effect sizes were calculated. To probe this question, it was necessary to eliminate the method by which ES's were estimated as zero, since this method obviously produced ES's different than the other procedures. Furthermore, although twelve methods were provided, nine were ultimately utilized. The method of calculating ES's based on correlations was never encountered, while change scores and raw data (both of which provided *M*'s and *SD*'s) elicited low frequencies and thus, for statistical purposes, were grouped under the

Table 10

Mean Effect Sizes, Standard Deviations, and Sample Sizes for Type of Treatment by Source of Outcome Measure

	Treatment					
	Behavioral			Nonbehavioral		
Source of Measure	*M*	*SD*	*n*	*M*	*SD*	*n*
With Outliers and Zero Effect Sizes						
Independent Observers	1.02_a	1.01	34	$.14_b$	.23	13
Teachers/school	$.66_a$	1.03	36	$.18_b$	.31	56
Peers	$.36_a$	.53	10	$.18_b$	.41	17
Subject/self-report	.63	.58	47	.26	.33	74
Performance measure	.63	.59	43	.50	1.06	23
Without Outliers and Zero Effect Sizes						
Independent Observers	.89	.62	27	.24	.25	8
Teachers/school	.82	.88	23	.29	.40	41
Peers	.41	.54	9	.24	.46	14
Subject/self-report	.76	.55	40	.32	.34	62
Performance measure	.79	.56	36	.28	.30	19

Note. Row means with different subscripts are significantly different at the .005 probability level based upon a test for simple main effects.

method utilizing means and standard deviations (see the coding scheme in Appendix B).

A one-way ANOVA was applied to the remaining unaggregated data (N = 576), revealing significant differences among the methods, $F(7, 575) = 3.57$, $p < .0009$. The Scheffe' Multiple Range Test was applied to these data which indicated a significant difference between only two methods. Interestingly, this difference was found between the method using posttest means and standard deviations and the method which utilized the pretest, posttest correction, with the former method producing larger ES's than the latter method (see Table 11 for M's and SD's).

Table 11

The Method by Which ES's Were Calculated

Method of Calculation	M	SD	n
Means & standard deviations[a]	.70	.93	219
ANOVA summary table	.54	.35	27
t-score	.72	.67	92
ANCOVA	.28	.25	12
Probit analysis	.74	.63	30
Estimate from p value	.53	.51	53
Holmes' method	.40	.55	32
Pretest-Posttest adjustment[a]	.36	.57	111

[a]Based on post-hoc analyses, these two methods are significantly different from each other, $p < .05$.

DISCUSSION

The aim of the present meta-analysis of school-based psychotherapy was to refine and improve upon Prout and DeMartino's (1985) recent research synthesis of the same topic. These improvements enhanced the validity of the present meta-analysis accordingly (Bryant, 1986):

1. External validity - Compared to Prout and DeMartino, this meta-analysis included almost five times as many published studies (33 vs. 160), was composed of 25% unpublished studies (n = 52), and specified the exact scope of the review by restricting the setting to the school exclusively.

2. Construct validity - The characteristics of the sampled studies were defined and provided.

3. Statistical conclusion validity - Compared to Prout and DeMartino's meta-analysis, this research synthesis treated the data statistically in a dual manner - with multiple effect sizes per study as well as with one average effect size per study. Furthermore, the method by which each effect size was calculated was coded. Finally, the primary data analyses were guided by specific hypotheses, providing a more focused review.

Overall, such procedures resulted in a more representative sample of the school-based literature as well as a more extensive exploration of the data from the sampled studies.

A primary outcome of this meta-analysis was that the overall ES, .48, was similar to Prout and DeMartino's. However, what sets this finding apart from Prout and DeMartino's is the fact that unpublished studies comprised 25% of the total sample of this meta-analysis and such studies were subsequently found to produce significantly lower ES's than published. If strictly published studies were included in this review, the overall Effect size would have been .66. Other findings would also seem to be affected by the exclusion of unpublished studies, which may have inflated the ES attributable to a particular variable. Based on this as well as other evidence of similar differences in ES between published and unpublished studies (Duzinski, 1987; Smith, 1980), it seems imperative to include a sample of unpublished studies within the meta-analysis.

Another main finding from this review was that behavioral treatment produced significantly larger ES's than nonbehavioral treatment. This difference remained consistent across all analyses. Exploratory analyses indicated how this occurred. For example, the nature of the outcome measure by which the treatment was assessed may

have contributed to the difference between behavioral and nonbehavioral therapy, as evidenced by the marginally significant interaction of these two variables ($p < .08$), and particularly by significant and marginally significant simple main effects. This analysis indicated that behavioral treatment produced significantly larger ES's when its impact was assessed by independent behavioral observations, normed rating scales/behavior checklists, and non-normed measurements; marginally significant differences were found with objective performance measures and achievement tests. In turn, nonbehavioral treatment produced larger ES's (though not significant) when its impact was assessed by other (cognitive and behavioral) performance measures.

In light of these findings, two interpretations are possible. Either the nature of behavioral problems predisposes them to better means of therapy relative to nonbehavioral problems (perhaps because behavioral problems and treatment are more readily observable and accountable events); or perhaps the measures used in assessing behavioral problems provide a more accurate or specific assessment of the problem of interest than measures used in assessing nonbehavioral problems.

One finding which offers partial support for the latter interpretation is that specific, behavioral measures produced significantly larger ES's in comparison to

specific, nonbehavioral measures. More evidence, however, was found when outliers and zero effect sizes were eliminated. This resulted in main effects for both variables so that specific measures of treatment impact yielded significantly higher ES's than general measures. Thus, in explaining the superiority of behavioral treatment, perhaps behavioral measures are more specific assessments of such treatment and, as a result yield larger ES's than nonbehavioral measures.

Behavioral treatment was also found to be more successful than nonbehavioral treatment for maladapting children, having its greatest impact (and nonbehavioral treatment having its lowest impact) on children whose problems were mild in nature. In contrast, children who had no problems benefited equally well from either treatment. Children in this latter category received treatment primarily for preventive purposes; that is, to facilitate or advance their present, normal conditions as a means of avoiding the development of any problems. Thus, when children do not have any apparent psychological problems, the probability of successful school-based therapy does not favor one particular type of treatment over another. In contrast, behavioral intervention appears to be the treatment of choice for maladapting children, at least when treatment is offered in the schools. However, it must be noted that when outliers and zero ES's were

eliminated behavioral treatment was significantly more beneficial for all levels of children's problem severity, including those who had no problems.

Based on this evidence, it does not appear as if children with severe problems are less susceptible to change as a result of school-based therapy than those with less severe problems, or (in some cases) than those who have no problems. Furthermore, behavioral treatment was beneficial across all levels of problem severity (with the possible exception of the group with no problems).

One surprising result was the failure to find any significant differences in ES as a function of the type of control group used in the study. Based on previous data (Duzinski, 1987; Landman & Dawes, 1982; Smith & Glass, 1977) it was presumed that nonspecific attention given to children would be beneficial. Perhaps the attention provided to control groups within school settings is not as strong as those provided within other settings. For example, the ease of communication within a school (due to the proximity of the students/subjects) may influence the amount of information control subjects learn about the treatment groups. For instance, compensatory rivalry may result when the no-treatment control subjects learn about the desirable therapy that the treatment group is receiving, and thus try to compete with the treatment groups' gains. Similarly, resentful demoralization may

result when the attention-placebo control group learns that they are not receiving the same treatment as those in the experimental group. Thus, the attention-placebo control group may try to retaliate by acting out more, giving up on difficult academic tasks, etc. These two "threats to internal validity" (Cook & Campbell, 1979) offer possible explanations regarding the nonsignificant differences among the control groups.

An evaluation of study quality was difficult to obtain since most of the studies included in this meta-analysis manifested appropriate design features, such as matching and/or random assignment to groups, 0-10% attrition, and obtaining subjects through problem-oriented measurements. As a result, the degree of relationship between design quality and effect size could not be discernable from a multiple regression analysis. Futhermore, differences in effect size appeared large among the group assignment procedures and the sources by which subjects were acquired, but such differences were nonsignificant, nor did these two variables produce an interaction.

In sum, although design quality could not be probed to the extent desired, such an effort allowed for an inspection of the nature of the studies which constitute this meta-analysis as well as those which constitute the school-based psychotherapy literature in general. The

majority of studies which contributed to the overall ES of school-based psychotherapy reflected appropriate design features, which enhanced the accuracy by which the true effect of school-based psychotherapy could be discerned. Thus, it is unfortunate that the full range of design quality could not be examined here, but it is equally fortunate that this meta-analysis is composed mainly of credible studies.

Finally, an examination was made of the various methods by which effect sizes were calculated, to determine whether or not the method influenced the ES value. The results revealed a significant difference between the method which used posttest M's and SD's and the method which appropriately corrected for or adjusted posttest M's and SD's with pretest values. The former method yielded significantly larger ES's than the latter method. It appears that the correction method is more conservative in the sense that it cancels out any gains or recognizes any losses the subjects might have manifested at pretest in comparison to the controls. The correction method thus yields the ES corresponding specifically to the true impact of the treatment. Based on this result, it appears that the pretest correction method is a necessary procedure which should be utilized in future meta-analyses, particularly when pretest data indicate a difference between the treatment and control groups.

LIMITATIONS

One of the primary concerns of this meta-analysis is the large number of zero ES's. While such a distribution is common among meta-analyses (Wolf, 1986), observance was still made regarding the robustness of such results. For the most part, major findings were unaffected by the large quantity of zeros and outliers, as evidenced by analyses excluding such variables. Essentially, such results demonstrate the robustness of obtained findings. Nevertheless, the zero ES's remain a focus of concern. It is suggested that future studies report sufficient data (M's and SD's) of measures yielding nonsignificant results, so that meta-analysts can obtain a more accurate account of effect sizes for nonsignificant outcomes to avoid a large frequency of zero effect sizes.

Another limitation of this meta-analysis regards the broad treatment classifications of behavioral and nonbehavioral. Such "superclasses," according to Presby (1978) "ignore important differences among the behavioral and nonbehavioral therapies, for example, the superior effects of rational-emotive therapy (RET) as compared to the others in that class" (p. 514). The point made here is that some particular forms of treatments within each group may have been more (or less) effective than others. That is, one subtype of treatment may be carrying most of the ES weight. Furthermore, delineating the various behavioral

and nonbehavioral treatments and their ES's would indicate the relative effectiveness of each treatment and thus assist practitioners in deciding which specific treatment to use when faced with a choice.

In hindsight, it appears that many of the treatments which made up these two categories in this review were not frequent and similar enough to develop such subcategories; the range and variety of treatments appear to be very large. Furthermore, in some cases, the authors did not thoroughly describe the treatment process, making potential classification difficult or very subjective. Thus, developing subcategories within the behavioral and nonbehavioral "superclasses" may have resulted in a wide range of treatments, and possibly a large number classified as ambiguous.

Nevertheless, it would be useful to discover if a few similar therapies consistently provide higher ES's within each treatment group. Providing such a breakdown of the two therapies within the school-based literature would require, prior to coding the studies, a thorough explication of rules and criteria which could be used in objectively determining the subgroup in which each treatment could fall.

In conclusion, school-based psychotherapy is considered moderately effective overall, with the average child in the treatment group benefiting more from the

therapy than 68% of those in the control group. The ES found in this review was identical to Prout and DeMartino's, however, the scope of this review was not the same and thus a direct comparison cannot be made. It was also found that behavioral treatment consistently exhibited more successful outcomes than nonbehavioral treatment. These results may be due in part to the type of assessment measures used, the nature of the child's problem (behavioral problems may be relatively easier to measure and change), or to the efficacy of the treatment itself. Thus, is it the nature of the measure, the problem, the treatment, or all three? This is a difficult question to answer, but it would be of great use to identify each variable's contribution. Such information cannot be answered based on the data in this meta-analysis, but we do know that overall behavioral treatment was more effective than nonbehavioral treatment, particularly when assessed by objective performance measures and independent behavioral observations. Hopefully, findings from this meta-analysis will help subsequent efforts attempting to uncover further variables which interact with or effect the efficacy of school-based psychotherapy.

REFERENCES

Achenbach, T. A. (1982). Developmental psychopathology (2nd ed.). New York: John Wiley and Sons, Inc.

Arkin, R., Cooper, H., & Kolditz, T. (1980). A statistical review of the literature concerning the self-serving bias in interpersonal influence situations. Journal of Personality, 48, 435-448.

Bakan, D. (1967). On method. San Francisco: Jossey-Bass.

Bangert-Drowns, R. L. (1986). Review of developments in meta-analytic method. Psychological Bulletin, 99(3), 388- 399.

Bryant, F. B. (1986). Improving the quality of research synthesis in program evaluation. Policy Studies Review, 5(4), 709-721.

Bryant, F. B. & Wortman, P. M. (1984). Methodological issues in the meta-analysis of quasi-experiments. In W. H. Yeaton & P. M. Wortman (Eds.),New directions in program evaluation: Vol. 24. Issues in data synthesis. San Francisco: Jossey-Bass.

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.D., & Campbell, D.T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Chicago, IL: Rand McNally.

Cook, T., & Leviton, L. Reviewing the literature: A comparison of traditional methods with meta-analysis. Journal of Personality, 48, 449-471.

Durlak, J. A. (1983). Providing mental health services to elementary school children. In C. E. Walker & M. C. Roberts (Eds.), Handbook of clinical child psychology. New York: Wiley.

Duzinski, G. A. (1987). The educational utility of cognitive behavior modification strategies with children: A quantitative synthesis. Dissertation Abstracts International, 48(339-A).

Eysenck, H. J. (1978). An exercise in mega-silliness. American Psychologist, 33, 517.

Feldman, K.A. (1971). Using the work of others: Some observations on reviewing and integrating. Sociology of Education, 44, 86-102.

Fiske, D.W. (1983). The meta-analytic revolution in outcome research. Journal of Consulting and Clinical Psychology, 51(1), p. 65-70.

Foulds, G. (1958). Clinical research in psychiatry. Journal of Mental Science, 104, 259-265.

Gallo, P.S., Jr. (1978). Meta-analysis-A mixed metaphor? American Psychologist, 33, 515-517.

Glass, G. V. & Kliegl, R. M. (1983). An apology for research integration in the study of psychotherapy. Journal of Consulting and Clinical Psychology, 51(1), 28-41.

Glass, G. V., McGaw, B., and Smith, M. L. (1981). Meta-analysis in Social Research. Beverly Hills, Calif.: Sage.

Greenwald, A.G. (1975). Consequences of prejudice against the null hypothesis. Psychological Bulletin, 82, 1-20.

Gresham, F. M. (1985). Utility of cognitive-behavioral procedures for social skills training with children: A critical review. Journal of Abnormal Child Psychology, 13(3), 411-423.

Gresham, F. M. & Lemanek, K. L. (1983). Social skills: A review of cognitive-behavioral training procedures with children. Journal of Applied Developmental Psychology, 4, 239-261.

Henry, S. E., & Kilmann, P. R. (1979). Student counseling groups in senior high school settings: An evaluation of outcome. Journal of School Psychology, 17(1), 27-46.

Hobbs, S. A., Moguin, L. E., Tyroler, M., & Lahey, B. B. (1980). Cognitive behavior therapy with children:

Has clinical utility been demonstrated? Psychological Bulletin, 87(1), 147-165.

Holmes, C. T. (1985). Effect size estimation in meta-analysis. Journal of Experimental Education, 52, 106-109

Kazdi n, A. E., & Wilson, G. T. (1978). Evaluation of behavior therapy: Issues, evidence and research strategies. Cambridge, Mass: Ballinger.

Kratochwill, T. R., Feld, J. K., Van Someren, K. R. (1986). The effectiveness of school psychological services. In S. N. Elliott and J. C. Witt (Eds.), The delivery of psychological services in schools (pp. 249-303). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Landman, J. T., & Dawes, R. M. (1982). Psychotherapy outcome: Smith and Glass' conclusions stand up under scrutiny. American Psychologist, 37, 504-516.

Lane, D.M., & Dunlap, W.B. (1978). Estimating effect size: Bias resulting from the significance criterion in editorial decisions. British Journal of Mathematical and Statistical Psychology, 31, 107-112.

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

Mansfield, R. S., Busse, T.V. (1977). Meta-analysis of research: A rejoinder to Glass. Educational Research, 6, 3.

Meltzoff, J., & Kornreich, M. (1970). Research in psychotherapy. New York: Atherton.

Presby, S. (1978). Overly broad categories obscure important differences between therapies. American Psychologist, 33, 514.

Prout, H. T. & DeMartino, R. A. (1986). A meta-analysis of school-based studies of psychotherapy. Journal of School Psychology, 24, 285-292.

Prout, H. T. & Harvey, J. B. (1978). Applications of desensitization procedures for school-related problems: A review. Psychology in the Schools, 15(4), 533-541.

Rachman, S.J., & Wilson, G.T. (1980). The effects of psychological therapy: Second enlarged edition. New York: Pergamon Press.

Rosenbaum, M. S. & Drabman, R. S. (1979). Self-control training in the classroom: A review and critique. Journal of Applied Behavioral Analysis, 12, 467-485.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. Psychological Bulletin, 86, 638-641.

Rosenthal, R. (1984). Meta-analytic procedures for social research. Beverly Hills, Calif.: Sage Publications.

Shapiro, D.A., & Shapiro, D. (1982). Meta-analysis of comparative therapy outcome studies: A replication and refinement. Psychological Bulletin, 92(3), 581-

604.

Shaw, M. C. & Wursten, R. (1965). Research on group procedures in schools: A review of the literature. Personnel and Guidance Journal, 44, 27-34.

Smart, R.G. (1964). The importance of negative results in psychological research. The Canadian Psychologist, 5, 225-232.

Smith, M.L. (1980). Publication bias and meta-analysis. Evaluation in Education, 4, 22-24.

Smith, M. L. & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. American Psychologist, 32, 752-760.

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

Sterling, T.D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance - or vice versa. Journal of the American Statistical Association, 54, 30-34.

Strube, M. J. (1981). Meta-analysis and cross-cultural comparison: Sex differences in child competitiveness. Journal of Cross-Cultural Psychology, 12, 3-20.

Strube, M. J., & Hartmann, D. P. (1982). A critical appraisal of meta-analysis. British Journal of Clinical Psychology, 21, 129-139.

Strube, M. J., & Hartmann, D. P. (1983). Meta-analysis:

Techniques, Applications, and Functions. Journal of Consulting and Clinical Psychology, 51(1), 14-27.

Sweeney, P.D., Anderson, K., & Bailey, S. (1986). Attributional style in depression: A meta-analytic review. Journal of Personality and Social Psychology, 50(5), 974- 991.

Weiner, I.B. (1982). Child and adolescent psychopathology. New York: Wiley.

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.

Wolf, F. M. (1986). Meta-analysis: Quantitative methods for research synthesis. Beverly Hills, CA: Sage Publications.

Wortman, P. M. (1983). Evaluation research: A methodological perspective. Annual Review of Psychology, 34, 223-260.

Wortman, P. M., and Bryant, F. B. (1985). School desegregation and black achievement. Sociological Methods and Research, 13(3), 289-324.

ENDNOTE

1 Although this was not explicitly stated in Prout and DeMartino's (1985) review, an investigation of the studies used in their meta-analysis indicated they had included studies which identified the therapist as "experimenter."

APPENDIX A

APPENDIX A

STUDIES INCLUDED IN THIS META-ANALYSIS

Abbott, A.H. (1978). Individual counseling with high-risk students: A practical approach. The School Counselor, 25, 206-208.

Albondy, N.V. (1975). The effects of experiential awareness, group-centered counseling, and teacher led group on the sociometric status of children in grade six. (Doctoral dissertation, The University of New Mexico, 1973). Dissertation Abstracts International, 35(7-A), 4150A.

Allen, R.P., Safer, D.J., Heaton, R., Ward, A., & Barrell, M. (1975). Behavior therapy for socially ineffective children. Journal of the American Academy of Child Psychiatry, 14, 500-510.

Alper, T.G., & Kranzler, G.D. (1970). A comparison of the effectiveness of behavioral and client-centered approaches for the behavior problems of elementary school children. Elementary School Guidance and Counseling, 5, 35.

Altmann, H.A., & Firnesz, K.M. (1973). A roleplaying approach to influencing behavioral change and self-esteem. Elementary School Guidance and Counseling, 7, 276-281.

Amerikaner, M., & Summerlin, M.L. (1982). Group counseling with learning disabled children: Effects of social skills and relaxation training on self-concept and classroom behavior. Journal of Learning Disabilities, 15, 340-343.

Anderson, E.C. (1967). Counseling and consultation versus teacher-consultation in the elementary school. Elementary School Guidance and Counseling, 2, 276-285.

Andrews, W.R. (1971). Behavioral and client-centered counseling of high school underachievers. Journal of Counseling Psychology, 18, 93-96.

Arbuckle, D.S., & Boy, A.V. (1961). Client-centered therapy in counseling students with behavior problems. Journal of Counseling Psychology, 8, 136-139.

Argulewicz, E.N. (1982). Effects of an instructional program designed to improve attending behaviors of learning disabled students. Journal of Learning Disabilities, 15, 23-27.

Arnold, S.C., & Forehand, R. (1978). A comparison of cognitive training and response cost procedures in modifying cognitive styles of impulsive children. Cognitive Therapy and Research, 2, 183-187.

Baker, B.L. (1973). Camp Freedom: Behavior modification for retarded children in a therapeutic camp setting.

American Journal of Orthopsychiatry, 43, 418-427

Bandura, A., Grusec, J.E., & Menlove, F.L. (1967). Vicarious extinction of avoidance behavior. Journal of Personality and Social Psychology, 5, 16-23.

Bandura, A., & Menlove, F.L. (1968). Factors determining vicarious extinction of avoidance behavior through symbolic modeling. Journal of Personality and Social Psychology, 8, 99-108.

Barabasz, A.F. (1973). Group desensitization of test anxiety in elementary school. The Journal of Psychology, 83, 295-301.

Barcai, A., Umbarger, C., Pierce, T.W., Chamberlain, P. (1973). A comparison of three group approaches to under-achieving children. American Journal of Orthopsychiatry, 43, 133-141.

Barclay, J.R. (1967). Effecting behavior change in the elementary classroom: An exploratory study. Journal of Counseling Psychology, 14, 240-247.

*Barke-Stein, J.A. (1976). A study of self concept in the elementary school using different methods of group counseling. (Doctoral dissertation, Brigham Young University, 1976). Dissertation Abstracts International, 37(2-A), 798A.

Barrows, R.S. (1971). Using audiotape playback in secondary school counseling. The School Counselor,

19, 115.

Bates, M. (1968). A test of group counseling. Personnel and Guidance Journal, 46, 749-753.

Bell, S., & Ledford, T. (1978). The effects of sociodrama on the behaviors and attitudes of elementary school boys. Group Psychotherapy, Psychodrama, and Sociometry, 31, 117-135.

Benson, R.L. & Blocher, D.H. (1967). Evaluation of developmental counseling with groups of low achievers in a high school setting. The School Counselor, 14, 215-220.

Bishop, I.H. (1977). Comparative effects of cognitive and affective group counseling on self-esteem of second grade children. Dissertation Abstracts International, 38(7-A),3948A. (University Microfilms, 77-28, 036)

Blackwood, R.O. (1970). The operant conditioning of verbally mediated self-control in the classroom. Journal of School Psychology, 8, 251-257.

Bleck, R.T., & Bleck, B.L. (1982). The disruptive child's playgroup. Elementary School Guidance and Counseling, 17, 137- 141.

Block, J. (1978). Effects of a rational-emotive mental health program on poorly achieving, disruptive high school students. Journal of Counseling Psychology, 25, 61-65.

*Blohm, A.A. (1978). Group counseling with moderately mentally retarded and learning disabled elementary school children. Dissertation Abstracts International, 39(6-A), 3362. (University Microfilms International No. 78-24-128)

Blythe, B.J., Gilchrist, L.D., Schinke, S.P. (1981). Pregnancy-prevention groups for adolescents. Social Work, 26, 503-504.

Brown, R.T. (1980). Impulsivity and psychoeducational intervention in hyperactive children. Journal of Learning Disabilities , 13, 19-24.

Brown, R.T., & Conrad, K.J. (1982). Impulse control or selective attention: Remedial programs for hyperactivity. Psychology in the Schools, 19, 92-97.

Burdsal, C., & Buel, C.L. (1980). A short term community based early stage intervention program for behavior problem youth. Journal of Clinical Psychology, 36, 226-241.

Calsyn, R.J., Quicke, J., & Harris, S. (1980). Do improved communication skills lead to increased self-esteem? Elementary School Guidance and Counseling, 15, 48-55.

Camp, W.L. (1967). Follow-up of a classroom program for potential high school dropouts. Psychology in the Schools, 4, 346-351.

Cangelosi, A., Gressard, C.F., & Mines, R.A. (1980). The

effects of a rational thinking group on self-concepts in adolescents. The School Counselor, 27, 357-361.

Carlisle, J.F. (1977). Effects of a symbolic modeling procedure on seventh-grade socially withdrawn children. Dissertation Abstracts International, 37(11-B), 5825B. (Xerox University Microfilms No. 77-11, 097)

Chennault, M. (1977). Improving the social acceptance of unpopular educable mentally retarded pupils in special classes. American Journal of Mental Deficiency, 72, 455-458.

Clark-Stedman, M., & Wolleat, P.L. (1979). A non-sexist group-counseling intervention: Moving toward androgyny. The School Counselor, 27, 110-118.

Clements, B.E. (1966). Transitional adolescents, anxiety, and group counseling. Personnel and Guidance Journal, 45, 67-71.

Coats, K.I. (1979). Cognitive self-instructional training approach for reducing disruptive behavior of young children. Psychological Reports, 44, 127-134.

Cobb, H.C., & Richards, H.C. (1983). Efficacy of counseling services in decreasing behavior problems of elementary school children. Elementary School Guidance and Counseling, 17, 180-187.

Cohen, N.J., Sullivan, J., Minde, K., Novak, C., & Helwig,

C. (1981). Evaluation of the relative effectivenesss of methylphenidate and cognitive behavior modification in the treatment of kindergarten-aged hyperactive children. Journal of Abnormal Child Psychology, 9, 43-54.

Cole, P.M., & Hartley, D.G. (1978). The effects of reinforcement and strategy training on impulsive responding. Child Development, 49, 381-384.

Cotugno, A.J. (1982). Cognitive control methods in the treatment of special education children. Psychology in the Schools, 19, 517-525.

Cradock, C., Cotler, S. & Jason, L.A. (1978). Primary prevention: Immunization of children for speech anxiety. Cognitive Therapy & Research, 2, 389-396.

Crow, M.L. (1969). A comparison of three group counseling techniques with sixth graders. Elementary School Guidance and Counseling, 6, 37-41.

Cullinan, D., Epstein, M.H., & Silver, L. (1977). Modification of impulsive tempo in learning-disabled pupils. Journal of Abnormal Child Psychology, 5, 437-444.

Deffenbacher, J.L., & Kemper, C.C. (1974). Counseling test-anxious sixth graders. Elementary School Guidance and Counseling, 9, 22-29.

DeLuca, F. (1976). The effects of group counseling as

influenced by group size on junior high school students in academic difficulty. Dissertation Abstracts International, 36(8-A), 5037-5038A. (Xerox University Microfilms No. 76-2979)

Digiuseppe, R., & Kassinove, H. (1976). Effects of a rational-emotive school mental health program on children's emotional adjustment. Journal of Community Psychology, 4, 382-387.

Dil, N., & Gotts, E.E. (1971). Improvement of arithmetic self concept through combined positive reinforcement, peer interaction, and sequential curriculum. Journal of School Psychology, 9, 462-471.

Downing, C.J. (1977). Teaching children behavior change techniques. Elementary School Guidance and Counseling, 11, 277-283.

Drowne, J.L. (1972). A study of three group counseling approaches and their effectiveness in modifying selected aspects of self-concept and selected personality characteristics of third grade children. Dissertation Abstracts International, 32(8-A), 4344. (Xerox University Microfilms No. 72-5625)

Drummond, D.J. (1975). Self-instructional training: An approach to disruptive classroom behavior. Dissertation Abstracts International, 35(8-B), 4167-4168B. (Xerox University Microfilms No. 75-3869)

Duke, D.B. (1982). The effects of relaxation training and cognitive behavior modification on locus of control and field independence-dependence. (Doctoral dissertation, California School of Professional Psychology, 1982). Dissertation Abstracts International, 43(4-B), 1250.

Eastman, B.G., & Rasbury, W.C. (1981). Cognitive self-instruction for the control of impulsive classroom behavior: Ensuring the treatment package. Journal of Abnormal Child Psychology, 9, 381-387.

Evers-Pasquale, W.L. (1978). The peer preference test as a measure of reward value: Item-analysis, cross validation, concurrent validation and replication. Journal of Abnormal Child Psychology, 6, 175-188.

Evers-Pasquale, W., & Sherman, M. (1975). The reward value of peers: A variable influencing the efficacy of filmed modeling in modifying social isolation in preschoolers. Journal of Abnormal Child Psychology, 3, 179-189.

Fink, A.M. (1973). Application of social learning principles to a high school psychology of adjustment curriculum: Effects of extra-classroom behavior. (Doctoral dissertation, University of Minnesota, 1973). Dissertation Abstracts International, 34(5-B), 2302.

Finney, B.C., & Van Dalsem, E. (1969). Group counseling for gifted underachieving high school students. Journal of Counseling Psychology, 16, 87-94.

Flowers, J., & Martson, A. (1972). Modification of low self-confidence in elementary school children. Journal of Educational Research, 66, 30-35.

Forman, S.G. (1980). A comparison of cognitive training and response cost procedures in modifying aggressive behavior of elementary school children. Behavior Therapy, 11, 594-600

Franz, W. K., Berning, L.W., & Reilly, E.M. (1976). The effects of sensory awareness training on interpersonal social distance in fourth graders. Psychology in the Schools, 13, 58-63.

Frazier, F., Matthes, W.A. (1975). Parent education: A comparison of adlerian and behavioral approaches. Elementary School Guidance and Counseling, 10, 31-38.

Friedling, C., & O'Leary, S.G. (1979). Effects of self-instructional training on second- and third-grade hyperactive children: A failure to replicate. Journal of Applied Behavior Analysis, 12, 211-219.

Fullerton, S. (1973). Self-concept changes of junior high students. Journal of Counseling Psychology, 20, 493-494.

Garwood, C.J.S. (1964). The development and utilization of

a group approach to counseling ninth graders. Dissertation Abstracts, 25(3), 1740. (University Microfilms No. 64-4804)

Geoghagen, J.L. (1970). An action approach to group counseling: An experimental study. (Doctoral dissertation, University of Tennessee, 1970). Dissertation Abstracts International, 31(9-A), 4458.

Gerler, E.R. (1980). A longitudinal study of multimodal approaches to small group psychological education. The School Counselor, 27, 184-189.

Glenwick, D.S., & Barocas, R. (1979). Training impulsive children in verbal self-control by use of natural change agents. Journal of Special Education, 13, 387-398.

Gottman, J. (1977). The effects of a modeling film on social isolation in preschool children: A methodological investigation. Journal of Abnormal Child Psychology, 5, 69-79.

Gourley, M.H. (1971). The effects of individual counseling, group guidance, and verbal reinforcement on the academic progress of underachievers. Dissertation Abstracts International, 31(8-A), 3873A. (University Microfilms No. 71-8559)

Graybill, D. & Gabel, H. (1982). A school-based counseling program for parents of learning disabled children

using paraprofessional co-counselors. Psychological Reports, 50, 729-730.

Greenleaf, D.O. (1982). The use of structured learning therapy and transfer programming with disruptive adolescents in a school setting. Journal of School Psychology, 20, 122-130.

Greenwell, C.C. (1980). The effects of script counseling on fifth and seventh grade students. (Doctoral dissertation, Texas Tech University, 1980). Dissertation Abstracts International, 41(4-A), 1403A.

Gresham, F.M., & Nagle, R.J. (1980). Social skills training with children: Responsiveness to modeling and coaching as a function of peer orientation. Journal of Consulting and Clinical Psychology, 48, 718-729.

Guerney Jr., B., Coufal, J., & Vogelsong, E. (1981). Relationship enhancement versus a traditional approach to therapeutic/preventative/enrichment parent-adolescent programs. Journal of Consulting and Clinical Psychology, 49, 927-939.

Hall, D.L. (1980). The use of symbolic vicarious reinforcement in facilitating social interaction in preadolescent children with low rates of interaction. (Doctoral dissertation, University of Virginia, 1978). Dissertation Abstracts International, 40(8-A), 4489A.

Halliwell, J.W., Musella, D.F., & Silvino, P.J. (1970). Effects of counseling on attitudes and grades with intermediate grade pupils designated as having poor attitudes. Elementary School Guidance and Counseling, 5, 113-123.

Hansen, J.C., Niland, T.M., & Zani, L.P. (1969). Model reinforcement in group counseling with elementary school children. Personnel and Guidance Journal, 47, 741-744.

Hardage, N.C. (1972). A comparison of the efficacy of treatments of classroom behavior management and group counseling for use with potential dropouts. Dissertation Abstracts International, 33(4-A), 1436. (Xerox University Microfilms No. 72-26, 551)

Harris, M.B., & Trujillo, A.E. (1975). Improving study habits of junior high school students through self-management versus group discussion. Journal of Counseling Psychology, 22, 515-517.

Harris, S.R. (1976). Rational-emotive education and the human development program: A guidance study. Elementary School Guidance and Counseling, 11, 113-122.

Hayes, E.J., Cunningham, G.K., & Robinson, J.B. (1977). Counseling focus: Are parents necessary? Elementary School Guidance and Counseling, 12, 8-14.

Heller, B. & Gurney, D. (1968). Involving parents in group counseling with junior high underachievers. The School Counselor, 15, 394-397.

Henthorn, J.W. (1980). An experimental study of a group counseling approach to improve the perceived self-concept and cognitive skills of Title I primary school children. (Doctoral dissertation, United States International University, 1980). Dissertation Abstracts International, 41(4-A), 1404.

Hervey, E.P. (1971). Comparison of three and six weeks of group model-reinforcement counseling for improving study habits and attitudes of junior high school students. (Doctoral dissertation, University of Wisconsin, 1970). Dissertation Abstracts International, 31(11-A), 5765A.

Hillman, B.W., & Runion, K.B. (1978). Activity group guidance: Process and results. Elementary School Guidance and Counseling, 13, 104-111.

Hinds, W.C., & Roehlke, H.J. (1970). A learning theory approach to group counseling with elementary school children. Journal of Counseling Psychology, 17, 49-55.

Hinze, W.H.A. (1971). The effects of group play activity upon the reading ability and anxiety of seven and eight year old boys with low reading achievement.

Dissertation Abstracts International, 31(7-A), 3269A. (Xerox University Microfilms No. 70-25,204)

Hoyser, E.E. (1971). Therapeutic non-directive play with low achievers in reading. (Doctoral dissertation, Oregon State University, 1971). Dissertation Abstracts International, 31(8-A), 3874A.

Jakibchuk, Z, & Smeriglio, V.L. (1976). The influence of symbolic modeling on the social behavior of preschool children with low levels of social responsiveness. Child Development, 47, 838-841.

*Jasnow, M. (1982). Effects of relaxation training and rational emotive therapy on anxiety reduction in sixth grade children. (Doctoral dissertation, Hoftra University, 1982). Dissertation Abstracts International, 43(12-B), 4149.

Jason, L., & Burrows, B. (1983). Transition training for high school seniors. Cognitive Therapy and Research, 7, 79-92.

Jensen, J.L. (1978). Eliminating self-defeating behavior group counseling: Its effects on self-concept and locus of control of economically disadvantaged high school students. (Doctoral dissertation, Brigham Young University, 1978). Dissertation Abstracts International, 39, 3375B.

Johnson, T., Tyler, Jr., V., Thompson, R., & Jones, E.

(1971). Systematic desensitization and assertive training in the treatment of speech anxiety in middle-school students. Psychology in the Schools, 8, 263-267.

Kanfer, F.H., Karoly, P., & Newman, A. (1975). Reduction of children's fear of the dark by competence-related and situational threat-related verbal cues. Journal of Consulting and Clinical Psychology, 43, 251-258.

Karlsruher, A.E. (1976). The influence of supervision and facilitative conditions on the psychotherapeutic effectiveness of nonprofessional and professional therapists. American Journal of Community Psychology, 4, 145-153.

Kavaler, S.I. (1974). The effects of dance therapy on mentally retarded children. Dissertation Abstracts International, 35(5-B), 2435. (University Microfilms No. 74-24, 640)

**Kellam, S.G., Branch, J.D., Agrawal, K.C., & Ensminger, M.E. (1975). Mental health and going to school: The Woodlawn program of assessment, early intervention, and evaluation. Chicago: University of Chicago Press.

Keller, M.F., & Carlson, P.M. (1974). The use of symbolic modeling to promote social skills in preschool children with low levels of social responsiveness.

Child Development, 45, 912-919.

Kelly, Jr., E.W., & Matthews, D.B. (1971). Group counseling with discipline-problem children at the elementary school level. The School Counselor, 18, 273-279.

Kennedy, C.D. (1981). Biofeedback and cognitive control in children with learning disabilities. (Doctoral dissertation, University of Washington, 1981). Dissertation Abstracts International, 42(1-B), 376-377.

Kern, R.M. (1971). The comparative effectiveness of a peer helper group counseling procedure and counselor oriented group counseling procedure on the adjustment of elementary school children. Dissertation Abstracts International, 31(8-A), 3877. (University Microfilms No. 71-4836)

Kern, R., & Kirby, J.H. (1971). Utilizing peer helper influence in group counseling. Elementary School Guidance and Counseling, 6, 70-73.

Kern, R.M., & Hankins, G. (1977). Adlerian group counseling with contracted homework. Elementary School Guidance and Counseling, 11, 284-291.

Kern, R.M., Kelley, J.D., & Downey, M. (1973). Group counseling versus halo consultation. Elementary School Guidance and Counseling, 7, 68-71.

Kilmann, P.R., Henry, S.E., Scarbro, H., & Laughlin, J.E. (1979). The impact of affective education on elementary school underachievers. Psychology in the Schools, 16, 217-223.

Klein, S.A., & Deffenbacher, J.L. (1977). Relaxation and exercise for hyperactive impulsive children. Perceptual and Motor Skills, 45, 1159-1162.

*Kolvin, I., Garside, R.F., Nicol, A.R., MacMillan, A., Wolstenholme, F., & Leitch, I.M. (1981). Help starts here: The maladjusted child in the ordinary school. New York: Tavistock.

Kondas, O. (1967). Reduction of examination anxiety and "stage-fright" by group desensitization and relaxation. Behavior Research and Therapy, 5, 275-281.

Korsh, N.B. (1977). Effects of interventions designed to improve cooperative social interaction and performance in small groups of third grade students in the open classroom. Dissertation Abstracts International, 37(10-B), 5438. (Xerox University Microfilms No. 77-6526)

Kranzler, G.D., Mayer, G.R., Dyer, C.O., & Munger, P.F. (1966). Counseling with elementary school children: An experimental study. Personnel and Guidance Journal, 44, 944-949.

Krauft, V.R. (1973). Transactional analysis group interactions with sixth grade behavioral problem boys. (Doctoral dissertation, University of Arkansas, 1973). Dissertation Abstracts International, 34(7-A), 3874-3875.

Ladd, G.W. (1981). Effectiveness of a social learning method for enhancing children's social interaction and peer acceptance. Child Development, 52, 171-178.

LaGreca, A.M., & Santogrossi, D.A. (1980). Social skills training with elementary school students: A behavioral group approach. Journal of Consulting and Clinical Psychology, 48, 220-227.

Lamia, M.C. (1977). A preventive and treatment program for adolescents: Psychological education. (Doctoral dissertation, California School of Professional Psychology -San Francisco, 1977). Dissertation Abstracts International, 38(4-B), 1888.

*Laxer, R.M., Quarter, J., Kooman, A., & Walker, K. (1969). Systematic desensitization and relaxation of high test-anxious secondary school students. Journal of Counseling Psychology, 16, 446-451.

Laxer, R.M., & Walker, K. (1970). Counterconditioning versus relaxation in the desensitization of test anxiety. Journal of Counseling Psychology, 17, 431-436.

Leal, L.L., Baxter, E.G., Martin, J., & Marx, R.W. (1981). Cognitive modification and systematic desensitization with test anxious high school students. Journal of Counseling Psychology, 28, 525-528.

Lewis, M.D. (1970a). Elementary school counseling and consultation: Their effects on teachers' perceptions. The School Counselor, 18, 49.

Lewis, M.D. (1970b). The effects of counseling and consultation upon the sociometric status and personal and social adjustment of third grade pupils. Elementary School Guidance and Counseling, 5, 44-52.

Liebroder, B.T. (1976). Evaluation of an elementary guidance program. (Doctoral dissertation, University of Utah, 1976). Dissertation Abstracts International, 37(6-A), 3422A.

Light, L.L., & Alexakos, C.E. (1970). Effect of individual and group counseling on study habits. The Journal of Educational Research, 63, 450.

Lilly, M.S. (1971). Improving social acceptance of low sociometric status, low achieving students. Exceptional Children, 37, 341-347.

Litrownik, A.J., Freitas, J.L., & Franzini, L.R. (1978). Self-regulation in mentally retarded children: Assessment and training of self-monitoring skills. American Journal of Mental Deficiency, 82, 499-506.

Little, S., & Jackson, B. (1974). The treatment of test anxiety through attentional and relaxation therapy. Psychotherapy: Theory, Research, and Practice, 11, 175-178.

Lott, S.J.B. (1975). A study of covert reinforcement as a treatment of test anxiety in male and female fifth graders. Dissertation Abstracts International, 36(4-A), 2028A. (Xerox University Microfilms No. 75-22, 510)

Mann, J. (1972). Vicarious desensitization of test anxiety through observation of videotaped treatment. Journal of Counseling Psychology, 19, 1-7.

Mann, J. & Rosenthal, T.L. (1969). Vicarious and direct counterconditioning of test anxiety through individual and group desensitization. Behavior Research and Therapy, 7, 359-367.

Marsh, D.T., Serafica, F.C., & Barenboim, C. (1980). Effect of perspective-taking training on interpersonal problem-solving. Child Development, 51, 140-145.

Mayer, G.R., Beggs, D.L., Fjellstedt, N., Forhetz, J., Nighswander, J.K., & Richards, R. (1970). The use of public commitment and counseling with elementary school children: An evaluation. Elementary School Guidance and Counseling, 5. 22-33.

McBrien, R.J., & Nelson, R.J. (1971). Experimental group strategies with primary grade children. Elementary School Guidance and Counseling, 6, 170-174.

McCollum, P.S. (1971). Group counseling as an adjunctive remediation technique for learning disabilities. (Doctoral dissertation, Texas Tech University, 1971). Dissertation Abstracts International, 32(4-B), 2404.

McCowan, R.J. (1968a). Group counseling underachievers and their parents. The School Counselor, 16, 30-41.

McCowan, R.J. (1968b). The effect of "brief contact" interviews with low-ability, low-achieving students. The School Counselor, 15, 386-389.

Meichenbaum, D.H., & Goodman, J. (1971). Training impulsive children to talk to themselves: A means of developing self-control. Journal of Abnormal Psychology, 77, 115-126.

Meredith, R., & Benninga, J.S. (1979). Counseling with parents to benefit children. Elementary School Guidance and Counseling, 14, 36-42.

Mestler, J.E. (1976). Behavioral changes of elementary students involved in the human development program: "Magic Circle". Journal of Clinical Child Psychology, 5, 18-20.

Mezzano, J. (1968). Group counseling with low-motivated male high school students--comparative effects of two

uses of counselor time. The Journal of Educational Research, 61, 222-224.

Michelson, L, & Wood, R. (1980). A group assertive training program for elementary schoolchildren. Child Behavior Therapy, 2, 1-9.

Miller, J.K. (1972). An investigation of group desensitization with test anxious seventh and eighth grade students. Dissertation Abstracts International, 32(8-A), 4354. (Xerox University Microfilms No. 72-7664)

Miller, N., & Kassinove, H. (1978). Effects of lecture, rehearsal, written homework, and IQ on the efficacy of a rational emotive school mental health program. Journal of Community Psychology, 6, 366-373.

Mink, E.B. (1979). The effectiveness of a group intervention based on transactional analysis in modifying locus of control expectancy of high school students. Dissertation Abstracts International, 40(4-A), 1968. (University Microfilms International No. 7921522)

Moates, H.L. (1969). The effects of activity group counseling on the self-concept, peer acceptance, and grade point average of disadvantaged seventh grade negro boys and girls. (Doctoral dissertation, Auburn University, 1969). Dissertation Abstracts

International, 30(9-A), 3795.

Moracco, J., & Kazandkian, A. (1977). Effectiveness of behavior counseling and consulting with non-western elementary school children. Elementary School Guidance and Counseling, 11, 244-251.

Moses, M.D. (1979). The effects of an Adlerian group counseling approach utilizing rehearsal with puppetry in decreasing the social isolation of withdrawn children. (Doctoral dissertation, Ohio University, 1979). Dissertation Abstracts International, 40(5-A), 2485.

Moulin, E.K. (1970). The effects of client-centered group counseling using play media on the intelligence, achievement, and psycholinguistic abilities of underachieving primary school children. Elementary School Guidance and Counseling, 5, 85-98.

Muller, S.D., & Madsen, Jr., C.H. (1970). Group desensitization for "anxious" children with reading problems. Psychology in the Schools, 7, 184-189.

Myrick, R.D., & Haight, D.A. (1972). Growth groups: An encounter with underachievers. The School Counselor, 20, 115-121.

Nagle, R.J., & Thwaite, B.C. (1979). Modeling effects on impulsivity with learning disabled children. Journal of Learning Disabilities, 12, 51-56.

Nauheim, J.L. (1980). The relative efficacies of anxiety management training, negative practice, and cognitive therapy in the treatment of test anxiety. (Doctoral dissertation, Hofstra University, 1980). Dissertation Abdtracts International, 42(1-B), 384.

O'Connor, R.D. (1969). Modification of social withdrawal through symbolic modeling. Journal of Applied Behavior Analysis, 2, 15-22.

O'Connor, R.D. (1972). Relative efficacy of modeling, shaping, and the combined procedures for modification of social withdrawal. Journal of Abnormal Psychology, 79, 327-334.

Oden, S., & Asher, S.R. (1977). Coaching children in social skills for friendship making. Child Development, 48, 495-506.

Omizo, M.M., & Michael, W.B. (1982). Biofeedback-induced relaxation training and impulsivity, attention to task, and locus of control among hyperactive boys. Journal of Learning Disabilities, 15, 414.

Parrish, J.M., & Erickson, M.T. (1981). A comparison of cognitive strategies in modifying the cognitive style of impulsive third-grade children. Cognitive Therapy and Research, 5, 71-84.

Perkins, J.A., & Wicas, E.A. (1971). Group counseling bright underachievers and their mothers. Journal of

Counseling Psychology, 18, 273-278.

Pihl, R.O., Parkes, M., Drake, H., & Vrana, F. (1980). The intervention of a modulator with learning disabled children. Journal of Clinical Psychology, 36, 972.

Platt, J.M. (1971). Efficacy of the Adlerian model in elementary school counseling. Elementary School Guidance and Counseling, 6, 86-91.

Poitras-Martin, D., & Stone, G.L. (1977). Psychological education: A skill-oriented approach. Journal of Counseling Psychology, 24(2), 153-157.

Ramage, R.L. (1979). A study of the effects of activity group counseling on the self-concepts, teachers' ratings of student behavior, and achievement of fourth and fifth grade students with learning disabilities. (Doctoral dissertation, University of Miami, 1979). Dissertation Abstracts International, 40(4-A), 1882.

Randolph, D.L., & Hardage, N.C. (1973). Behavioral consultation and group counseling with potential dropouts. Elementary School Guidance and Counseling, 7, 204-209.

Randolph, D.L., & Saba, R.G. (1973). Changing behavior through modeling and consultation. Elementary School Guidance and Counseling, 8, 98-106.

Randolph, D.L., & Wallin, K.R. (1973). A comparison of behavioral consultation and behavioral consultation

with model-reinforcement group counseling for children who are consistently off-task. The Journal of Educational Research, 67, 103-107.

Reckless, W.C., & Dinitz, S. (1972). The prevention of juvenile delinquency: An experiment. Columbus, OH: Ohio State University Press.

Redfering, D.L., & Bowman, M.J. (1981). Effects of a meditative- relaxation exercise on non-attending behaviors of behaviorally disturbed children. Journal of Clinical Child Psychology, 10, 126-127.

Ribordy, S.C., Tracy, R.J., & Bernotas, T.D. (1981). The effects of an attentional training procedure on the performance of high and low test-anxious children. Cognitive Therapy and Research, 5, 19-28.

Ridley, C.A., & Vaughn, S.R. (1982). Interpersonal problem solving: An intervention program for preschool children. Journal of Applied Developmental Psychology, 3, 177-190.

Ridley, C.A., Vaughn, S.R., & Wittman, S.K. (1982). Developing empathic skills: A model for preschool children. Child Study Journal, 12(2), 89-97.

Russell, M.L. & Roberts, M.S. (1979). Behaviorally-based decision-making training for children. Journal of School Psychology, 17, 264-269.

Sackles, J.A. (1981). An evaluation of three treatment

programs for anger control in young adolescents. (Doctoral dissertation, Hofstra University, 1980). Dissertation Abstracts International, 42(3-B), 1189.

Sage, C.K. (1980). Using self-instructional techniques to teach children positive assertion. Dissertation Abstracts International, 41(1-B), 367. (University Microfilms International No. 8013991)

Sarason, I.G., & Sarason, B.R. (1981). Teaching cognitive and social skills to high school students. Journal of Consulting and Clinical Psychology, 49, 908-919.

Sawyer, A.R. (1974). The effectiveness of token reinforcement, modeling, and traditional teaching techniques on achievement and self-concept of underachievers. (Doctoral dissertation, University of Northern Colorado, 1973). Dissertation Abstracts International, 34(7-A), 3888A.

Schinke, S.P., Blythe, B.J., & Gilchrist, L.D. (1981). Cognitive-behavioral prevention of adolescent pregnancy. Journal of Counseling Psychology, 28, 451-454.

Schinke, S.P., Blythe, B.J., Gilchrist, L.D., & Burt, G.A. (1981). Primary prevention of adolescent pregnancy. Social Work With Groups, 4, 121-135.

Schmieding, O. (1966). An investigation of efficacy of counseling and guidance procedures with failing junior

high students. The School Counselor, 14, 74-80.

Schumacher, R.R. (1974). The influence of differential group composition on the effectiveness of group counseling with second, third, fourth, and fifth grade male children. (Doctoral dissertation, University of Virginia, 1974). Dissertation Abstracts International, 35(4-A), 1992-1993.

Shearn, D.F., & Randolph, D.L. (1978). Effects of reality therapy methods applied in the classroom. Psychology in the Schools, 15, 79-83.

Smith, R.L., & Troth, W.A. (1975). Achievement motivation: A rational approach to psychological education. Journal of Counseling Psychology, 22, 500-504.

Smith, W. (1981). The effect of project turnabout on the self-concept and attitude toward school of low-achieving students. (Doctoral dissertation, University of Toledo, 1980). Dissertation Abstracts International, 42(1-A), 149.

Soper, E.T. (1978). The effects of a "think aloud" program on locus of control in young boys. (Doctoral dissertation, University of Utah, 1977). Dissertation Abstracts International, 38(7-A), 3966.

Stanley, S.F. (1978). Family education to enhance the moral atmosphere of the family and the moral development of adolescents. Journal of Counseling

Psychology, 25, 110-118.

Stilwell, W.E., Brown, P.W., & Barclay, J.R. (1973). Effects of three classroom management methods on classroom interaction of fifth graders. Psychology in the Schools, 10, 365-372.

Stone, G.L., & Noce, A. (1980). Cognitive training for young children: Expanding the counselor's role. The Personnel and Guidance Journal, 58, 416-420.

Stone, G.L., Hinds, W.C., & Schmidt, G. W. (1975). Teaching mental health behaviors to elementary school teachers. Professional Psychology, 6, 35-40.

Stormer, G.E. (1967). Milieu group counseling in elementary school guidance. Elementary School Guidance and Counseling, 1, 240-254.

Subotnik, L.S. (1975). Client-centered group therapy compared with behavior modification in changing inappropriate behavior of elementary school children. Psychotherapy and Psychosomatics, 25, 138-141.

Taylor, E.E. (1978). A comparison of the effects on the behavior of elementary school children with identifiable behavior and adjustment problems of counseling the children with that of counseling the mothers of the children. Dissertation Abstracts International, 38(7-A), 3966-3967A. (Xerox University Microfilms No. 77-29, 453)

Taylor, W.F., & Hoedt, K.C. (1974). Classroom-related behavior problems: Counsel parents, teachers, or children? Journal of Counseling Psychology, 21, 3-8.

Thombs, M.R., & Muro, J.J. (1973). Group counseling and the sociometric status of second grade children. Elementary School Guidance and Counseling, 7, 194-197.

Thompson, D.G., & Hudson, G.R. (1982). Values clarification and behavioral group counseling with ninth-grade boys in a residential school. Journal of Counseling Psychology, 29, 394-399.

Tosi, D.J., Swanson, C., & McLean, P. (1970). Group counseling with nonverbalizing elementary school children. Elementary School Guidance and Counseling, 4, 260-266.

Tosi, D.J., Upshaw, K., Lande, A., & Waldron, M.A. (1971). Group counseling with nonverbalizing elementary students: Differential effects of Premack and social reinforcement techniques. Journal of Counseling Psychology, 18, 437-440.

Tyler, F.B., & Gatz, M. (1977). Development of individual psychosocial competence in a high school setting. Journal of Counseling and Clinical Psychology, 45, 441-449.

Vogelsong, E.L. (1978). Relationship enhancement training

for children. Elementary School Guidance and Counseling, 12, 272.

Vogrin, D., & Kassinove, H. (1979). Effects of behavior rehearsal, audiotaped observation, and intelligence on assertiveness and adjustment in third-grade children. Psychology in the Schools, 16, 422-429.

Warner, R.W., & Hansen, J.C. (1970). Verbal-reinforcement and model-reinforcement group counseling with alienated students. Journal of Counseling Psychology, 17, 168-172.

Weissberg, R.P., Gesten, E.L., Rapkin, B.D., Cowen, E.L., Davidson, E., de Apodaca, R.F., & McKim, B.J. (1981). The evaluation of a social problem-solving training program for suburban and inner-city third grade children. Journal of Consulting and Clinical Psychology, 49, 251-261.

Weithorn, C.J., & Kagen, E. (1979). Training first graders of high-activity level to improve performance through verbal self-direction. Journal of Learning Disabilities, 12, 23-29.

Wen, S. (1974). Modifying children's impulsivity through redundancy and variability training. Psychology in the Schools, 11, 440-444.

Winkler, R.C., Teigland, J.J., Munger, P.F., & Kranzler, G.D. (1965). The effects of selected counseling and

remedial techniques on underachieving elementary school students. Journal of Counseling Psychology, 12, 384-387.

Wirth, S. (1977). Effects of a multifaceted reading program on self-concept. Elementary School Guidance and Counseling, 12, 33-40.

Yates, L.E.L. (1976). The use of sociometry as an identifier of research sample for psychological treatment and quantifier of change among second grade students. Group Psychotherapy and Psychodrama, 29, 102-110.

Zelie, K., Stone, C.I., & Lehr, E. (1980). Cognitive-behavioral intervention in school discipline: A preliminary study. Personnel and Guidance Journal, 59, 80-83.

*Two separate studies from these references were utilized in this analysis.

**Four separate studies from these references were utilized in this analysis.

APPENDIX B

APPENDIX B

Coding Scheme For Meta-analysis Of Psychotherapy With Children

I. Study Characteristics

1. Study ID# (001-999) (1-3)

2. Year of publication (code last two digits) (4-5)

3. Source (1-5) (6)
   1=published article
   2=book
   3=dissertation
   4=conference paper
   5=other

4. Total number of treatment groups (7-8)

5. Total number of comparisons (9-10)

6. Total number of outcome measures (11-12)

7. Follow-up data available (1-2) (13)

   1=yes
   2=no

II. Design Characteristics

8. Type of design (1-5)

   1=Pretest-Posttest with nonequivalent control group (NECG) (14)
   2=posttest only with NECG
   3=randomized true experiment
   4=other (e.g. matching)
   5=not available

9. Group assignment procedure (1-6) (15)

1=random
2=matching
3=available intact
4=voluntary self-selection
5=other
6=not available

10. Total sample size-assigned (16-18)
(all treatment groups and control groups)

11. Total sample size-completed posttest (19-21)
(all treatment groups and control groups)
Not ascertainable code 00

12. Overall quality code for this study (22)

III. Subject Information

13. Number of males in total sample (23-25)
Number unknown code 99

14. Mean age of subjects to the nearest tenth yr. (26-27)
Number unknown code 00

15. Ethnic sample characteristics (1-4) (28)

1=majority or all white
2=majority or all minority
3=mixed
4=unknown

16. Special sample characteristics (1-6) (29)

1=retarded
2=learning disabled
3=underachievers
4=other
5=unknown
6=none

17. Source of subjects (1-7) (30)

1=clinical inpatients
2=clinical outpatients seeking treatment
3=volunteers for special project
4=subjects chosen through problem-oriented observation, measurement, or recommendation
5=convenient 6=mixed/other 7=unknown

18. General seriousness of problem (1-4) (31)

1=none
2=mild
3=of uncertain nature/degree
4=moderate to severe

19. Target problem (1-15) (32-33)

1=social isolate
2=fears/phobias
3=anxiety
4=enuresis
5=somatic problems
6=depression
7=other or mix of 1-6

(1-7 indicate internalizing symptomatology)

8=impulsive/hyperactive
9=non-compliant/management problem/behavior problem
10=psychotic/autistic
11=other or mix of 8-10
12=social skills, undefined

(8-12 indicate externalizing symptomatology)

13=mix of 1-12
14=none
15=unknown

20. Academic learning problems (1-3) (34)

1=present
2=absent
3=unknown

IV. Therapist Characteristics

21. Number of therapists (35-36)
(code 0 for unknown)

22. Experience level of therapist (1-8) (37)

1=mental health professionals (PhD in Psychology, social work; MD in Psychiatry; school guidance counselor)
2=professional trainees (graduate students in psychology, interns, practicum students, psychiatric residents)
3=parents
4=teacher
5=other non-professionals
6="experimenter" 7=mixed 8=unknown

V. Comparison Information

23. Comparison Number (38-39)

24. Type of Comparison (1-4) (40)

1=treatment vs. control
2=behavioral vs. nonbehavioral
3=individual vs. group
4=combination

25. Type of Control Group (1-6) (41)

1=none
2=no treatment (assume if not stated)
3=wait-list
4=attention-placebo
5=other
6=not available

26. Sample size of treatment group for this comparison (42-44)

27. Sample size of control group for this comparison (45-47)

VI. Treatment Characteristics

28. Type of treatment (1-4) (48)

1=behavioral
2=nonbehavioral
3=mixed
4=unknown

29. Method of delivery (1-4) (49)

1=individual
2=group
3=mixed
4=unknown

30. Number of treatment sessions (50-51)

(code 0 for unknown)

31. Average length of treatment sessions in minutes (52-54)

(code 99 for unknown)

32. Treatment setting (1-9) (55)

1=school
2=home
3=mental health, community mental health or psychology/psychiatry clinic
4=general hospital or dental clinic
5=residential treatment center (psychiatric or special school)
6=camp
7=combination of at least two of the above
8=other
9=unknown

VII. Characteristics of Outcome Measures

33. Type of outcome measure (1-9) (56)

1=independent behavioral observation
2=nonindependent behavioral observation
3=peer sociometric
4=normed rating scale or behavioral checklist (or psychometrically adequate - someone else has used it before)
5=nonnormative/experimenter constructed instrument
6=achievement test or intellectual measure
7=other performance measure (e.g. MFF)
8=school grades
9=objective performance measure (e.g. days in school, arrests, approaching feared object)

34. Source of outcome measure (1-10) (57-58)

1=independent observers
2=parents
3=therapist
4=teachers/school
5=peers
6=subject self-report
7=subject performance measure (on an achievement, IQ, or cognitive measure)
8=other (expert judges, not independent observers, or therapists, or 1-7)
9=mixed
10=unknown

35. Dimension of adjustment (1-10) (59-60)

1=fear/anxiety
2=cognitive skills
3=global adjustment
4=social adjustment/social skills
5=achievement
6=personality
7=self-esteem
8=bed=wetting
9=mixed
10=unknown

36. Specific or generalized impact of treatment (1-2) (61)

1=specific
2=generalized

37. Type of adjustment or change measured (1-8) (62)

1=behavioral
2=personality
3=academic performance
4=sociometric
5=cognitive tempo
6=cognitive problem-solving skills
7=physiological measure
8=other

VIII. Effect Size Information

38. Reliability of measure (63)

(code 999 if not available)

39. Effect size at posttreatment (64-67)

40. Length of follow-up (in weeks) (68-70)

41. Effect size at follow-up (71-75)

42. How effect size was calculated (1-12) (76-77)

1=means/standard deviations
2=anova summary table
3=*t* score
4=raw data
5=ANCOVA
6=chi square/nonparametric
7=change scores
8=estimate from *p*
9=correlations
10=nonsignificant and no statistical info.
11=Holmes method
12=posttest adjustment

43. Source of data (1-3) (78)

1=standard information provided
2=data drawn from graphs
3=2-week test-retest reliabilities used with change scores

44. Number of this outcome measure (79-80)

APPROVAL SHEET

The thesis submitted by Jill Carmody has been read and approved by the following committee:

Dr. Joseph A. Durlak, Director
Professor, Psychology, Loyola

Dr. Fred Bryant
Associate Professor, Psychology, Loyola

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the Committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Arts.

April 17, 1985
Date

Joseph A. Durlak
Director's Signature